



Fakultät II – Informatik, Wirtschafts- und Rechtswissenschaften  
Department für Informatik

# **Constraint-Handling mit Supportvektor-Dekodern in der verteilten Optimierung**

Dissertation zur Erlangung des Grades eines  
Doktors der Naturwissenschaften

vorgelegt von

**Dipl.-Inform. Jörg Bremer**

Gutachter:

**Prof. Dr. Michael Sonnenschein**

**Jun.-Prof. Dr. Oliver Kramer**

**Prof. Dr. Wolfgang Renz**

Tag der Disputation: 4. Februar 2015



---

## Zusammenfassung

Die elektrische Energieversorgung befindet sich aktuell im Umbruch hin zu einem auf erneuerbaren Quellen aufbauenden, verteilten System. Ziel aktueller politischer Entscheidungen ist die schrittweise Transformation einer bisher eher zentralistisch orientierten elektrischen Energieversorgung durch wenige Großkraftwerke hin zu einem dezentralen Aufbau. Gleichzeitig muss ein Umdenken bezüglich der Einsatzplanung für die stetig steigende Zahl an elektrischen Erzeugern erfolgen. Eine zunehmende Dezentralisierung der Netzsteuerung bedingt durch die stetig wachsende Zahl kleiner regenerativer und insbesondere individuell konfigurierter Erzeuger führt zur Notwendigkeit dezentraler Algorithmen, um der wachsenden Komplexität entgegenzuwirken. Markterfordernisse lassen zudem die Entwicklung zu einem sich regelmäßig anpassenden System erwarten, bei dem sich die verteilten Anlagen selbsttätig je nach Marktlage produktorientiert zusammenfinden. Selbstorganisationsansätze werden daher aktuell als Möglichkeit diskutiert, der Komplexität und den Adaptionserfordernissen zu begegnen.

Diese Dynamik macht aber auch die Zusammenarbeit von verschiedenen sich unbekanntenen Anlagen erforderlich. Ohne zentrale Steuerungseinheit (mit zentraler, statischer Modellierung) müssen im Rahmen von Optimierungsansätzen die individuellen Suchräume mit den verschiedenen Möglichkeiten der beteiligten Anlagen ad hoc und automatisiert zu einem Modell des aktuell zu lösenden Optimierungsproblems integriert werden können.

Jede individuell betriebene Anlage hat einen ihr eigenen Satz an Möglichkeiten, den Verlauf der Erzeugung (oder auch des Verbrauchs) im Sinne einer Einsatzplanung anzupassen. Diese Flexibilität hängt ab von der individuellen, aktuellen Konfiguration, einer Reihe von (technischen) Nebenbedingungen, aktuellen Betriebsumständen sowie gegebenenfalls der aktuellen Einbindung in einen gekoppelten Anlagenkontext – bei einem Blockheizkraftwerk beispielsweise vom Wärmebedarf des nebenläufig zu heizenden Objekts.

Die Modellierung des Optimierungsproblems bei dem für jede Anlage einer Gruppe ein kostenoptimaler Fahrplan gesucht wird, sodass das Gesamtziel der Gruppe möglichst gut erreicht wird, kann nicht statisch vorgegeben werden. Vielmehr muss bedingt durch die dynamische Reorganisation der Gruppen auch die Modellierung des jeweiligen Optimierungsproblems einschließlich der Integration der individuellen Flexibilitäten nach jeder Gruppenbildung automatisiert erfolgen können.

Hierzu wird ein abstrakter Modellierungsansatz für die Flexibilitäten verschiedenster Anlagen sowie ihrer Nebenbedingungen nebst universeller Integration in Planungs- und Optimierungsalgorithmen benötigt. Entscheidend für die Akzeptanz eines solchen Planungsansatzes ist jedoch die Fähigkeit Lösungen zu finden, die von den jeweiligen Anlagen auch umgesetzt werden können. Hierfür ist eine detaillierte Kenntnis der Möglichkeiten einer jeden einzelnen Anlage unabdingbar. Diese müssen systematisch durchsucht und bewertet werden können. Hierfür einen durchgehenden Modellierungs- und Optimierungsansatz einschließlich eines Verfahrens zur systematischen Erzeugung zulässiger Lösungen zu entwickeln, ist Ziel dieser Arbeit.



---

## Abstract

The system for electrical power supply is currently undergoing a transition towards a decentralized provision from renewable resources. Political decisions target a change from traditional control schemes for rather few, large power plants to a power grid with decentralized control. At the same time, a solution has to be found for the growing number of (hardly predictable) energy generators. Grid control by decentralized and individually configured units leads to a need for new algorithms in order to cope with growing complexity. An integration of markets into the system leads to a steadily adapting system with units concurrently re-organizing into groups according to current market situation and desired commodities. Thus, approaches of self-organization are currently discussed as a prospective solution for a steadily adapting system.

This dynamics will lead to necessary co-operations of unacquainted energy units. Without central control (with a central, static model) individual search spaces of different units – representing the individual capabilities within a group – have to be integrated to a model for the optimization problem at runtime.

Each individually operated unit has its own set of different schedules to offer for a scheduling algorithm. This flexibility depends on the current, individual configuration of the unit, several (technical) constraints for operation, current operational state, and, if applicable, on state and requirements of coupled units – e. g. on the thermal demand of a house in case of a co-generation plant.

Modeling the optimization problem that determines a cost optimal operation schedule for each member of the group such that the group as a whole achieves the given targets cannot be attained a priori as a static model. Due to the dynamic nature of the problem and permanent re-organization, the model of the actual optimization problem (one for each group) has to be derived automatically right after forming a new group of electrical units. Individual flexibilities and constraints of these units have to be integrated on the fly.

Crucial for the acceptance of any optimization algorithm in the field of energy management and provision planning is the ability to produce feasible solutions with schedules that can be operated by the units afterwards. In this thesis, a new approach for flexibility modeling of arbitrary units on an abstract level is developed by using a support vector method and integrated with a decoder approach that guides algorithms where to look for feasible solutions and how to evaluate them.



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Hintergrund . . . . .	1
1.2	Problemstellung . . . . .	4
1.3	Lösungsansatz . . . . .	5
1.4	Aufbau der Arbeit . . . . .	7
<b>I</b>	<b>Grundlagen</b>	<b>9</b>
<b>2</b>	<b>Problemstellung im Kontext Smart Grid</b>	<b>11</b>
2.1	Einleitung . . . . .	11
2.2	Wandel der Energieversorgung . . . . .	13
2.3	Koordinationsaufgaben im Verteilnetz . . . . .	15
2.4	Problemstellung der verteilten Lastplanung in dynamischen Verbänden . . . . .	19
2.5	Zusammenfassung . . . . .	22
<b>3</b>	<b>Verwandte Arbeiten</b>	<b>23</b>
3.1	Einleitung . . . . .	23
3.2	Optimierte Wirkleistungsbereitstellung . . . . .	24
3.2.1	Unit Commitment . . . . .	25
3.2.2	Virtuelle Kraftwerke . . . . .	27
3.2.3	Verteilte Ansätze . . . . .	29
3.3	Kritische Diskussion verwandter Arbeiten . . . . .	31
3.4	Zusammenfassung . . . . .	33
<b>4</b>	<b>Constraint-basierte Optimierung</b>	<b>35</b>
4.1	Einleitung . . . . .	35
4.2	Grundlagen . . . . .	35
4.3	Evolutionäre Lösungsansätze . . . . .	37
4.4	Constraints im Kontext Smart Grid . . . . .	40
4.4.1	Hard Constraints . . . . .	40
4.4.2	Soft Constraints . . . . .	41
4.4.3	Globale Constraints . . . . .	42
4.5	Constraintbehandlung während der Optimierung . . . . .	42
4.5.1	Lagrange'sche Multiplikatoren und die Karush-Kuhn-Tucker-Bedingungen . . . . .	45
4.5.2	Mehrzielansätze . . . . .	47
4.5.3	Penaltyansatz . . . . .	47
4.5.4	Dekoderansatz . . . . .	48

4.5.5	Spacemapping und Surrogatmodelle . . . . .	50
4.6	Zusammenfassung . . . . .	51
<b>5</b>	<b>Supportvektor Maschinen und Kern-basierte Verfahren</b>	<b>53</b>
5.1	Einleitung . . . . .	53
5.2	Supportvektor Maschinen . . . . .	54
5.2.1	Grundprinzip . . . . .	55
5.2.2	Lösungsverfahren . . . . .	57
5.2.3	Anwendungen Kern-basierter Verfahren . . . . .	58
5.3	Supportvektor Data Description . . . . .	60
5.3.1	Grundprinzip . . . . .	61
5.3.2	Geometrische Interpretation . . . . .	64
5.3.3	Lösung des Optimierungsproblems zur Bestimmung der Hypersphäre	66
5.3.4	Parametrierung . . . . .	68
5.3.5	Anwendungen für die SVDD . . . . .	70
5.4	Zusammenfassung . . . . .	72
<b>II</b>	<b>Problemformalisierung &amp; Lösungsansatz</b>	<b>73</b>
<b>6</b>	<b>Formalisierung und Modellierung der Wirkleistungsplanung im Smart Grid</b>	<b>75</b>
6.1	Einleitung . . . . .	75
6.2	Beteiligte Akteure . . . . .	75
6.3	Modellierung . . . . .	77
6.3.1	Fahrplanmodellierung und Wirkleistungsprodukt . . . . .	78
6.3.2	Anlagenmodellierung und Randbedingungen . . . . .	80
6.4	Kennzahlen zur Lösungsbewertung . . . . .	84
6.5	Problemdefinition . . . . .	86
6.6	Zusammenfassung . . . . .	89
<b>7</b>	<b>Lösungskonzept</b>	<b>91</b>
7.1	Einleitung . . . . .	91
7.2	Lösungskonzept . . . . .	92
7.2.1	Suchraummodell . . . . .	94
7.2.2	Systematische Lösungsgenerierung . . . . .	95
7.2.3	Prozesskette . . . . .	95
7.3	Voraussetzungen an die Anlagen . . . . .	96
7.4	Zusammenfassung . . . . .	98
<b>8</b>	<b>SVDD als Blackbox-Modell</b>	<b>101</b>
8.1	Einleitung . . . . .	101
8.2	Erstellen einer Trainingsmenge . . . . .	102
8.3	Modell . . . . .	110
8.3.1	Suchraummodell . . . . .	111
8.3.2	Illustration . . . . .	112
8.3.3	Einbettung und Anwendbarkeit . . . . .	114
8.4	Zusammenfassung . . . . .	116

<b>9</b>	<b>Supportvektor-Dekoderansatz</b>	<b>117</b>
9.1	Einleitung . . . . .	117
9.2	Problemstellung und Lösungsidee . . . . .	117
9.3	Dekoderabbildung auf Basis der empirischen Kernabbildung . . . . .	119
9.3.1	Voraussetzungen . . . . .	120
9.3.2	Schritt 1: Abbildung in den Supportvektor RKHS $\mathcal{H}^{(\ell)}$ mittels empirischer Kernabbildung . . . . .	121
9.3.3	Schritt 2: Rejustierung im RKHS . . . . .	122
9.3.4	Schritt 3: Urbildbestimmung . . . . .	124
9.3.5	Illustration . . . . .	125
9.4	Integration in Optimierung . . . . .	128
9.5	Zusammenfassung . . . . .	131
<b>10</b>	<b>Modellierung von Kosten und Kennzahlen</b>	<b>133</b>
10.1	Einleitung . . . . .	133
10.2	Integrationsansatz . . . . .	134
10.3	Verwendete Kennzahlen . . . . .	135
10.3.1	Produktionskosten . . . . .	135
10.3.2	Umweltkennzahlen . . . . .	137
10.3.3	Erhalt von Freiheitsgraden . . . . .	139
10.4	Simulationsergebnisse . . . . .	140
10.5	Zusammenfassung . . . . .	143
<b>III</b>	<b>Anwendung und Evaluation</b>	<b>145</b>
<b>11</b>	<b>Evaluation</b>	<b>147</b>
11.1	Einleitung . . . . .	147
11.1.1	Strukturanalyse . . . . .	148
11.1.2	Informationsanalyse . . . . .	150
11.2	Verwendete Evaluationsansätze . . . . .	152
11.3	Modellierung der Szenarien . . . . .	153
11.3.1	Szenarien . . . . .	153
11.3.2	Anlagenmodelle . . . . .	154
11.3.3	Relaxation diskreter Fahrplanräume . . . . .	159
11.4	Komponententests . . . . .	160
11.4.1	Klassifikator . . . . .	161
11.4.2	Dekoder . . . . .	168
11.4.3	Kennzahlen für die Mehrzieloptimierung . . . . .	174
11.5	Zusammenfassung . . . . .	176
<b>12</b>	<b>Umsetzung mit zentralen Optimierungsansätzen</b>	<b>177</b>
12.1	Einleitung . . . . .	177
12.2	Verwendete Verfahren . . . . .	178
12.3	Abbildungskonzept . . . . .	180
12.4	Simulative Evaluation . . . . .	182
12.4.1	Ergebnisse . . . . .	183
12.4.2	Vergleich mit Penalty-Ansatz . . . . .	191

12.4.3 Die Rolle von Batteriespeichern . . . . .	194
12.5 Zusammenfassung . . . . .	196
<b>13 Umsetzung mit dezentralen Optimierungsansätzen</b>	<b>199</b>
13.1 Einleitung . . . . .	199
13.2 Verfahren . . . . .	199
13.2.1 Ein verteilter Mehrziel-Greedy-Ansatz . . . . .	200
13.2.2 COHDA . . . . .	205
13.3 Zusammenfassung . . . . .	207
<b>14 Integration von Prognoseunsicherheiten für eine robuste Planung</b>	<b>209</b>
14.1 Einleitung . . . . .	209
14.2 Erweiterung des Modells . . . . .	209
14.3 Simulative Evaluation . . . . .	211
14.4 Fazit . . . . .	215
<b>15 Faire Gewinnverteilung</b>	<b>217</b>
15.1 Einleitung . . . . .	217
15.2 Lösungsansatz . . . . .	218
15.2.1 Shapley-Werte in der Wirkleistungsplanung . . . . .	219
15.2.2 Bestimmung der Verteilungskoeffizienten . . . . .	220
15.3 Zusammenfassung . . . . .	223
<b>16 Interoperabilität</b>	<b>225</b>
16.1 Einleitung . . . . .	225
16.2 Integration in bestehende Standards . . . . .	225
16.2.1 61850 . . . . .	226
16.2.2 Common Information Model . . . . .	227
16.2.3 OPC-UA . . . . .	228
16.3 Vergleichbarkeit der Kostenkennzahlen . . . . .	228
16.4 Zusammenfassung . . . . .	229
<b>17 Generalisierung des Ansatzes</b>	<b>231</b>
17.1 Einleitung . . . . .	231
17.2 Erweiterungen des Ansatzes . . . . .	231
17.2.1 Einsatz für komplexe Leistungswerte . . . . .	231
17.2.2 Operationen auf dem Suchraum . . . . .	233
17.2.3 Verbessertes Sampling . . . . .	237
17.3 Generalisierung des Ansatzes . . . . .	239
17.4 Grenzen des Ansatzes . . . . .	239
17.5 Zusammenfassung . . . . .	242
<b>18 Fazit und Ausblick</b>	<b>243</b>
18.1 Zusammenfassung . . . . .	243
18.2 Fazit . . . . .	245
<b>Literatur</b>	<b>247</b>

# Abbildungsverzeichnis

2.1	Entwicklung des Beitrags erneuerbarer Energieträger in Deutschland . . .	12
2.2	Potenziale für Regelleistung durch DSM-Maßnahmen im Haushaltssektor; nach (Dena 2010), verändert. . . . .	15
4.1	Lösungsraum der gültigen Lösungen $\mathcal{F}$ als Unterraum des durch die Para- meter aufgespannten Suchraums $\mathcal{S}$ . . . . .	37
4.2	Vereinfachtes, lineares Dekoderbeispiel. . . . .	49
4.3	Dekoderansatz nach (Koziel & Michalewicz 1999) . . . . .	50
5.1	Lage der möglichen Diagonalsegmente bei der SMO . . . . .	58
5.2	Geometrische Zusammenhänge in $\mathcal{H}$ und Lage der Hyperkugel, nach (Kim et al. 2008), verändert. . . . .	65
5.3	Abbilden einer Trainingsmenge mittels SVDD . . . . .	65
5.4	Bedeutung der Bandbreite für die Klassengrenze . . . . .	69
5.5	Bedeutung der Bandbreite für nicht zusammenhängende Teilregionen . . .	71
6.1	Vereinfachendes, 2-dimensionales Beispiel für den Suchraum eines modu- lierenden BHKW mit drei exemplarischen Constraints. . . . .	81
6.2	Vereinfachtes, 2-dimensionales Beispiel für einen nicht konvexen Lösungs- raum für ein BHKW (verbleibende weiße Fläche). . . . .	82
7.1	Prozesskette für die dynamische Integration von Flexibilitäten in die (ver- teilte) Optimierung mittels Dekoder. . . . .	96
8.1	Abhängigkeiten der Realisierbarkeit von der Fahrplandauer . . . . .	103
8.2	Ablaufdiagramm für das sukzessive Sampling von Anlagensimulationsmo- dellen (Bremer & Sonnenschein 2013e). . . . .	106
8.3	Ausgleich der selektiven Stichprobe durch Brown'sche Bewegungen. . . . .	107
8.4	Temperaturverläufe in einem Warmwasseruntertischgerät . . . . .	108
8.5	Histogramme empirischer Verteilungen . . . . .	109
8.6	Vorgehen bei der Erstellung des Modells mittels SVDD. . . . .	111
8.7	Schematische Darstellung des Klassifikationsprozesses. . . . .	113
8.8	Illustratives Beispiel zweier Suchraummodelle. . . . .	114
8.9	Reparatur von Löchern im Suchraummodell. . . . .	115
8.10	Vorläufiges rein Klassifikator-basiertes Integrationsmodell . . . . .	116
9.1	Prinzipskizze des Supportvektor-Dekoderansatzes. . . . .	119
9.2	Dekoderausgabe für das artifizielle Beispiel eines Doppelrings. . . . .	125
9.3	Beispiel einer Dekoderabbildung am artifiziiellen Beispiel. . . . .	126

9.4	Beispiele für hochdimensionale Dekoderabbildungen . . . . .	127
9.5	Darstellung des Skalierungseffektes bei der Rejustierung im RKHS . . . . .	128
9.6	Dekoderabbildung eines regelmäßigen Gitters . . . . .	129
9.7	Ein erstes, artifizien Dekoder-Optimierungsbeispiel . . . . .	130
10.1	Beispielhafter Effizienzverlauf eines BHKW . . . . .	137
10.2	Vergleich der Suchraumgrößen für BHKW . . . . .	139
10.3	Vergleich der Suchraumgrößen für Redox-Flow-Batterien . . . . .	140
10.4	Indikatornäherungsfunktion für die Bewertung des Batterieendzustandes. . . . .	140
10.5	Rekonstruktion einer artifizien Kennzahl . . . . .	141
10.6	Beispiele für die Rekonstruktion artifizien Kennzahlen. . . . .	142
10.7	Durch Kennzahlenintegration zusätzlich benötigte Supportvektoren. . . . .	142
11.1	Korrelationscharakteristik der Fitnesslandschaften . . . . .	149
11.2	Modell und Modellkomponenten für das verwendete BHKW-Modell. . . . .	155
11.3	Exemplarische elektrische Fahrpläne für ein Wärmepumpenmodell zur Deckung des thermischen Profils aus Abbildung 11.4. . . . .	157
11.4	Thermische Heizleistung zu elektrischen BHKW-Fahrplänen . . . . .	157
11.5	Vergleich der Fahrpläne bei der Relaxation zu kontinuierlichen, gemittelten Leistungswerten am Beispiel einer Wärmepumpe. . . . .	159
11.6	Zusammenhang zwischen Samplegröße und Korrektklassifikationsrate . . . . .	164
11.7	Samplegröße für Batteriespeicher . . . . .	165
11.8	Anzahl Supportvektoren für ein BHKW . . . . .	166
11.9	Anzahl Supportvektoren für Wärmepumpe und Boiler . . . . .	167
11.10	Visueller Vergleich von Simulationsmodell und Suchraummodell für hochdi- mensionale Fahrpläne. . . . .	168
11.11	Vergleich der Modelle für eine Wärmepumpe. . . . .	170
11.12	Vergleich von Simulationsmodell und Dekoder für ein BHKW . . . . .	171
11.13	Hausdorffabstand zwischen Wärmepumpenmodell und Dekoder . . . . .	172
11.14	Ausschnittvergrößerung mit resultierenden Fahrplanmengen (durch den jeweils resultierenden Dekoder zufällig generiert). . . . .	173
11.15	Vergleich von Hausdorffabstand und Korrektklassifikationsraten . . . . .	174
12.1	Genereller Ablauf für Simulated Annealing mit Dekoder. . . . .	180
12.2	Beispiel für Pfadabbildung . . . . .	181
12.3	Zweites Beispiel für Pfadabbildung . . . . .	182
12.4	Konvergenzverhalten der Standardverfahren . . . . .	185
12.5	Beispielergbnis mit Simulated Annealing 1 . . . . .	186
12.6	Beispielergbnis mit Simulated Annealing 2 . . . . .	187
12.7	Konvergenzverhalten des Simulated Annealing . . . . .	190
12.8	Beispielergbnis mit Anlagenconstraints . . . . .	192
12.9	Rolle der Batteriespeicher für die Einsatzplanoptimierung . . . . .	195
13.1	Aktualisierungsschritt eines verteilten Greedy-Verfahrens . . . . .	201
13.2	Grundprinzip des Greedy-Verfahrens . . . . .	201
13.3	Beispielergbnis für Mehrzieloptimierung 1. . . . .	203
13.4	Beispielergbnis für Mehrzieloptimierung 2. . . . .	204
13.5	Beispielergbnis aus der Integration von Dekoder und COHDA. . . . .	206

14.1	Vergleich der Entscheidungsgrenzen an zwei Beispielen mit unterschiedlicher Prognoseabweichung. . . . .	213
15.1	Zeitkomplexität bei der verteilten Bestimmung von Shapley-Werten . . .	222
16.1	Minimalbeispiel für ein Suchraummodell in IEC61970 CIM. . . . .	227
17.1	P-Q-Leistungsfläche der Synchronmaschine . . . . .	232
17.2	Menge von Fahrplänen mit gleichem Anlagenendzustand am Beispiel eines BHKW. . . . .	234
17.3	Verteilung kombinierter BHKW-Trainingsmengen für verschiedene Anlagenzahlen. . . . .	236
17.4	Beispiel für kombinierte Suchräume zweier BHKW . . . . .	237
17.5	Beispiel produktspezifischen Samplings . . . . .	238
17.6	Beispielfahrpläne für ein BHKW für einen ganzen Tag in Minutenauflösung.	240
18.1	Zusammenfassung der erreichten Ziele. . . . .	245



# Tabellenverzeichnis

8.1	Exemplarische Größe verschiedener Lösungsräume am Beispiel von BHKW und Untertischwarmwassergerät . . . . .	104
11.1	Komplexitätsbetrachtungen für drei ausgewählte Problemszenarien. . . .	151
11.2	Vergleich verschiedener Klassifikatorkennzahlen für ein Untertischwarmwassergerät . . . . .	162
11.3	Vergleich verschiedener Klassifikatorkennzahlen für eine exemplarische Wärmepumpe . . . . .	162
11.4	Vergleich verschiedener Klassifikatorkennzahlen für einen fast vollständig geladenen Batteriespeicher . . . . .	163
11.5	Korrektklassifikationsrate für höher dimensionale Fahrpläne verschiedener Anlagen . . . . .	164
11.6	Größe des Fehler bei der Kennzahlenrekonstruktion durch den Dekoder am Beispiel dreier artifizierlicher Kennzahlen. . . . .	176
12.1	Ergebnisse für die Optimierung mit ABC für Verbände mit 10 Anlagen und einer Produktlänge von 8 Zeitintervallen. . . . .	183
12.2	Ergebnisse für die Optimierung mit PSO für Verbände mit 10 Anlagen und einer Produktlänge von 8 Zeitintervallen. . . . .	184
12.3	Ergebnisse für die Optimierung mit HS für Verbände mit 10 Anlagen und einer Produktlänge von 8 Zeitintervallen. . . . .	184
12.4	Simulationsergebnisse für Simulated Annealing 1 . . . . .	188
12.5	Simulationsergebnisse für Simulated Annealing 2 . . . . .	188
12.6	Simulationsergebnisse für Simulated Annealing 3 . . . . .	188
12.7	Simulationsergebnisse für Simulated Annealing 4 . . . . .	189
12.8	Simulationsergebnisse für Simulated Annealing 5 . . . . .	189
12.9	Simulationsergebnisse für Simulated Annealing 6 . . . . .	189
12.10	Gültigkeit der Lösungen . . . . .	191
12.11	Gültigkeit der Lösungen für 96 Zeitintervalle . . . . .	191
12.12	Ergebnisse für die Optimierung mittels Penalty-Ansatz für verschiedene Gewichtungsverhältnisse . . . . .	193
13.1	Ergebnisse für den verteilten Greedyalgorithmus für Verbände mit 10 Anlagen. . . . .	202
13.2	Ergebnisse für den verteilten Greedyalgorithmus für Verbände mit 100 Anlagen. . . . .	202
13.3	Ergebnisse für den verteilten Greedyalgorithmus für Verbände mit 1000 Anlagen. . . . .	203

14.1	Vergleich der Klassifikatoren für einen Warmwasserboiler mit verschiedenen (prognostizierten) Zapfprofilen. . . . .	212
14.2	Vergleich der Dekoderfehler (Anteil falsch rekonstruierter Fahrpläne) für verschieden große mittlere Prognoseabweichungen am Beispiel eines Boilers.	213
14.3	Vergleich der Trefferquoten für weitere Szenarien mit unsicheren Prognosen.	214
14.4	Vergleich der Korrektklassifikationsraten bei der Klassifikation von Warmwasserbereitermodellen. . . . .	214
14.5	Vergleich der Korrektklassifikationsraten für Warmwasserbereiter für 24-dimensionale Fahrpläne . . . . .	214
15.1	Verteilungsschlüssel für Erzeugerverbünde mit gleichartigen BHKW . . . .	222
17.1	Verschiedene Rechenzeiten für die Dekodererstellung . . . . .	241
17.2	Verschiedene Rechenzeiten für die Dekoderbenutzung . . . . .	241

# Verwendete Formelzeichen und Abkürzungen

$\alpha_i, \beta_i$ .....	Lagrange-Multiplikatoren
$C$ .....	Constraint
$d_H(A, B)$ .....	Gerichteter Hausdorffabstand
$\widetilde{d}_H^{(n)}$ .....	Diskreter, fehlerbereinigter Hausdorffabstand
$\mathcal{E}$ .....	Einsatzplan
$\mathcal{F}$ .....	Lösungsraum
$\mathcal{F}_{[0,1]}$ .....	Skalierter Lösungsraum
$\mathcal{F}[z_0]$ .....	Lösungsraum in Abhängigkeit vom Startzustand $z_0$
$\mathcal{H}$ .....	Hilbertraum
$\mathcal{I}^{(Q)}$ .....	Indikatorabbildung zur Kenngröße $Q$
$\mathcal{I}^{(Q)}$ .....	Funktionale Beschreibung von Kenngröße $Q$
$k$ .....	Kern
$K$ .....	Kern-Gram-Matrix
$\mathcal{L}$ .....	Fitnesslandschaft
$\mathbf{p}, \mathbf{p}^{(r)}$ .....	Fahrplan bzw. Wirkleistungsverlauf
$\mathbf{p}_{[0,1]}$ .....	Skalierter Fahrplan
$P, \zeta^{(P)}$ .....	Wirkleistungsprodukt, Ziellastgang $\zeta^{(P)} \in \mathbb{R}^d$
$\varphi$ .....	Shapley-Wert
$\Phi$ .....	Feature Map
$\Psi_x$ .....	Bild von $x$ in $\mathcal{H}$
$Q$ .....	Kenngröße
$R(x), R_S$ .....	Abstand des Bildes von $x$ vom Zentrum der Hyperkugel in $\mathcal{H}$ und Radius der Hyperkugel
$S$ .....	Freier Suchraum
$\mathcal{S}$ .....	Hypersphäre zur Klassifikation von Lastgängen
$\sigma$ .....	Standardabweichung bzw. Bandbreite des Gaußkerns
$\Sigma$ .....	freier Suchraum
$SV$ .....	Menge von Supportvektoren $SV = s_i$
$U$ .....	Einheit
$\mathcal{V}$ .....	Verbund (Koalition aus Agenten)
$w$ .....	Gewichtungsvektor für SVDD-Klassifikator
$x \sim U(0, 1)$ .....	Gleichverteilte Zufallsgröße aus $[0, 1]$

## Verwendete Formelzeichen und Abkürzungen

---

$\mathcal{X}$ .....	Stichprobe von gültigen Lastgängen
$\xi_i$ .....	Schlupfvariable
$Z_U$ .....	Menge der Betriebszustände von Einheit $u$
$z_0$ .....	Startzustand
$\langle \cdot, \cdot \rangle$ .....	Skalarprodukt
$2^N$ .....	Potenzmenge der Menge $N$
$x_1 \succ x_2$ .....	Relation $x_1$ besser als $x_2$
$x_1 \succeq x_2$ .....	Relation $x_1$ besser als oder gleich gut wie $x_2$
ABC .....	Artificial Bee Colony
BHKW .....	Blockheizkraftwerk
EA .....	Evolutionärer Algorithmus
EEG .....	Erneuerbare-Energien-Gesetz
HS .....	Harmony Search
iid .....	independent and identical distributed
KKT .....	Karush-Kuhn-Tucker (Bedingungen)
KWK .....	Kraft-Wärme-Kopplung
PSO .....	Partikelschwarmoptimierung
PV .....	Photovoltaik
RKHS .....	Reproducing Kernel Hilbert Space
SVDD .....	Supportvektor Data Description
SVM .....	Supportvektor-Maschine
VK .....	Virtuelles Kraftwerk

# 1. Einleitung

*Wer das Breichen mit Knoblauch würzt,  
findet sein Baby auch im Dunkeln.*

---

*(Fritz-J. Schaarschuh)*

## 1.1. Hintergrund

Eine schrittweise Transformation von den heutigen, zentralen Markt- und Netzstrukturen in der Energieversorgung hin zu einem dezentralen Aufbau ist nicht nur Ziel aktueller politischer Entscheidungen (European Parliament & Council 2009; Deutsche Bundesregierung (Bundeskabinett) 2010) sondern auch dringende Notwendigkeit für eine erfolgreiche Integration regenerativer Stromerzeugung (Ramchurn et al. 2012; Buchholz et al. 2008; Consentec GmbH & r2b energy consulting GmbH 2010). Bisher erfolgt die Erzeugung der täglich benötigten elektrischen Energie durch wenige Großkraftwerke, für die dafür jeweils Einsatzpläne basierend auf dem erwarteten Verbrauch erstellt werden. Dieses Schema ist einem rasanten Wandel unterlegen, da täglich mehr und mehr kleine, regenerative Erzeuger wie Blockheizkraftwerke (BHKW), Windenergieanlagen, Biomassekraftwerke oder Photovoltaik-Anlagen ans Netz gehen (Sonnenschein et al. 2011). Wenn diese nach und nach die Erzeugungsleistung übernehmen, so muss auch ein Weg gefunden werden, diese Anlagen in die jeweiligen Planungsprozesse für die Netzführung einzubeziehen. Dann aber würde im Falle der Beibehaltung etablierter Planungsmethoden die wachsende Problemgröße durch die schiere Menge der zu verplanenden Anlagen ein Berechenbarkeitsproblem verursachen (Saravanan et al. 2013).

Die Bündelung der Potenziale kleiner Anlagen durch einen (informationstechnischen) Zusammenschluss zu Gruppen, die dann als Einheiten mit hinreichend Leistungsvermögen betrachtet werden können, stellt einen aktuell intensiv betrachteten Forschungsgegenstand dar (Lukovic et al. 2010; Bitsch et al. 2002). Man spricht hier auch von Virtuellen Kraftwerken (Awerbuch & Preston 1997). Ein entscheidender Faktor für die Bündelung wird die Ausbildung von Verbänden von verteilten Erzeugern und Verbrauchern auf allen Stufen des Netzes sein (Nieße et al. 2012a). Verbundbildung bedeutet in diesem Zusammenhang insbesondere, dass eine Gruppe von mehreren Anlagen eine Aufgabe gemeinsam übernimmt, wenn einzelne Anlagen alleine keine ausreichende Dimensionierung besitzen oder aber allein nicht die notwendige Flexibilität bei der Steuerung aufweisen. Dies ist z. B. der Fall, wenn einzelne Anlagen allein zu wenig Nennleistung aufweisen oder nicht genügend Freiheitsgrade für eine Erzeugungsanpassung aufweisen (Tröschel 2010). Zunächst wird diese Verbundbildung noch zur Überwindung von Markteintrittsbarrieren durch Bündelung von Ressourcen erfolgen. Im weiteren Verlauf der Transformation werden diese Verbünde jedoch in Zusammensetzung und Zielrichtung ebenfalls einem Wandel hin zu kleineren,

## 1. Einleitung

---

dafür aber stärker stochastischen und fluktuierenden Einheiten unterzogen sein, da sie beispielsweise vermehrt Photovoltaikanlagen o. ä. enthalten werden. Sowohl diese Volatilität bei den fluktuierenden Einspeisern als auch weitere Einflussgrößen wie sich ändernde Marktsituationen oder Netzanforderungen lassen erwarten, dass Verbünde über die Zeit nicht stabil in ihrer Zusammensetzung sind sondern sich regelmäßig neu ausbilden, um den aktuellen Erfordernissen gerecht zu werden (Nieße et al. 2014). Solche Verbünde werden nach und nach die Aufgaben der heutigen Großkraftwerke aus allen Bereichen der Netzsteuerung übernehmen müssen (Nieße et al. 2014). Eine dieser Aufgaben wird die gemeinschaftliche Bereitstellung eines geplanten Wirkleistungslastgangs zur Deckung einer antizipierten Nachfrage beispielsweise für den Folgetag sein (Kraftwerkseinsatz- oder Wirkleistungsplanung).

In der klassischen Wirkleistungsplanung (Saravanan et al. 2013; Franch et al. 2009) wird – meist für den Folgetag – der Einsatz von zur Verfügung stehenden elektrischen Erzeugern geplant. Für jeden Generator wird ein Fahrplan bestimmt, sodass zum einen der prognostizierte Verbrauch gedeckt ist und zum anderen die Betriebsmittel möglichst kostenoptimal eingesetzt werden. Damit Verbünde von Anlagen nach ihrer Bildung diese Aufgabe – die Bereitstellung des gewünschten Wirkleistungslastgangs – erfüllen können, werden Verfahren benötigt, die Last auf die einzelnen Teilnehmer so zu verteilen, dass einerseits der Summenlastgang der einzelnen Anlagen den gewünschten Ziellastgang möglichst gut annähert und andererseits aber auch globale (z. B. Minimierung nachteiliger Umwelteinflüsse) sowie insbesondere auch lokale (z. B. Maximierung des individuellen Ertrags) Optimierungspotenziale genutzt werden. Es ist also je Anlage ein Fahrplan gesucht, sodass die gemeinsame Aufgabe nach gegebenen Optimalitätskriterien möglichst gut gelöst wird (Hinrichs et al. 2013a; Bremer & Sonnenschein 2012a; Bremer 2012; Bremer & Sonnenschein 2013b).

Die besonderen Schwierigkeiten hierbei ist die Berücksichtigung der individuellen Fähigkeiten der verschiedenen Anlagen. Jede Anlage hat ihre eigenen meist technisch bedingten Einschränkungen unter denen sie betrieben werden kann. Beispiele hierfür sind Beschränkungen des Leistungsbereichs, Speicherfüllstände oder einzuhaltende Vorgaben etwa zur nebenläufigen Wärmeerzeugung in gekoppelten Erzeugern (Sonnenschein et al. 2013). Aber auch wirtschaftliche oder umwelttechnische Gründe können gewisse Betriebsweisen zumindest unerwünscht machen, wenn diese zu teuer oder ineffizient sind (Bremer & Sonnenschein 2013d). Solche Einschränkungen – oder genauer: ihre Auswirkungen – sind nicht statisch, sondern abhängig vom aktuellen Betriebszustand, aktuellen Prognosen aber möglicherweise auch von Lieferverpflichtungen, die einen Teil der Anlagenkapazität bereits für andere Zwecke belegen. Der Betriebszustand wiederum kann auch ganze Anlagenkomplexe umfassen wenn beispielsweise die fragliche Anlage in eine größere Produktionsanlage o. ä. eingebettet ist (Sudeikat & Heitmann 2013). Die Nebenbedingungen und somit die Möglichkeiten die der Anlage verbleiben, ihr Leistungsprofil (Verlauf von Stromerzeugung oder auch -verbrauch) für einen gegebenen Zeitraum zu variieren, sind also für jeden neuen Optimierungsvorgang individuell zu modellieren bzw. zu bestimmen (Bremer & Sonnenschein 2013f). Man spricht hier auch von der Flexibilität einer Anlage. Waren die Nebenbedingungen in der Vergangenheit aufgrund der relativen Gleichartigkeit der Generatoren noch eher einfach (und auch zentral) zu modellieren (Saravanan et al. 2013), so entsteht durch die Vielfalt und die Individualität der kleinen, regenerativen Erzeuger eine neue Komplexität bei der Modellierung bzw. Darstellung der Flexibilitäten.

Diese individuellen Bedingungen und Möglichkeiten sind als Suchraum jeweils in das Optimierungsmodell einzubetten, da hier nur solche Fahrpläne für die Anlagen berücksichtigt werden dürfen, die später von der Anlage im jeweiligen Betrieb auch umsetzbar sind. Ansonsten hat die gefundene Lösung der Wirkleistungsplanung keinen Wert (Ahmad 2010).

Hierfür werden neue Verfahren benötigt, die sowohl in der Lage sind, Zustandsinformationen einer großen Anzahl von Anlagen im Rahmen der Optimierung der Fahrpläne zu nutzen als auch flexibel auf Änderungen zu reagieren. Aufgrund der volatilen Umgebung (mit ständig wechselnder Verbundzugehörigkeit) muss die Modellierung bereits auf einer hinreichend abstrakten Ebene ermöglicht werden, so dass keinerlei Anpassung am System notwendig wird sobald sich die Zusammensetzung eines Verbundes oder das Profil einer Anlage ändern (Sonnenschein et al. 2014; Sudeikat & Heitmann 2013). Änderungen des Marktes stellen ähnliche Anforderungen an die Bestimmung der Flexibilität, beispielsweise durch Änderungen der Zeithorizonte oder der zeitlichen Auflösung bei unterschiedlichen Stromprodukten (Nieße et al. 2014). Bisherige Ansätze arbeiten entweder mit einer zentralen Modellierung der Nebenbedingungen oder integrieren spezifische Anlagenmodelle direkt in den Optimierungsansatz (beispielsweise Makkonen & Lahdelma 2006; Milano 2010; Chassin et al. 2008). Hierdurch werden weder die dynamische Rekonfiguration von Verbänden noch eine hinreichende Abstraktion für die automatisierte Integration neuer Anlagen (bisher unbekanntem Typs) zur Laufzeit unterstützt.

Beinahe alle Optimierungsfragestellungen aus der Praxis haben das Problem, dass der Suchraum der gültigen Lösungen durch eine Reihe von (oft nicht-linearen) Nebenbedingungen eingeschränkt wird. Nebenbedingungen (engl. Constraints) stellen Bedingungen dar, die die Lösung eines Optimierungsproblems neben der Optimalität erfüllen muss, damit sie auch einen praktischen Wert hat (Eiben & Ruttkay 1997). Lösungen, die harte Nebenbedingungen nicht erfüllen, lassen sich nicht praktisch umsetzen und sind daher nicht anwendbar. Optimierungsansätze benötigen spezielle Strategien, um Nebenbedingungen bei der Suche nach dem Optimum handhaben zu können. Gegebenenfalls ist aber eine explizite (mathematische) Beschreibung der Nebenbedingungen gar nicht verfügbar. Eine systematische Integration effizienter Strategien zur Handhabung solcher Constraints ist speziell auf dem Gebiet evolutionärer Optimierungsstrategien bisher eher unbeachtet geblieben (Kramer 2010). Im Falle von verteilten Optimierungsansätzen – wie sie beispielsweise aus Gründen der Skalierbarkeit bei der Koordination von Fahrplänen für die dezentrale Energieerzeugung erwartet werden können – kann je nach verwendetem Verfahren erschwerend hinzukommen, dass der eingeschränkte Suchraum oder alternativ die Menge der gültigen Lösungen (der Handlungsspielraum eines Akteurs) effizient an andere Akteure eines verteilten Algorithmus – beispielsweise innerhalb eines Multiagentensystems – kommuniziert werden muss (Bremer et al. 2010). Im Falle einer zentralen Planungsstelle ist die Kommunikation des Suchraums in jedem Fall erforderlich, wenn die zentrale Stelle nicht über eine eigene Modellierung aller Anlagen verfügt; was in größeren Szenarien auch nicht praktikabel umsetzbar ist.

Für eine effiziente Kommunikation wiederum ist eine kompakte Darstellung erforderlich. Für die Fahrplanbestimmung muss jede der verteilten Anlagen den Raum der (eigenen) möglichen, alternativen Handlungsweisen bestimmen und kommunizieren können. Die Prognose von möglichen Lastgängen muss in den Anlagen selbst erfolgen, da hierfür

neben den eigenen anlagespezifischen Betriebsparametern auch die Kenntnis des Bedarfs sowie des aktuellen Betriebszustandes notwendig ist.

In dynamischen Szenarien, d.h. in solchen in denen die Zusammensetzung der Gruppe von Anlagen noch nicht a priori bekannt ist, treten weitere Schwierigkeiten auf. In dynamisch gebildeten Verbänden mit a priori unbekanntem Anlagen kann in jedem Fall keine zentrale Modellierung statisch vorgegeben werden. Dies ist immer dann der Fall, wenn beispielsweise die Wirkleistungsplanung marktorientiert durchgeführt wird und sich Anlagen am Markt regelmäßig produktorientiert passend zu neuen Verbänden zusammensetzen (Nieße et al. 2014). In derartigen Fällen muss das mathematische Modell des Optimierungsproblems für die Gruppe als Ganzes automatisiert erstellt werden können; und das regelmäßig bei jeder neuen Verbundbildung. Jeder Anlagentyp verfügt über eine individuelle, nicht allgemein bekannte Modellierung seiner Flexibilität sowie seiner individuellen technischen Einschränkungen. Hieraus automatisiert ein Gesamtmodell abzuleiten ist eine nicht triviale Aufgabe, insbesondere, wenn zukünftig auch bisher unbekannte Anlagentypen zur Laufzeit eingebunden werden müssen, deren Modellierung oder Constraints erst zur Laufzeit bekannt werden (Sonnenschein et al. 2014).

## 1.2. Problemstellung

Im Smart Grid muss für die Wirkleistungsplanung zunächst also der Raum der realisierbaren Lastgänge individueller, dezentraler Anlagen effizient modelliert werden, sodass zum einen von beliebigen Constraints (auch von ggf. unbekanntem Anlagen) abstrahiert werden kann und zum anderen auch individuelle Kosten- oder Effizienzkennzahlen zur Bewertung einzelner Fahrpläne integriert werden können. Hier wird eine unabhängige Beschreibung auf abstrakter Ebene benötigt, die verschiedenste Anlagen, ihre Flexibilität und Constraints einheitlich auf eine Weise modelliert, bei der sich die Modelle verschiedener Anlagen automatisiert zu einem Gesamtoptimierungsmodell ad hoc zusammenfügen lassen.

Die effiziente Beschreibung der Suchräume einzelner Anlagen kann beispielsweise über Surrogatmodelle erfolgen, welche das technische Modell einer Anlage so abstrahieren, dass nur noch das beobachtbare Verhalten (ohne Ursachen) näherungsweise abgebildet wird. Prinzipiell kommen hier u. a. eine Vielzahl von Regressions- und Klassifikationsverfahren in Frage, welche als Blackbox-Modell eine Entscheidung über die Realisierbarkeit eines Lastgangs ermöglichen (Bremer et al. 2010). In der Regel kann durch solche Modelle für jede dezentrale Einheit der Handlungsspielraum (die Flexibilität der Anlage) durch eine verhältnismäßig kleine Menge von Beispielen beschrieben werden, ohne dass eine explizite Modellierung der Constraints auf Seiten der Koordinationsstelle notwendig wird. Diese Darstellung allein ermöglicht jedoch noch nicht die Optimierung in den dezentralen Handlungsspielräumen. Sie stellt lediglich eine kompakte Darstellung für die Übertragung oder Speicherung dar (Bremer et al. 2011a).

Unbeantwortet ist somit noch die Frage nach einem geeigneten Gesamtkonzept zur einheitlichen Integration solcher Blackbox-Modelle in (beispielsweise evolutionäre) Optimierungsstrategien. Eine besondere Schwierigkeit ist hierbei der extrem kleine Anteil zulässiger Lösungen. Dies gilt in besonderem Maße bei höherdimensionalen Problemstellungen

wie sie beispielsweise in der Wirkleistungsplanung im Smart Grid meist unvermeidbar sind (Bremer & Sonnenschein 2013e). Hierdurch wird es äußerst schwierig, existierende Heuristiken beispielsweise durch Ergänzen einer Penaltyfunktion zur Bestrafung ungültiger Lösungen einzubinden (Bremer & Sonnenschein 2012a). Die Wahrscheinlichkeit, im Zuge der Suchraumexploration eine gültige Lösung zu erhalten, ist aufgrund des Verhältnisses der Anzahl der ungültigen Lösungen viel zu gering. Eine vorherige Einschränkung der Suche auf die Region gültiger Lösungen gestaltet sich zunächst ebenfalls schwierig, da zum einen keine funktionale Beschreibung der Randbedingungen vorliegt und zum anderen gegebenenfalls auch in nicht zusammenhängenden Regionen gesucht werden muss. Es ist mit den zuvor beschriebenen Blackbox-Modellen lediglich möglich, im Nachhinein zu überprüfen, ob ein gefundener Lösungskandidat zur Lösungsmenge gehört oder nicht. Dies macht bereits das Finden einer initialen, gültigen Lösung schwierig. Auf eine mathematische Beschreibung der Nebenbedingungen für ein Standardvorgehen bei der Integration in bestehende Optimierungsansätze kann nicht zurückgegriffen werden (Bremer et al. 2011b).

Zum einen ist aufgrund der dynamischen Struktur, wodurch das exakte Optimierungsproblem vorab nicht bekannt ist, auch eine vorab vorgenommene Beschreibung der Constraints nicht möglich. Zudem beinhalten die Constraints gegebenenfalls Informationen über die Betriebsweise der Anlage, sodass diese – zumindest nicht in mathematischer Klartextform – außer Haus gegeben werden sollen oder dürfen. Gefragt ist an dieser Stelle also eine Möglichkeit, gültige Lösungskandidaten systematisch aus den vorhandenen Blackbox-Modellen generieren zu können.

Die Integration von den einzelnen Lösungen zugeordneten, individuellen Kostenkennzahlen stellt eine zusätzliche Schwierigkeit dar, ist aber ebenfalls unabdingbar (Bremer et al. 2011b), beispielsweise zur fairen Aufteilung von Kosten (Bremer & Sonnenschein 2013c).

### 1.3. Lösungsansatz

Einen eleganten Lösungsansatz für die skizzierte Problematik bieten Kern-basierte Verfahren mit denen parallel eine geometrische Beschreibung der Region gültiger Lösungen (als Unterraum des Suchraums) und die funktionale Zuordnung zu beschreibenden Kennzahlen anhand exemplarischer Lösungen erlernt werden kann. Dieser Ansatz wird zusammen mit einem hierauf aufbauenden, neuen Dekoderverfahren zu einer Gesamtmethodik für die Abbildung auf unterschiedliche Probleminstanzen entwickelt und an dem Smart Grid Beispiel evaluiert. Bei einem solchen Dekoder handelt es sich um eine Technik zur Behandlung von Constraints während der Optimierung, die die systematische Erzeugung von zulässigen Lösungen erlaubt (Coello Coello 2002). Bisherige Dekoderstrategien müssen jeweils problemspezifisch (meist von Experten) angepasst werden (Coello Coello 2002; Koziel & Michalewicz 1999), was in einem dynamischen Umfeld die Automatisierbarkeit erschwert. Hier wird eine Weiterentwicklung vorgenommen.

Als Gesamtkonzept werden zunächst die eingeschränkten Suchräume der verteilten Akteure mit Hilfe eines Supportvektor-Surrogat-Modells kodiert, das von jedweder individueller, mathematischer Beschreibung der Anlagenfähigkeiten und -constraints abstrahiert.

Solche Stützvektormethoden (engl. support vector) für die Datenmodellierung nutzen zur Beschreibung eines Datenraums eine Trainingsmenge mit Beispieldaten aus diesem Raum und wählen hieraus eine Teilmenge (die Supportvektoren) mit denen eine die Daten einschließende Hüllkurve berechnet werden kann (Tax & Duin 2004; Ben-Hur et al. 2001). Ziel ist die Nutzung einer Supportvektor Darstellung als Blackbox-Modell. Diese erlaubt im hier skizzierten Anwendungsfall die effiziente Darstellung von Mengen potenziell realisierbarer Anlagenfahrpläne (inklusive Kosten) zur Kommunikation von individuellen Handlungsoptionen.

Aufbauend hierauf wird ein Ansatz entwickelt, den unbeschränkten Suchraum mithilfe der beschreibenden Supportvektoren so auf den beschränkten Teilraum der gültigen Lösungen abzubilden, dass eine Definition von beispielsweise Nachbarschaftsbeziehungen im unbeschränkten Raum möglich wird während gleichzeitig die Evaluierung der Zielfunktion unter der Abbildung im eingeschränkten Lösungsraum vorgenommen werden kann. Hierbei handelt es sich um eine spezielle Methode (ein so genannter Dekoder) zur Handhabung von Constraints während der Optimierung (Koziel & Michalewicz 1999), die in der Regel speziell auf die jeweilige Instanz des Optimierungsproblems angepasst werden muss (Coello Coello 2002). In dem hier vorgestellten Ansatz wird es möglich sein, den Dekoder automatisiert zu erstellen. Zugeordnete Kostenkennzahlen werden bei diesem Vorgehen geeignet mit transformiert, sodass eine entsprechende Bewertung auch in dem unbeschränkten Raum möglich wird. Auf diese Weise wird es möglich, Restriktionen und individuelle Kosten so in einen beliebigen Optimierungsansatz zu integrieren, dass diese nicht mehr explizit modelliert werden müssen. Dies ist neben der automatisierten Ableitung eine weitere Erweiterung der Funktionalität bestehender Dekoderansätze. Das Optimierungsproblem wird durch den Dekoder in ein unbeschränktes Problem transformiert. Ein solcher Dekoder ist aber auch im vollständig verteilten Optimierungsszenario mit rein lokaler Entscheidungsfindung von großem Nutzen, um systematisch zulässige (Teil-)Lösungen einschließlich Bewertung zu generieren.

Hierzu muss ein geeignetes Verfahren für die Abbildung entwickelt werden. Mit dieser Abbildung liegt neben der Supportvektor-basierten Kodierung der Handlungsspielräume eine zweite essentielle Voraussetzung zur Entwicklung eines Gesamtkonzepts vor. Im Ergebnis entsteht ein individuell anpassbares Metakonzept, welches auf verschiedene Probleminstanzen mit dezentralen Handlungsspielräumen abgebildet werden kann.

Ziel der Arbeit ist somit die Entwicklung eines hybriden Ansatzes welcher die Integration modellunabhängiger Beschreibungen beliebiger elektrischer Anlagen mit verschiedenen Optimierungsverfahren ermöglicht. Hierzu wird zunächst ein Dekoder entwickelt, welcher auf einem Supportvektor-Modell zur Beschreibung möglichen Anlagenhaltens aufsetzt. Die Hybridisierung mit verschiedenen Optimierungsansätzen ergibt sich durch die Reformulierung der Problemstellung. Bei dieser Umformulierung wird durch den entwickelten Dekoder das Optimierungsproblem dadurch vereinfacht, dass alle technischen Modelle der Anlagen einschließlich der individuellen Constraints in Dekodern gekapselt werden. Der Lösungsalgorithmus selbst braucht daher keine Nebenbedingungen mehr zu beachten. Hierdurch wird eine constraintfreie Formulierung möglich, wodurch verschiedene Standardansätze zur Optimierung ohne spezielle Anpassung an das Problem und ohne ansatzspezifische Constraintmodellierung eingesetzt werden können.

## 1.4. Aufbau der Arbeit

Diese Arbeit gliedert sich wie folgt. Ausgehend von einer Diskussion aktueller Herausforderungen für die zukünftige Energieversorgung, die maßgeblich durch den (gewollten) Ausbau verteilter, regenerativer und dadurch schwerer vorhersagbarer Energieerzeuger geprägt ist, werden im Anschluss zunächst verwandte Arbeiten auf dem Gebiet der (verteilten) Planung von Wirkleistungserzeugern und -verbrauchern betrachtet. In beiden Kapiteln sind zunächst Hintergründe und Verfahren von Interesse, die sich mit der bedarfsgerechten und marktorientierten Bereitstellung von Wirkleistung – oder genauer: deren vorausschauender Planung – befassen; aber auch die Einbettung in den Gesamtkontext wird hier vorgenommen.

Der übrige Grundlagenteil wird sich in den Kapiteln Kapitel 4 ab Seite 35 und Kapitel 5 ab Seite 53 mit zwei grundlegenden Technologien befassen, die für diese Arbeit von Bedeutung sind: die Behandlung von Constraints während der Optimierung sowie Supportvektorverfahren. Während erstere den Forschungskontext aufspannen, in den sich der Optimierungsteil dieser Arbeit einbettet, bereitet der Supportvektorteil die spätere Umsetzung vor, die maßgeblich auf dieser Technik fußt.

Der zweite Teil dieser Arbeit befasst sich mit der Formalisierung der Problemstellung und der Umsetzung in einen Lösungsansatz. Da sich die Gesamtlösung als Prozesskette mit verschiedenen, einzeln betrachtbaren Teillösungen beschreiben lässt, erläutert Kapitel 7 auf Seite 91 zunächst den Überblick über die einzelnen Schritte sowie das Zusammenspiel der einzelnen Komponenten. Im Anschluss werden die einzelnen Komponenten kapitelweise genauer beschrieben. Zunächst befasst sich Kapitel 8 ab Seite 101 mit der Erstellung eines abstrakten Modells der Flexibilität einzelner Anlagen, mit dem deren Möglichkeiten als Alternativlösungen für die Optimierung beschrieben werden. Kapitel 9 ab Seite 117 erläutert im Anschluss, wie dieses Modell genutzt werden kann, um systematisch gültige Lösungen während der Optimierung zu generieren. Ein wichtiger Punkt der Wirkleistungsplanung ist die Fähigkeit eine Bewertung der Lösungen hinsichtlich verschiedener Kriterien vornehmen zu können. Neben dem Erreichen des Planungsziels spielen auch Kosten, Robustheit des Plans u.ä. eine Rolle. Kapitel 10 ab Seite 133 befasst sich mit der Möglichkeit, Fahrpläne individuell mit solchen Bewertungskennzahlen versehen zu können. Die Beschreibung der einzelnen Komponenten beschränkt sich in diesen Kapiteln zunächst noch auf die Beschreibung und Illustration der jeweiligen Funktionsweise.

Die Evaluation der Funktionalität sowohl der einzelnen Komponenten als auch des Anwendungsfalls Wirkleistungsplanung im Gesamtkontext findet sich im dritten Teil dieser Arbeit. Hier wird in Kapitel 11 ab Seite 147 als erstes eine Betrachtung der Fitnessfunktion für das fragliche Wirkleistungsplanungsproblem vorgenommen, um eine Entscheidung für die Verwendbarkeit bestimmter Optimierungsverfahren treffen zu können. In der weiteren Beschreibung des verwendeten Evaluationsansatzes werden sowohl die betrachteten Szenarien hergeleitet als auch die verwendeten Anlagenmodelle. Die einzelnen Komponenten werden zunächst einzeln bezüglich ihrer spezifischen individuellen Funktion evaluiert. Die Einbettung in den Anwendungskontext und die Anwendbarkeit für das betrachtete Wirkleistungsplanungsproblem werden in den zwei nachfolgenden Kapiteln getrennt für zentrale und für dezentrale Optimierungsverfahren untersucht.

Einige über das Basisplanungsproblem hinausgehende Fragestellungen werden eben-

## 1. Einleitung

---

falls noch angerissen. Da diese Fragestellungen jedoch noch nicht abschließend vollumfänglich gelöst sind obwohl bereits erste Ergebnisse vorliegen, sollen hier neben der Hinführung zur Problemstellung lediglich Wege aufgezeigt werden, wie die jeweilige Problematik angegangen werden kann. Dies betrifft Fragen nach dem Umgang mit Prognoseunsicherheiten (Kapitel 14 ab Seite 209), die Frage nach der fairen Aufteilung eines gemeinschaftlich erwirtschafteten Gewinns (Kapitel 15 ab Seite 217) sowie die Sicherstellung der Interoperabilität im Kontext mit möglichen Kommunikationsstandards, die im Smart Grid zum Einsatz kommen könnten (Kapitel 16 ab Seite 225).

Der dritte Teil schließt mit einer Betrachtung weiterer Anwendungsfälle im Smart Grid sowie der Übertragbarkeit auf andere Anwendungsdomänen und einem Fazit der Arbeit.

**Teil I.**

**Grundlagen**



## 2. Problemstellung im Kontext Smart Grid

*Electricity is really just organized lightning*

---

*(George Carlin)*

### 2.1. Einleitung

Die Problemstellung für die im Rahmen dieser Arbeit ein Lösungskonzept erarbeitet und vorgeschlagen wird lässt sich am geeignetsten anhand eines konkreten Anwendungsszenarios aus der Praxis motivieren und illustrieren. Hierfür wurde die automatisierte Wirkleistungsplanung durch autonome, verteilte Einheiten in einem Smart Grid gewählt. Bevor zu Ende dieses Kapitels die Fragestellung auf Basis der Anwendung konkretisiert werden kann, werden im Folgenden zunächst hinleitend das Thema Smart Grid und die darin enthaltene Wirkleistungsplanungsthematik eingeführt.

Die elektrische Energieversorgung nicht nur in Deutschland ist derzeit einem rasanten Wandel unterlegen. Ein Paradigmenwechsel sowohl bei der Erzeugungsstruktur als auch bei der informationstechnischen Einbindung neuer Anlagen und Systeme zeichnet sich bereits deutlich ab. Gefordert ist heute mehr denn je eine Klima- und Ressourcenschonende Erzeugung elektrischer Energie und deren sichere, effiziente und preisgünstige Bereitstellung – auch in Zukunft.

Konsequenterweise und nicht zuletzt aufgrund diverser Förderungen gehen immer mehr Anlagen ans Netz welche elektrische Energie aus erneuerbaren Ressourcen gewinnen (vgl. auch Abbildung 2.1). Die 2009 beschlossene Richtlinie der EU (European Parliament & Council 2009) legt verbindliche Zielwerte für den Anteil erneuerbarer Energie in den einzelnen Mitgliedsstaaten fest, um EU-weit einen Anteil von mindestens 20% bis zum Jahr 2020 zu erreichen. Als Reaktion hierauf brachte die Bundesregierung den Nationalen Aktionsplan (Deutsche Bundesregierung (Bundeskabinett) 2010) für die Bevorzugung erneuerbarer Energien auf den Weg. Hierzu gehört auch das Erneuerbare-Energien-Gesetz (EEG) (vgl. z. B. Deutsche Bundesregierung (Bundeskabinett) 2012)

Der zunehmende Anteil erneuerbarer Energieressourcen verursacht aber auch eine Reihe neuer Herausforderungen, die es zu meistern gilt. Diese sind sowohl ökonomischer als auch technischer Natur. Eine verteilte Einspeisung nunmehr auch auf den unteren Spannungsebenen wird zu einer sich kontinuierlich ändernden teilweisen Umkehr des Stromflusses und zu Problemen bei der Einhaltung zulässiger Spannungswerte führen, die Stabilität des Versorgungsnetzes wird zukünftig ebenfalls verteilt sichergestellt werden müssen, die stochastische Natur der Erzeugung aus regenerativen Quellen macht

## 2. Problemstellung im Kontext Smart Grid

die Stromversorgung schwerer planbar und die bloße Zahl verteilter Erzeuger lässt die Problemgrößen in schwierig zu beherrschende Dimensionen anwachsen (Ramchurn et al. 2012; Nieße et al. 2014; Lopes et al. 2007; Horenkamp et al. 2007).

Bereits im Jahr 1999 wurde in dem Bericht der EPRI (1999) auf die zusätzlichen Potenziale durch die Integration intelligenter Mehrwertdienste und neuer Stromprodukte hingewiesen (Rohlfing 2012). Hieraus lässt sich bereits erahnen, dass für die vollständige Umstellung der Elektrizitätsversorgung auf dezentrale, erneuerbare Quellen neue und intelligente Steuerungsansätze erforderlich sein werden (Ramchurn et al. 2012), auch um die benötigte Energie bedarfsgerecht bereitstellen zu können.

Genau diese Herausforderungen und Lösungsansätze und die hieraus erwachsenden Anforderungen sind Thema dieses Kapitels als Hinleitung zum Wirkleistungsplanungsproblem, welches das Evaluationsszenario für den hier entwickelten Ansatz stellen wird.

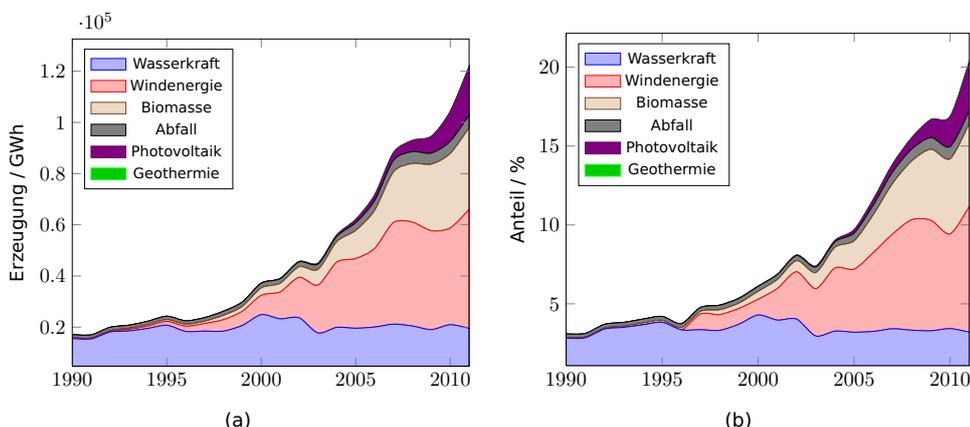


Abbildung 2.1.: Entwicklung des Betrags ausgewählter erneuerbarer Energieträger in Deutschland, als Jahresproduktion 2.1(a) bzw. als Anteil am jeweiligen gesamten Stromverbrauch 2.1(b). Eigene Darstellung nach Daten aus (BMWi 2012)

Für ein Verständnis des Problemkontextes ist es zunächst wichtig, den strukturellen Wandel in der Versorgung mit elektrischer Energie und die hiermit einhergehenden technischen aber auch organisatorischen Probleme zu betrachten. Als Teil der Gesamtproblematik ergeben sich hieraus neue Anforderungen auch an die vorausschauende Planung der Bereitstellung von Wirkleistung als pro-aktive Maßnahme zur Sicherstellung der Deckung einer zu erwartenden Stromnachfrage. Betraf diese Einsatzplanung in der Vergangenheit lediglich eine überschaubare Menge von Großkraftwerken, die es auf eine prognostizierbare Verbrauchskurve abzustimmen galt, so wird diese Aufgabe in Zukunft von einer Heerschar individuell agierender, dezentraler Akteure mit übernommen werden müssen (Nieße et al. 2012a). Hierdurch entstehen sowohl Fragen bezüglich der Modellierung des Planungssystems als auch der Koordinierungsansätze, die hier aufgezeigt werden sollen.

Abschließend werden in diesem Kapitel erste, abstrakte Formulierungen der Arbeitsziele dieser Arbeit sowie der damit einhergehenden Forschungsfragen vorgenommen.

## 2.2. Wandel der Energieversorgung

Die Sicherstellung einer zukünftigen Energieversorgung basierend auf der Verbrennung fossiler Energieträger scheint aus zwei wesentlichen Gründen nicht zu erreichen und schwerlich akzeptabel. Zum einen herrscht mittlerweile wissenschaftlicher Konsens über die klimaschädliche Wirkung von durch die Verbrennung fossiler Energieträger freigesetzter Treibhausgase wie  $CO_2$  (Doran & Zimmerman 2009; Brown et al. 2008), zum anderen erfordert der stetige Rückgang der Förderung aus den derzeit erschlossenen fossilen Quellen die Erschließung neuer Ressourcen (Ramchurn et al. 2012). Einsparungen allein werden dem UN-Weltklimareport zufolge nicht in der Lage sein, die aus sowohl dem Bevölkerungsanstieg als auch aus dem Wohlstandswachstum resultierenden Mehrbelastungen zu kompensieren (IPCC 2011).

Die Erschließung wenig rentabler Ressourcen führt zum Einsatz von Gewinnungstechnologien, welche weitere Umweltbelastungen nach sich ziehen und setzt häufig gerade die Bevölkerung in ohnehin unterversorgten Regionen unnötigen Gesundheitsgefahren aus (Friedman 2009). Ein Beispiel hierfür ist die Gewinnung von Rohöl aus Teer- bzw. Ölsanden in Kanada (Casey-Lefkowitz et al. 2010; Charpentier et al. 2009). Nicht zuletzt werden die immer knapper werdenden Mengen und weniger lukrativen Fördermethoden zu einer weiteren Verteuerung von Energie führen.

Stellt man den Klimaschutz und somit das Ziel der  $CO_2$ -Reduktion in den Vordergrund so wird klar, dass eine Umstellung der Energieversorgung auf umweltschonende Erzeugung aus erneuerbaren Energieträgern trotz einer gleichzeitigen Forcierung effizienterer Verbraucher eine erhöhte Nachfrage nach elektrischer Energie vermuten lässt. Dies liegt in dem Umstand begründet, dass bisherige fossile Brenn- und Treibstoffe zumindest teilweise durch Elektrizität substituiert werden. Im Bereich der häuslichen Heizwärmeerzeugung ist beispielsweise mit einer verstärkten Verbreitung von elektrisch betriebenen Wärmepumpen zu rechnen.

Einen weitaus größeren Anteil am Anstieg der Nachfrage nach elektrischer Energie dürfte der zu erwartende (und aus Klimasicht durchaus sinnvolle) starke Ausbau des elektrischen Individualverkehrs zusammen mit elektrisch betriebenen Fernzügen ausmachen. In Deutschland wird der Schienenverkehr zwar über ein separates Stromnetz versorgt, international ist dies aber nicht generell der Fall.

Die Dena Netzstudie II geht zwar von einem Rückgang der Nachfrage nach elektrischer Energie von 8-10% bis zum Jahr 2020 (-25% bis 2050) aus (Dena 2010), rechnet hierin aber keinen zusätzlichen Bedarf durch beispielsweise Elektromobilität ein. Die eEnergy-Studie von Franz et al. (2006) sieht bei sinkender Gesamtnachfrage nach Energie eine Steigerung der Nachfrage nach Strom von 4% bis 2030 (gegenüber 2002). Die Studie von Boßmann et al. (2012) geht davon aus, dass ab ca. 2030 genug Einsparpotenzial mobilisiert werden kann, um die steigende Nachfrage durch Elektromobilität, Heizwärme, etc. aufzufangen. In jedem Fall werde der Bedarf zum Großteil durch regenerative Energiequellen wie Wind, Photovoltaik, Wasserkraft oder Biomasse gedeckt werden müssen.

Die grundlegende Struktur des elektrischen Versorgungsnetzes hat sich seit seinen Ursprüngen strukturell nur wenig geändert. Es wurde entworfen und konstruiert für den effizienten Transport elektrischer Energie von wenigen Großherzeugern zu einer Vielzahl

von verteilten Groß- und Kleinverbrauchern. Die zentrale Idee war, Energie nach Bedarf zu produzieren von einer überschaubaren Anzahl gut steuerbarer Großerzeuger basierend auf Standardlastprofilen für gut bekannte und prognostizierbare Verbraucher. Große Regelkraftwerke gleichen unvermeidbare Schwankungen aus.

In Deutschland erfolgt die Verteilung elektrischer Energie über eine hierarchisch aufgebaute Netzstruktur welche auch die verschiedenen Spannungsebenen für die unterschiedlichen Verteilungsaufgaben widerspiegelt. Beschreibungen finden sich in (Heuck et al. 2007; Bendrat 2010). Auf oberster Ebene dient ein 380/420 kV (Nominalspannung  $U_n$  bzw. zulässige Maximalspannung  $U_m$ ) Höchstspannungsnetz als Redundanz, um etwaige Kraftwerksausfälle durch Import von elektrischer Energie ausgleichen zu können. Als Variante kann dieses auch als 220 kV Netz ausgeführt sein. Dieses ist als Maschennetz aufgebaut, um große Versorgungssicherheit bei gleichzeitiger großer Ausdehnung zu gewährleisten und eine ausgeglichene Netzlast zu ermöglichen. Großkraftwerke stehen üblicherweise zentral in der Nähe großer Verbraucherzentren und speisen in das 110/123 kV Hochspannungsnetz ein, welches die Verteilung in die Ballungszentren hinein übernimmt. Diese sind als Ringnetz ausgeführt, was eine selektive Abschaltung ermöglicht und gute Spannungshaltungseigenschaften aufweist. Dort wird auf 20/24 kV für das Mittelspannungsnetz (Maschen- oder Ringnetz) für die Verteilung innerhalb des Ballungszentrums (bzw. 10 kV in rein städtischen Gebieten) heruntertransformiert bevor schließlich das 400 V Niederspannungsnetz den Transport zu den Kleinabnehmern (Haushalte, Kleingewerbe, u. ä.) übernimmt. Letzteres ist üblicherweise (insbesondere in ländlichen Gegenden) als Strahlennetz ausgeführt (Bendrat 2010). Dieser einfache Aufbau ist zwar kostengünstig, weist aber schlechte Spannungshaltungseigenschaften auf. Zudem ist keine Redundanz bei der Versorgung vorhanden.

Mit steigender Erzeugung durch erneuerbare Energieträger wandelt sich dieses gut prognostizierbare System mehr und mehr zu einer hochdynamischen Struktur. Dezentrale Erzeuger sind in der Regel vergleichsweise kleine Erzeugungseinheiten. Mit steigender dezentraler Erzeugung (speziell durch eingeschränkt prognostizierbare Erzeuger) wird sich ein sowohl in Stärke als auch Richtung kontinuierlich schwankender Stromfluss ergeben. Einspeisung elektrischer Energie wird es zukünftig mehr und mehr auch auf den unteren Netzebenen geben. Häusliche Photovoltaik-Anlagen und Mikro-BHKW haben ihren Anschluss im Niederspannungsnetz; Windenergie, größere BHKW und Solarparks hingegen auch im Mittelspannungsnetz. Eine zeitweise Umkehr des Stromflusses je nach Situation im Netz ist die logische Folge. Dennoch müssen jederzeit die technischen Nebenbedingungen des Netzes (beispielsweise die Einhaltung von Spannungsband und maximalem Stromfluss) und der einzelnen Anlagen (z. B. Anfahrtrampen) für einen sicheren Betrieb eingehalten werden.

Der Betrachtungsgegenstand bei der Beschäftigung mit dem Thema Smart Grid geht üblicherweise über die rein technischen Aspekte des Stromnetzes hinaus. Ziel ist die Gesamtheit der Ansätze, welche darauf abzielen eine verstärkte Integration von Energie aus erneuerbaren Quellen zu ermöglichen. Hierzu muss die fluktuierende Erzeugung aus regenerativen Quellen nicht nur in das Netz sondern auch in den Markt integriert werden (Bundesnetzagentur 2011). Insbesondere weist die Bundesnetzagentur in ihrem Eckpunktepapier (Bundesnetzagentur 2011) auf die starke wechselseitige Abhängigkeit zwischen rein technischen Integrationslösungen und Marktintegrationsansätzen hin. Für

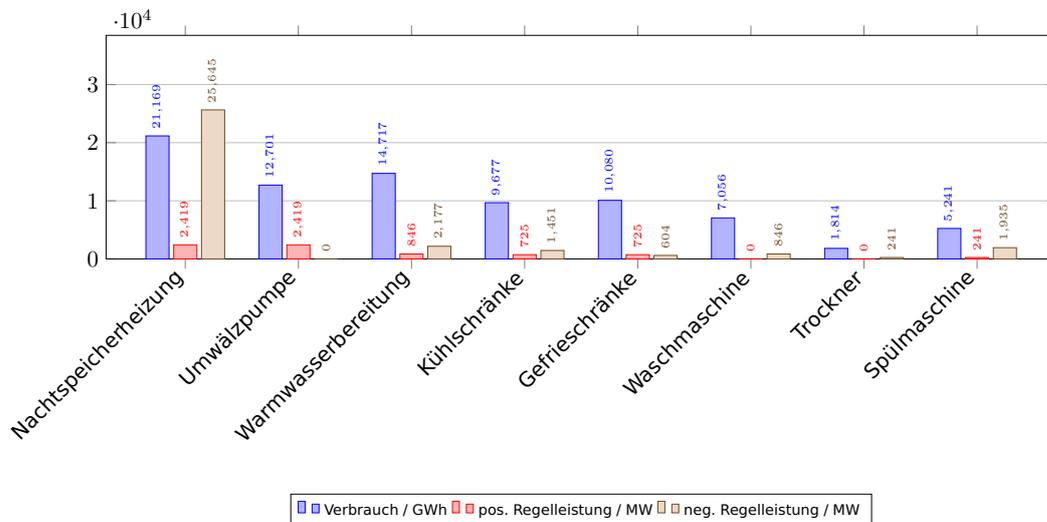


Abbildung 2.2.: Potenziale für Regelleistung durch DSM-Maßnahmen im Haushaltssektor; nach (Dena 2010), verändert.

eine Umsetzung der Energieversorgung als Smart Grid müssten immer beide Aspekte betrachtet werden. Insbesondere müssten Aufgaben, bei denen die Aussicht bestünde, dass Konkurrenz, Heterogenität und Differenzierung die besseren Lösungen lieferten (Bundesnetzagentur 2011, S. 6), über einen entsprechenden Markt geregelt werden.

Zur technischen Reife erneuerbarer Energien gehört auch deren Einordnung in Marktgegebenheiten sowie die zunehmende Übernahme von netzstützenden Maßnahmen. Hierfür müsse laut Bundesnetzagentur auch die Erzeugerseite künftig zunehmend auf Marktsignale reagieren. Geschehe dies nicht werde der weiter zurückgehende konventionelle Erzeugungsanteil einer weiter zunehmenden Preisvolatilität ausgesetzt (Bundesnetzagentur 2011, S. 10). Ohne Markorientierung müssten die Netze stärker ausgebaut werden, um einen hinreichenden Mindestbestand an konventionellen Reserve- und Regelkraftwerken zu erreichen.

## 2.3. Koordinationsaufgaben im Verteilnetz

Bisher erfolgte die Stromerzeugung üblicherweise durch wenige vorhersagbare Großkraftwerke mit aufeinander abgestimmten Antwortzeiten bezüglich der Regelbarkeit. Hiermit wurde in Summe der Bedarf gedeckt. Die Erzeugung aus regenerativen Quellen wie Wind oder Sonnenenergie jedoch folgt anderen Notwendigkeiten. Die Erzeugung ist abhängig von der jeweils aktuellen Wetterlage und kann kaum sinnvoll geregelt werden. Lediglich ein Abregeln der Anlagen ist in gewissen Grenzen möglich. Schlechte Kurzzeitprognosen von Wind und Wolken führen zu stochastisch erscheinender Einspeisung. Andere Erzeuger wie beispielsweise Blockheizkraftwerke müssen ihre Erzeugung an einen parallelen Wärmebedarf koppeln

Eines der wesentlichen Prinzipien bei der Führung konventioneller Stromversorgungs-

systeme, die Lastorientierung bei der die aktuelle Produktion jederzeit der aktuellen Last angepasst wird, wird in Zukunft ein zweites Betriebsprinzip ergänzend zur Seite gestellt bekommen (Sonnenschein et al. 2011): Das Prinzip der Lastanpassung (Last folgt Erzeugung). Bei diesem Betriebsprinzip unterstützen Verbraucher die Netzführung indem sie ihre Lasten steuerbar machen und somit neue Potenziale (vgl. Abbildung 2.3 auf der vorherigen Seite) durch Bereitstellung zusätzlicher Freiheitsgrade für die Planung und Regelung erschließen. Regelungspotenziale, die verloren gehen da eine immer größer werdende Zahl an Erzeugern nicht (sinnvoll) regelbar ist (z. B. Windenergie), können so teilweise kompensiert werden. Dieses neue Steuerungsparadigma, welches alle Maßnahmen zur Beeinflussung (im Sinne des Netzbetriebs) der Last auf Verbraucherseite zusammenfasst, wird auch als Demand-Side Management bezeichnet (Sonnenschein et al. 2010). Die Idee der Nachfragebeeinflussung ist nicht neu (Gellings & Chamberlin 1988), allerdings war die ursprüngliche Intention eine Reduktion der Lastschwankungen. Heute werden jedoch eine Reihe weiterer Anwendungsfälle und Möglichkeiten der Einflussnahme als Lastmanagementstrategien diskutiert (Stadler 2005; Charles River Associates 2005; Gellings 2009):

**Peak Clipping:** Hier werden seitens des Energieversorgers Lastspitzen gekappt. Dabei ist sowohl das gezielte Abschalten nicht dringend benötigter Verbraucher denkbar als auch eine geeignete Bepreisung im Rahmen eines Spitzenlastmanagements (Sonnenschein et al. 2010; Heindler 2001).

**Valley Filling:** In diesem gegenläufigen Ansatz ist die Absicht das Auffüllen von Lasttälern zur besseren Kraftwerksauslastung. Erreicht wird dies durch Schaffung einer zusätzlichen Nachfrage zu üblichen Schwachlastzeiten, z. B. durch Nachtspeicherheizungen (Stadler 2005).

**Load Shifting:** Die größte Flexibilität bietet die Verschiebung von Lasten. Hier wird der Zeitpunkt der Leistungsaufnahme verschoben (üblicherweise von Starklast- zu Schwachlastzeit). Im Gegensatz zu den beiden anderen Methoden ändert sich hier die Nachfrage (zumindest über mittelfristige Zeiträume betrachtet) nicht (Sonnenschein et al. 2011).

Für große Verbraucher (Industrieanlagen) ist eine prinzipielle Mitwirkung an der Laststeuerung schon seit längerem (wenngleich auch in begrenztem Umfang) vorgesehen. Energieversorger bieten Großverbrauchern oft günstigere Tarifkonditionen für die Option ggf. abgeschaltet zu werden, um bei einem Engpass auszuhelfen. Hier ist allerdings keine feingranulare Steuerung insbesondere nicht auf kurzen Zeitskalen möglich. Ebenfalls im Industriesektor seit längerem gebräuchlich ist der Einsatz eines Spitzenlastmanagements. Durch die Bepreisung mit Leistungspreisen (für die Bereitstellung der Kapazität während der Lastspitzen) lohnt es sich für Unternehmen, IT-gestützt Lasten schalten zu lassen, um die eigene Spitzenlast zu reduzieren (Heindler 2001).

Akzeptables Potenzial verspricht man sich jedoch auch von einer Vielzahl von Geräten im Haushaltssektor (vgl. Abbildung 2.3) – auch wenn diese aufgrund ihrer geringen Leistungsklasse einzeln kein oder kaum Potenzial aufweisen (Borenstein et al. 2002; Stadler 2003; Charles River Associates 2005; Seebach et al. 2009). Es ist zu erwarten, dass sich diese Potenziale durch geschicktes Zusammenfassen (Pooling) zu Gruppen erschließen

lassen, ohne dass es zu unbeherrschbaren Problemen bei der Skalierung der Problemgröße kommt (Lünsdorf & Sonnenschein 2010).

Methodisch kann allgemein zwischen direkten und indirekten Verfahren unterschieden werden. Bei den indirekten Verfahren wird häufig durch das Setzen von Anreizen (etwa über Preissignale) versucht, die Nachfrage zu beeinflussen. Eine entsprechende Reaktion auf die Preise kann auch automatisiert durch geeignete Energiemanagementsysteme erfolgen (Bendel et al. 2006).

Virtuelle Kraftwerke (Lukovic et al. 2010) stellen ein seit längerem erforshtes Konstrukt zur Integration dezentraler Energieanlagen in das Verteilnetz dar. Das Konzept sieht vor, die Gesamtsystemkomplexität zu mildern, indem dezentrale Anlagen zu Gruppen zusammengefasst und als Einheit betrachtet werden. Kapitel 3 ab Seite 23 wird sich eingehender mit diesem Konzept befassen. Die zu erwartende Notwendigkeit einer Integration existierender aber auch zukünftiger Energiemärkte lässt aktuelle Forschungen bereits an Erweiterungen arbeiten, welche auf ein Gesamtsystem mit sich selbsttätig, dynamisch rekonfigurierenden aber dafür stets bedarfsgerecht ausgerichteten Gruppierungen abzielen (Nieße et al. 2014).

Die Automobilbranche steht in den Startlöchern zu einem Wandel hin zur Elektromobilität. Erste kommerzielle Elektrofahrzeuge sind am Markt verfügbar oder werden derzeit erprobt. Als Beispiel sei hier die Studie *Erforschung und Erprobung neuer Fahrzeugkonzepte zur Elektromobilität* (Knorre 2011) genannt, in der auch die Großserienreife bestätigt wurde. Die Ziele der Bundesregierung (nationaler Entwicklungsplan Elektromobilität) sehen bis zum Jahr 2020 einen Ausbau des Bestands auf 1 Mio. Elektrofahrzeuge vor. Andere Studien erwarten jedoch durchaus geringere Zahlen (Schott et al. 2013).

Die typische Anschlussleistung für Elektrofahrzeuge liegt zwar nur bei maximal etwa 3,7 kW (einphasiger Anschluss, 16 A), es ist jedoch zu erwarten dass bei weiterem Ausbau und technischer Weiterentwicklung ein relevanter Anteil der Fahrzeuge auch mit Schnellladesystemen von bis zu 44 kW ausgestattet sein wird (Sonnenschein et al. 2011). Ohne intelligente Steuerung des Ladevorgangs ist ein hoher Gleichzeitigkeitsfaktor bei den Ladevorgängen zu erwarten (beispielsweise zur Feierabendzeit). Hierdurch wären ohne steuerndem Eingriff negative Auswirkungen auf die Spannungsqualität im Niederspannungsnetz (Nischler et al. 2011; Kahlen et al. 2010) und eine erhöhte Lastgangspreizung (Reiner et al. 2009) zu erwarten.

Der Einsatz intelligenter Steuerungen und passender Kommunikationstechnik mit dem Fahrzeuge während des Ladevorgangs kontrolliert werden, wird es ermöglichen den Ladewunsch termingerecht abzuschließen bei gleichzeitiger Zwischennutzung des Akkus zum Lastausgleich für das Netz. Man spricht bei dieser Nutzung von Elektrofahrzeugen im Netz von Vehicle-to-Grid (V2G) Ansätzen (z. B. Tröschel et al. 2011). Derartige Ansätze werden durchaus als zukünftige Schlüsseltechnologien im Rahmen eines Lastausgleichs gehandelt (Ramchurn et al. 2012). Aufgrund der Mobilität handelt es sich jedoch um einen Akteur, der dynamisch zu integrieren ist.

Durch ihre Fähigkeit, elektrische Energie sowohl aus dem Netz aufzunehmen (Verbraucher beim Laden) als auch bei Bedarf ans Netz abzugeben (Erzeugung durch Rückspeisung) gehören Elektrofahrzeuge wie auch andere Speichertechnologien zu einer besonderen Kategorie von Akteuren im Netz. Im Englischen hat sich hierfür der bereits in den 70er

Jahren von Toffler (1970) geprägte Begriff *Prosumer* (aus Producer und Consumer) eingebürgert. Dieser Begriff wird auch in dieser Arbeit als Sammelbegriff für alle Anlagen verwendet, die sowohl elektrische Energie ins Netz einspeisen als auch elektrische Energie aufnehmen können. Im Rahmen eines intelligent gesteuerten Ladevorgangs stellt ein Fahrzeugakkumulator somit einen Speicher dar, der prinzipiell sowohl positive, als auch negative Regelleistung für das Netz zur Verfügung stellen kann (Sonnenschein et al. 2011).

Die Akteure, die in Zukunft und auch schon heute im Versorgungsnetz aufeinander abzustimmen sind, sind nicht nur viel zahlreicher als in der Vergangenheit sondern weisen auch eine deutlich größere Diversität sowohl in ihrem Verhalten als auch ihren Randbedingungen, die mögliche Betriebsweisen diktieren, auf. Nicht nur eine stetig steigende Zahl sondern auch immer mehr verschiedene Typen sind in die Planungs- und Kontrollstrukturen des Netzes einzubinden.

Dadurch, dass alle steuernden Aufgaben im Netz perspektivisch von einer Vielzahl kleiner Anlagen zu übernehmen sind, ist davon auszugehen, dass dies nur durch Zusammenschluss zu bewältigen ist. Eine Kommunikationsverbindung, mit dem die Anlagen untereinander kommunizieren können sowie ein geeigneter Verständigungsstandard werden unabdingbare Voraussetzung hierfür sein. Als weiterer Effekt der Dezentralisierung kommt ein Wegfall zentraler Kontrollinstanzen hinzu. Hierdurch wird eine ad hoc Zusammenarbeit verschiedener, durch unabhängige Eigentümer betriebener Anlagen notwendig; im Gegensatz zu heutigen Strukturen mit Großbetreibern mit Entscheidungsgewalt über eine Vielzahl eigener Generatorblöcke.

Alles in allem ergeben sich hieraus für die Zukunft neue Anforderungen an die Steuerung dieses komplexen Systems. Die zu erwartende riesige Anzahl an individuell konfigurierten Erzeugern sowie ihre Variabilität legen nahe, dass diese weitgehend autonom agieren müssen und dabei zwar durch den Menschen überwacht werden aber nicht unter seiner direkten Kontrolle stehen (Ramchurn et al. 2012).

Es gibt also gute Gründe, unser derzeitiges elektrisches Versorgungsnetz zu transformieren in ein Netz, welches

- ... die notwendigen Kontrollmechanismen für eine Vielzahl dezentraler Akteure zur Seite stehen hat.
- ... sich automatisch überwacht und schützt (Mayer & Rohjans 2010).
- ... einen bidirektionalen Informationsaustausch – auch peer-to-peer – zwischen den verschiedenen Akteuren (Anlagen) im Netz unterstützt (Schönberg 2012; Ekanayake et al. 2012).
- ... diese Informationen automatisiert instrumentalisiert und so für einen möglichst selbstorganisierten Ausgleich von Erzeugung, Speicherung und Verbrauch sorgt (*European SmartGrids Technology Platform 2006*).
- ... den ökologischen Nutzen in den Vordergrund stellt (*European SmartGrids Technology Platform 2006*).

Nun ermöglicht Kommunikation alleine noch keine Koordination zwischen den verschiedenen Akteuren im Netz (Meister 2012). Hierzu werden auch neue Planungs- und Kontrollalgorithmen benötigt. Eine derartige Neugestaltung lässt eine Reihe technischer

Herausforderungen erwarten und wird zu einem Paradigmenwechsel bei den Steuerungsansätzen führen.

Der Betrieb des elektrischen Versorgungsnetzes verlangt eine Vielzahl an Steuerungs- und Kontrollaufgaben und die Integration erneuerbarer Energiequellen sowie die damit einhergehende Transformation des Versorgungsnetzes zu einem dezentral geführten Gebilde mit verteilten Entscheidungsträgern bedarf neuer Forschungsansätze auf diversen Ebenen. Diese Arbeit wird sich im Folgenden mit dem Thema der Wirkleistungsplanung weiter beschäftigen.

Die Aufgabe der Wirkleistungsplanung ist die vorausschauende Einsatzplanung für eine Gruppe von elektrischen Erzeugern (perspektivisch auch in Kooperation mit steuerbaren Verbrauchern). Ein erwarteter elektrischer Verbrauchslastgang soll durch eine entsprechende Erzeugung gedeckt werden. Bisher wurde hier meist auf Erfahrungswerte aus der Vergangenheit zurückgegriffen. Ziel der Wirkleistungsplanung ist ein Einsatzplan, der für jede Anlage einen Fahrplan festlegt mit dem in Summe genau der erwartete Lastgang gedeckt wird. Übliche Intension ist es, den möglichst kostengünstigsten Mix der zur Verfügung stehenden Mittel zu finden.

Um einen solchen Plan erstellen zu können, ist es notwendig, von jeder Anlage die zur Verfügung stehenden Möglichkeiten zu kennen. Jede Anlage kann üblicherweise den Verlauf der Erzeugung in gewissen Grenzen variieren und somit verschiedene Fahrpläne als Alternativen für die Planung anbieten. Gegebene technische Beschränkungen, aber z. B. auch ökonomische Erwägungen, hingegen verbieten die Umsetzbarkeit eines Großteils der theoretisch denkbaren Fahrpläne. Als Information muss für die Planung (welche verschiedene Kombinationen von Fahrplänen durchsucht und bewertet) entweder von jeder Anlage die Menge der durch die Anlage realisierbaren Fahrpläne direkt oder aber eine mathematische Beschreibung des Modells der Anlage zur Verfügung stehen, mit welchem die Realisierbarkeit eines Fahrplans geprüft werden kann.

Ein wichtiger Bestandteil der Fragestellung werden die Nebenbedingungen sein, unter denen eine Anlage überhaupt betrieben werden kann. Für diese Arbeit sind hierbei weniger die technischen Anschlussbedingungen, welche die Zulassung zum Anschluss an das Stromnetz regeln, als vielmehr die technischen Rahmenbedingungen von Interesse unter denen eine Anlage in der Lage ist, ihren Erzeugungsfahrplan zu variieren. Analoges gilt für Verbraucher.

## **2.4. Problemstellung der verteilten Lastplanung in dynamischen Verbänden**

Basierend auf den zuvor dargelegten Herausforderungen, die es im Zusammenhang mit der Realisierung eines Smart Grid anzugehen gilt, soll nunmehr die Problemstellung, der sich diese Arbeit widmet abgesteckt werden. Diese stammt aus dem Bereich der verteilten Last- bzw. Einheiteneinsatzplanung und gründet sich ebenfalls auf Fragestellungen des verteilten, optimierten Ausgleichs von Erzeugung, Speicherung und Verbrauch von elektrischer Energie.

Dezentrale Erzeuger sollen sich dynamisch mit Speichern und steuerbaren Verbrau-

chern zusammenschließen können, um gemeinsam einen definierten (ausgeschriebenen) Wirkleistungsverlauf für einen gegebenen Zeitraum nach Fahrplan zu realisieren. Betrachtet werden hierbei also Day-Ahead-Szenarien, wobei der Zeitraum auch kürzer als 24 Stunden sein kann. Annahmen über eine konkrete zeitliche Auflösung werden ebenfalls nicht getroffen. Auch deutlich längere Zeiträume und variable zeitliche Auflösungen sollen Gegenstand der Untersuchungen sein. Essentielle Voraussetzungen für die Erstellung eines optimalen Gesamtfahrplans für eine Gruppe von dezentralen Energieerzeugern sowie Verbrauchern ist die Kenntnis der aktuellen, individuellen Handlungsspielräume (Mengen realisierbarer Alternativlastgänge) der verschiedenen Anlagen. Diese müssen durch die Anlagen selbst prognostiziert und gegebenenfalls kommuniziert werden. Die Prognose muss in den Anlagen selbst erfolgen, da hierfür neben den eigenen anlagespezifischen Betriebsparametern auch die Kenntnis des Bedarfs sowie des aktuellen Betriebszustandes notwendig ist. Ferner sollen Interneta über die eigene Betriebsführung der jeweiligen Anlage möglicherweise nicht offengelegt werden. Der Handlungsspielraum soll Eingang in die Selbstbeschreibung einer Anlage finden, indem die Lastgangmöglichkeiten auf einer Metaebene anlagenunabhängig modelliert werden. Durch eine solche abstrakte Modellierung wird es möglich, Lösungsansätze mit Fokus auf die Optimierungsstrategien zu betrachten ohne Anlagenmodellierungen direkt integrieren zu müssen. Eine solche allgemeinere Problemdefinition ermöglicht auch eine deutlich flexiblere Abbildung der Lösungsansätze auf verschiedene Einsatzszenarien sowie auf variable Verbundzusammenstellungen, Optimierungsziele, Zeiträume und zeitliche Auflösungen sowie eine bessere Anpassung an neue Märkte im Rahmen zunehmender Dezentralisierung. Zudem wird durch die hierdurch ebenfalls verstärkte Modularisierung die Wartung deutlich verbessert, da nachträgliche Änderungen oder Erweiterungen eines Optimierungsansatzes keinerlei Änderungen aufgrund von in das Verfahren integrierten Anlagenmodellierungen erforderlich machen. Die Integration neuer, bisher unbekannter Anlagen muss standardisiert möglich sein.

Das Problem, das in dieser Arbeit vorrangig untersucht und gelöst werden soll, lässt sich zunächst wie folgt an einem Beispiel beschreiben: Eine heterogene Menge von steuerbaren, dezentralen Erzeugern (bspw. Mikro-KWK Anlagen) und steuerbaren Verbrauchern (bspw. adaptive Kühlgeräte) soll in enger Zusammenarbeit mit in den Verbund integrierten Speichern ihre individuellen Fahrpläne für einen gegebenen Zeithorizont so aufeinander abstimmen, dass der Summenlastgang einen gewünschten Wirkleistungsverlauf annimmt. Einschränkungen bzgl. der Realisierbarkeit einzelner Fahrpläne müssen hierbei ebenso berücksichtigt werden wie individuelle Präferenzen und die Kosten oder der Nutzen der einzelnen Alternativen für die jeweiligen Anlagen. Der Spezialfall bei dem Erzeugung und Verbrauch möglichst exakt aufeinander abgestimmt werden (Supply-Demand-Matching) ist hierin bereits enthalten. Der gewünschte Summenlastgang ist in diesem Fall die Nulllinie. Der Suchraum, in dem nach dem Optimum gesucht wird ergibt sich durch die Überlagerung der einzelnen Handlungsspielräume der verschiedenen Anlagen. Der Begriff Handlungsspielraum bezeichnet hierbei die Menge der verschiedenen, alternativen Fahrpläne (Erzeugung und/ oder Verbrauch), die eine Anlage für einen gegebenen Zeitrahmen fahren könnte ohne hierbei gegebene Restriktionen (Leistungsbeschränkungen, Pufferspeicherfüllstand o. ä. ) zu verletzen: Der Handlungsspielraum ist somit der Suchraum *einer* Anlage. Dieser Raum realisierbarer, alternativer Fahrpläne muss im Rahmen einer Optimierung möglichst allen am jeweiligen Verfahren beteiligten Einheiten in geeigneter Form zur Verfügung stehen, da er Teil des Suchraums ist. Zumindest muss das Optimie-

ungsverfahren in irgendeiner Form Zugriff haben. Das gleiche gilt für die Zuordnung von lokalen Indikatoren (z. B. Kosten oder Präferenzen) zu den verschiedenen Lastgängen. Andernfalls sind globale Optimierungsziele nur schwerlich zu erreichen.

Im Falle eines verteilten (auf einem Agentenansatz beruhenden) Lösungsansatzes muss jede der dezentralen Einheiten den Raum der möglichen, alternativen Handlungsweisen selbst bestimmen und gegebenenfalls (für global optimale Lösungen) kommunizieren können, da nicht jeder an der Einsatzplanbestimmung beteiligte Agent eine Modellierung aller Anlagen praktikabel vornehmen kann, um deren Suchraum zu bestimmen. Offen ist derzeit noch die Frage nach einem geeigneten Optimierungsansatz mit dem die als Meta-Modell kodierten Handlungsspielräume nach geeigneten Lastgängen durchsucht werden können. Zudem steigt die Suchraumgröße mit wachsender Zahl an Anlagen sehr schnell an. Gefragt sind hierbei Optimierungsansätze, welche auf der Selbstbeschreibung von Fahrplänen (Meta-Modellierung der Constraints) der einzelnen Akteure beruhen und die mit diesen Beschreibungen direkt arbeiten können. Hier sollen parallel zwei allgemeine Lösungsansätze betrachtet werden. Bei einem zentralistischen Ansatz würde ein Mitglied der Gruppe mit genügend Rechenleistung eine Führungsrolle übernehmen und das Optimierungsproblem lösen. Dieser Ansatz dient auch als Referenz. Im Rahmen dieser Arbeit werden aber auch verschiedene dezentrale Ansätze untersucht, bei denen die Rechenleistung intelligent auf die verschiedenen Knoten verteilt wird. Zu denken wäre hierbei beispielsweise an verteilte, individuenbasierte Algorithmen. Der hier skizzierte Anwendungsfall lässt sich auch verallgemeinert ausdrücken: Eine Menge von verschiedenen Einzelakteuren soll die individuellen Handlungsweisen aufeinander abstimmen, um in Summe ein gewünschtes Verhalten aufzuweisen.

Im Rahmen des Projektes Smart Nord (<http://smarnord.de/>) werden begleitend und weiterführend nicht nur verschiedene Optimierungsalgorithmen für die Wirkleistungsplanung, sondern auch unterschiedliche Organisationsformen zur Koordination untersucht. Alle betrachteten Lösungsansätze werden nach einer theoretischen Betrachtung implementiert und in einer Simulationsstudie einander gegenübergestellt. Im Rahmen dieser Arbeit sollen Antworten auf die folgenden Forschungsfragen gefunden werden:

- Wie kann der Suchraum der durch eine Anlage realisierbaren Fahrpläne so modelliert werden, dass von konkreten Anlagenmodellen und Constraints abstrahiert werden kann?
- Welches Vorgehen erlaubt eine günstige (möglichst universale) Einbindung des Suchraummodells in verschiedene Optimierungsverfahren? Aufgrund der Komplexität durch die Dimensionalität der zu erwartenden Problemstellungen sind hier insbesondere Heuristiken zu berücksichtigen.
- Was sind die Randbedingungen an ein solches Metaverfahren? Welche Einschränkungen müssen an die Abbildbarkeit auf verschiedene Problemstellungen gemacht werden? Kann im Wirkleistungsplanungsfall die Parametrisierung automatisiert erfolgen?

Als Hauptergebnis liegt abschließend eine vollständige, mathematische Beschreibung des Konzeptes (einschließlich des formalen Vorgehens zur Abbildung auf konkrete Problemstellungen) für die Lösung von Problemen aus der eingangs vorgestellten Problemklasse vor. Dieses Konzept ist als Metavorgehen abbildbar auf verschiedene Probleminstanzen.

Zur Evaluierung wird eine Abbildung auf das skizzierte Problem zur Wirkleistungsbereitstellung verteilter Anlagen vorgenommen.

## 2.5. Zusammenfassung

Auch wenn das zu entwickelnde Konzept für die gemeinschaftliche Optimierung der Gesamthandlung individueller Akteure mit verteilten, beschränkten Handlungsspielräumen nicht auf konkrete Anwendungsfelder ausgelegt ist, sondern sich auf verschiedene Fälle abbilden lassen wird, muss für eine Evaluierung ein konkretes Szenario gefunden werden.

In diesem einleitenden Kapitel wurde zunächst als späteres Anwendungsfeld das Konzept Smart Grid erklärend abgesteckt. Die Wirkleistungsplanung, also die Erstellung eines geeigneten und hinsichtlich gegebener Kostenkriterien günstigen Einsatzplans der jeder Anlage eines Verbundes einen auch unter Nebenbedingungen realisierbaren Fahrplan zuordnet, ist ein wichtiger Baustein einer auf Verbundbildung basierenden Smart Grid Realisierung. Große Bedeutung kommt hierbei der Gültigkeit des Planungsergebnisses zu. Ein Einsatzplan, der Fahrplanvorgaben für Anlagen enthält, die von den Anlagen aufgrund technischer Beschränkungen nicht umgesetzt werden können ist in Gänze nutzlos. Im Vergleich zu den früheren Großkraftwerken ist aber gerade diese Umsetzbarkeit eines Fahrplans für kleinere Anlagen wie BHKW mit Wärmespeichern als Subsystem, Wärmepumpen o. ä. schwerer zu modellieren. Insbesondere steigt die Diversität der Anlagen und damit auch die der zu berücksichtigenden Constraints.

Ohne eine abstrakte Beschreibung der Suchräume (Menge realisierbarer Fahrpläne) bzw. der beschränkenden Constraints wird es schwerlich möglich eine allgemeingültige Formulierung des Wirkleistungsproblems umzusetzen, bei der nicht a priori die Menge der Anlagen in einem Verbund (und damit auch deren Modellierungen für die Planung) bekannt sein muss. Sich dynamisch zusammenfindende Verbünde sind aber darauf angewiesen auch vorab unbekannte Anlagen (auch unbekanntem Typs) aufzunehmen und in die Planung zu integrieren.

Hierzu wird ein abstrakter Modellierungsansatz nebst universeller Integration in Planungs- und Optimierungsalgorithmen benötigt. Die Entwicklung eines solchen Ansatzes wird Ziel dieser Arbeit sein.

## 3. Verwandte Arbeiten

*When you look at the dark side,  
careful you must be  
...for the dark side looks back;*

---

*(Yoda)*

### 3.1. Einleitung

Das hier betrachtete Problem der Wirkleistungsplanung ist nicht neu und wird beispielsweise als Kraftwerkseinsatzplanung (engl.: unit commitment problem) seit langem auch im Rahmen der bisherigen Struktur des Energieversorgungsnetzes betrachtet: Gegeben ist ein erwarteter Lastverlauf für beispielsweise den nächsten Tag und eine Menge eigener Kraftwerksblöcke soll kostenoptimal für diesen Einsatz verplant werden. Meist werden hier jedoch längere Zeiträume betrachtet.

Mehr oder weniger statische Szenarien mit einer überschaubaren Anzahl (eigener) Kraftwerksblöcke mit bekannter Modellierung machten bisher die Aufstellung zentraler, statischer Modelle für die Optimierung einfach. Eine Änderung der Kraftwerkskonfiguration war eher selten nötig und wenn, dann eher einfach zu modellieren. Mit der Notwendigkeit der Integration einer Vielzahl dezentraler, eher kleiner Erzeugungsanlagen geht wie im vorigen Kapitel erläutert auch die Notwendigkeit einher, neue Konzepte für die Wirkleistungsplanung und damit für die Koordination der verteilten Anlagen zu erforschen. Allein die große Zahl zu steuernder Anlagen verbietet aufgrund der entstehenden Problemgröße bereits die Verwendung konventioneller Ansätze.

Beinahe alle aktuell diskutierten Koordinationsansätze für dezentrale Erzeugungsanlagen (auch gemeinsam mit steuerbaren Verbrauchern, oder so genannten Prosumern, z. B. Speichern) zielen darauf ab, Hierarchien aufzubauen oder Anlagen zu gruppieren. Gute Übersichten über bestehende Ansätze geben u. a. beispielsweise (Tröschel 2010; Lünsdorf 2012b). Das bekannteste Konstrukt in diesem Zusammenhang ist das Virtuelle Kraftwerk (VK), welches einen informationstechnischen Zusammenschluss einer Menge von Anlagen darstellt, damit diese so gesteuert werden können, dass sie nach außen wie ein einzelnes (Groß-)Kraftwerk wirken.

Nach einem allgemeinen Überblick und einer Beschreibung des Unit Commitment Problem im Allgemeinen, wird daher zunächst ein Blick auf die Steuerungs- und Planungskonzepte von Virtuellen Kraftwerken geworfen. Eine Betrachtung von auf Agentensystemen fußenden Ansätzen schließt dieses Kapitel ab. Der besondere Fokus liegt in diesem Kapitel auf der Eignung der Ansätze für dynamische Zusammenschlüsse von Anlagen und damit auf der Eignung für dynamisch erst zur Laufzeit feststehende Optimierungsmodelle für

neu gebildete Anlagengruppen. Die Beachtung beliebiger Anlagentypen und insbesondere auch ihrer individuellen Constraints ist ein weiteres, wichtiges Kriterium.

## 3.2. Optimierte Wirkleistungsbereitstellung

Das berechnungsschwere Problem zur optimalen Wirkleistungsplanung kann durch verschiedene Architekturen angegangen werden. Im einfachsten Fall findet ein zentrales Verfahren Anwendung. Hierbei ist eine einzelne Instanz (Agent, Zentralrechner, o. ä. ) für die gesamte Optimierung zuständig. Sie muss entsprechend über alle notwendigen Informationen verfügen – auch darüber, welche Anlage welche Fahrpläne umsetzen kann. Diese einfache Architektur ist nur für eher kleine Problemgrößen mit wenigen Anlagen praktikabel umsetzbar.

Eine Verteilung der Rechenlast kann auf verschiedene Weise geschehen. Hierarchische Ansätze zerlegen das Gesamtproblem in Teilprobleme indem verschiedene Ebenen betrachtet werden. Eine derartige Zerlegung ist aufgrund der hierarchischen Struktur des elektrischen Versorgungsnetzes durchaus natürlich: zuerst erfolgt eine Planung auf Teilnetzebene, dann wird (die gleiche Planung) auf nächsthöherer Ebene mit aggregierten Teilnetzen durchgeführt usw. Auf der jeweiligen Ebene ist meist wieder eine zentrale Instanz zuständig. Vorteil ist die Zerlegung des großen Problems in handhabbare Teilprobleme gleicher Art. Dezentrale Ansätze schließlich zerlegen das Gesamtproblem der Wirkleistungsplanung in kleine anlagenspezifische Teilprobleme, die von in den Anlagen integrierten kleinen Recheneinheiten weitgehend parallel bearbeitet werden können. Kommunikationskanäle sorgen für die Abstimmung zwischen den Anlagen. Bei diesem Modell existieren weder eine koordinierende Instanz noch ein zentraler Speicher. Der Vorteil liegt in der Skalierbarkeit, da mit wachsender Problemgröße durch weitere Anlagen auch immer die entsprechende Rechenleistung für dieses neue Teilproblem hinzukommt. Im Folgenden sollen einige solcher Ansätze mit ihren Vor- und Nachteilen näher betrachtet werden.

Im Rahmen heutiger (zentraler) Kraftwerkseinsatzplanung kommen üblicherweise bereits verschiedene Näherungsverfahren zum Einsatz. Beispiele aus der Forschung hierzu sind (Mao & Li 2008) und (Xiong et al. 2008). Die Bestimmung des exakten, globalen Optimums ist aufgrund von Unsicherheiten in den Daten (Prognosefehler) ohnehin erst ex post möglich. In der Praxis ist in Deutschland und Europa die Software BoFIT der Firma ProCom GmbH (<http://www.procom.de>) weit verbreitet (Franch et al. 2009), welche die Modellierung eines Kraftwerksparks als gemischt-ganzzahliges Modell notwendig macht. Nebenbedingungen wie Mindeststillstandszeiten einzelner Kraftwerksblöcke oder Speicherbilanzen sowie Anlagenprofile sind bei diesen Ansätzen üblicherweise integraler Bestandteil des mathematischen Modells und damit nicht dynamisch austauschbar. Grundsätzlich sind aber die Ansätze zur Optimierung auch auf dezentrale Modelle anpassbar.

Die Koordination dezentraler Erzeuger und Verbraucher in der Absicht einen gewünschten Gesamtwirkleistungslastgang bereitzustellen, ähnelt in der Zielstellung zunächst einem Virtuellen Kraftwerk. Generell können die Koordinationsansätze (das Scheduling) zur Bestimmung individueller Wirkleistungsverläufe für die einzelnen, teilnehmenden Anlagen unterteilt werden in zentrale und verteilte Verfahren. Der zentrale Ansatz dominiert derzeit immer noch die Diskussion um die geeignetste Architektur für Virtuelle

Kraftwerke (Tröschel & Appelrath 2009). Für einen zentralen Ansatz spricht beispielsweise., dass hiermit für die einzelnen Anlagenbetreiber im Mittel ein besserer Gewinn erzielt werden kann, wenn alle Informationen zentral verarbeitet werden können (Gatterbauer 2010). Zentrale Ansätze wurden bisher im Kontext statischer Verbünde von dezentralen Anlagen betrachtet, was allerdings erhebliche Restriktionen bezüglich Skalierbarkeit und Flexibilität nach sich zieht. So gesehen, finden bei der Steuerung Virtueller Kraftwerke bewährte Steuerungsansätze Anwendung, um durch derartige Verbünde das Verhalten konventioneller Kraftwerke nachzuahmen, wobei allerdings der potenzielle Vorteil der Flexibilität vernachlässigt wird. Im Rahmen von Virtuellen- oder Schwarmkraftwerken wurden aber auch bereits dezentrale Ansätze diskutiert. Beispiele hierfür sind (Kok et al. 2008; Wedde et al. 2007; Conrad et al. 2005). Als Steuerungsmechanismen kommen häufig marktbasierende Algorithmen zum Einsatz (Ygge 1998; Ygge et al. 1999).

Dezentrale Ansätze können noch weiter unterteilt werden in kommunikationsbehaftetes Vorgehen und in Ansätze, bei denen die einzelnen Anlagen ihre Entscheidungen für einen bestimmten (i.d.R. eher momentanen) Wirkleistungsverlauf kommunikationslos und allein anhand eines gemessenen Netzzustandes (meist Frequenz oder Spannung) unabhängig von den anderen Anlagen treffen wie beispielsweise in (Kaestle 2007). Derartige Ansätze haben derzeit das Ziel der Ausregelung des momentanen Gesamtwirkleistungsverlaufs und sind so für Szenarien mit vorausschauender Planung weniger geeignet. Das Problem der verteilten Optimierung der Fahrpläne (insbesondere unter Kenntnis des globalen Suchraums) zur Erstellung optimaler Lastgänge innerhalb fluktuierender Verbünde ist jedoch noch nicht zufriedenstellend gelöst. Arbeiten in Bezug auf die verteilte planerische Erstellung von Fahrplänen für einzelne Anlagen in einem Verbund haben bisher oft nur statische Verbünde betrachtet oder zahlreiche Vereinfachungen vorgenommen. Als Beispiel sei hier (Palensky 2001) aufgeführt. Statische Verbünde weisen aber keine hinreichende Flexibilität bzgl. sich ändernder (ggf. von der Verbundzusammensetzung abhängiger) Optimierungsziele bzw. ihrer Zusammensetzung (oder ihrer Topologie) auf. Verbünde werden in bisherigen Arbeiten zu Virtuellen Kraftwerken als vorkonfiguriertes, festes Gebilde betrachtet. Die Anzahl der Anlagen in den Verbänden ist meist eine limitierende Größe bei der Umsetzbarkeit von Optimierungsverfahren. In Zukunft werden aber Lösungen benötigt, die durch geschickte Modellierung auch die Zustandsinformationen einer großen Zahl von dezentralen Anlagen in die Optimierung einbeziehen können. Gleichzeitig muss das Verfahren aber eine hinreichende Flexibilität besitzen, um in Rahmen dynamischer Verbundrekonfigurationen auch bisher unbekannte Anlagen einzubeziehen oder sich an neue Zielvorgaben (z. B. durch den Markt) anzupassen.

Ausgehend von der ursprünglichen Standardformulierung für den zentralen Fall mit wenigen Großkraftwerken als Unit Commitment Problem sollen nachfolgend einige wichtige Lösungsansätze näher betrachtet werden.

#### 3.2.1. Unit Commitment

In der klassischen Kraftwerkseinsatzplanung erfolgt die vorausschauende Planung der bedarfsgerechten Wirkleistungsbereitstellung auf verschiedenen Zeitskalen. Die tägliche Einsatzplanung der Generatoren, genauer gesagt das Scheduling der Ein- und Ausschaltzeitpunkte, ist die Aufgabe des Unit Commitment auf Stundenbasis (Saravanan et al.

2013). Primäres Ziel beim Unit Commitment ist es, die kostengünstigste Kombination von verschiedenen Generatoren zu finden, um eine prognostizierte Nachfrage nach elektrischer Energie zu decken. Die konventionelle Formulierung des Problems lautet (vgl. z. B. Saravanan et al. 2013; Guan et al. 2003; Kurban & Filik 2009):

$$\sum_{t=1}^{N_t} \sum_{i=1}^{N_U} [C_i(P(i,t))I(i,t) + SU(i,I(i,t)) + SD(i,t)] \rightarrow \min. \quad (3.1)$$

Als Kosten gehen hierbei die Brennstoffkosten  $C_i$  von Generator  $i$  abhängig von der im Zeitintervall  $t$  erzeugten Leistung  $P(i,t)$ , die Kosten für das Hochfahren  $SU(i,t,I(i,t))$  abhängig vom Betriebszustand der jeweiligen Anlage und die Kosten für das Herunterfahren  $SD(i,t)$  ein.  $I(i,t) \in \{0,1\}$  kennzeichnet lediglich den Betriebszustand AN oder AUS von Generator  $i$  im Zeitintervall  $t$ .

Hierbei sind verschiedene Constraints zu beachten. Als globaler Constraint ist zunächst der Energiebedarf vorgegeben, der zu jedem Zeitintervall innerhalb einer gegebenen Epsilonumgebung durch die Erzeugung gedeckt sein muss. Maximal zulässige Emissionen der Anlagen werden (auch vor dem Hintergrund regulatorischer Vorgaben) gegebenenfalls ebenfalls als globaler Constraint vorgegeben (Saravanan et al. 2013). Netzrestriktionen (Spannungsbänder und Leitungskapazitäten) und die Vorhaltung einer minimalen rotierenden Reserve sind weitere Beispiele. Generatorspezifische Constraints sind üblicherweise Anfahrampen, minimale Lauf- und Standzeiten sowie Kapazitätsbeschränkungen. Alle diese Constraints sind von ihrer Formulierung her für alle Anlagen identisch bei der Integration in das Optimierungsmodell, wobei eine individuelle, anlagentypische Modellierung nicht vorgesehen ist.

Die typische Formulierung des Optimierungsproblems erfolgt als gemischt-ganzzahliges Programm, welche als Spezialfall der linearen Programmierung davon ausgeht, dass alle oder zumindest einige der Entscheidungsvariablen diskret sind. Im Fall der klassischen Formulierung (Generator aus, an oder gegebenenfalls auf einem von wenigen diskreten Leistungsniveaus) ist dies auch der Fall. Constraints sind ebenfalls linear formuliert. Komplexere Constraints sind in dieser Darstellung nicht einfach integrierbar; insbesondere solche nicht, die in ihrer Ausprägung zu einem Zeitpunkt  $t_j$  von Entscheidungen bezüglich der Betriebsweise bzw. deren Auswirkungen abhängen, die zum Zeitpunkt  $t_i < t_j$  den Betrieb der Anlagen beeinflussen. Dies ist beispielsweise bei indirekt über Speicherfüllstandsänderungen (verursacht durch frühere Betriebszustände) von zwischengeschalteten thermischen Pufferspeichern in gekoppelten elektrisch-thermischen Systemen (z. B. Blockheizkraftwerke) der Fall.

Das Unit Commitment Problem ist ein NP-schweres Problem (Guan et al. 2005). Eine Vielzahl verschiedener (exakter, heuristischer und hybrider) Ansätze sind in der Vergangenheit bereits entwickelt und untersucht worden (Saravanan et al. 2013; Bhardwaj et al. 2012; Kurban & Filik 2009). Beispiele sind Prioritätslistenmethoden (Delarue et al. 2013), dynamische Programmierung, Branch and Bound, Integerprogrammierung, Simulated Annealing (Kirkpatrick et al. 1983), Lagrange Relaxation (Guan et al. 2003), Tabu Search (Glover & Laguna 1997) oder das Innere-Punkte-Verfahren. Aber auch Heuristiken wie etwa schwarmbasierte Verfahren wurden untersucht (Xiong et al. 2008). Ein häufiges Problem der Verfahren ist die Skalierbarkeit bei wachsender Problemgröße. So identifizierten Saravanan et al. (2013) beispielsweise auch, dass aufgrund der Constraints viel Zeit durch die

Evaluation ungültiger Lösungen verwendet wird. Eine Modellierung welche ausschließlich zulässige Lösungen betrachtet wäre hier also von Vorteil.

Die Deregulierung des Marktes bringt weitere Schwierigkeiten durch komplexere (auf verteilten Informationen aufbauende) Zielfunktionen mit sich (Saravanan et al. 2013); siehe hierzu auch (Gatterbauer 2010). Folglich identifizieren Saravanan et al. (2013) ein unvermindertes Forschungsinteresse und weiteren Forschungsbedarf, um steigende Problemgrößen, weitere Anwendungsfälle und eine wachsende Verteiltheit des Problems durch Marktintegration und die steigende Unsicherheit in den Prognosen in den Griff zu bekommen.

#### 3.2.2. Virtuelle Kraftwerke

Das operative Management eines Energiesystems beinhaltet eine Reihe komplexer Aufgaben von der technischen Überwachung bis hin zu organisatorischen und geschäftlichen Maßnahmen. Traditionell wird dieses Management zentral umgesetzt. Aufgrund der schieren und stetig weiterwachsenden Zahl verteilter Energiesysteme (Erzeuger wie auch Betriebsmittel und flexibilisierbare Verbraucher) in einem sich vom statischen zu einem dynamisch rekonfigurierenden Energiesystem individueller Entscheidungsträger wandelnden Verteilnetz (Ilić 2007), lässt es unwahrscheinlich erscheinen, dass die etablierten zentralen Steuerungsansätze mit der rapide wachsenden Problemgröße schritthalten können. Folglich wurde von Wu et al. (2005) ein dringender Bedarf für eine dezentrale Kontrolle aufgezeigt.

Von der International Energy Agency (IEA) kam bereits Ende der Neunziger der Vorschlag für einen virtuellen Energieversorger (virtual utility) als möglicher Weg zur verteilten Kontrolle (Awerbuch & Preston 1997; International Energy Agency 2002). Folgerichtig wurde aufbauend auf diesem Konzept das Konstrukt des Virtuellen Kraftwerks sowohl für technische als auch kommerzielle Betriebsziele entwickelt und intensiv erforscht (Lukovic et al. 2010; Tröschel & Appelrath 2009). Eine Liste von Umsetzungen findet sich beispielsweise in (Coll-Mayor et al. 2004; Nikonowicz & Milewski 2012).

Das Konzept Virtuelles Kraftwerk lässt sich beschreiben als informationstechnischer Zusammenschluss einer Menge von dezentralen Energiesystemen (Tröschel 2010). Grundsätzlich können in einem VK sowohl Erzeuger, Verbraucher als auch Prosumer (meistens Batteriespeicher) zu einer Gruppe zusammengefasst sein; aber auch reine Erzeuger-VK sind üblich. Die ursprüngliche Idee hinter einem solchen Zusammenschluss ist die Bündelung der verteilt vorhandenen Kapazitäten zu einem gemeinschaftlich steuerbaren Ganzen, um so ein den herkömmlichen Großkraftwerken vergleichbares Pendant zu erhalten (Bitsch et al. 2002). Ziel ist also die Koordination einer stetig wachsenden Zahl dezentraler Erzeuger auf IKT-Ebene (Tröschel 2010).

Mittlerweile wird der Begriff Virtuelles Kraftwerk je nach konkretem Anwendungsfall für verschiedene Konstrukte verwendet. Die diskutierten Konzepte unterscheiden sich nach ihrer Architektur ihrer Zielsetzung oder auch nach der Konfiguration der beteiligten Anlagen. Einige exemplarische Sichten sind (Nikonowicz & Milewski 2012; Tröschel 2010):

**Spitzenlastreduktion** Die Bereitstellung von Leistung für das Abfangen von Lastspitzen ist immer mit hohen Kosten verbunden (Malinowski & Kaderly 2004). Die Reduktion

solcher Spitzen kann somit zu einem Verzicht eines gewissen Anteils von Spitzenlastkraftwerken führen. Durch gezielten Betrieb dezentraler Erzeugung wird daher versucht Spitzenlasten (insbesondere im Verteilnetz) entgegenzuwirken (Tröschel 2010; Pielke & Kurrat 2008). Aber auch eine gesteuerte Beeinflussung des Verbrauch wurde bereits untersucht (z. B. Gutschl & Stigler 2006). Üblicherweise kommt hierbei wie auch in (Pandiaraj et al. 2002) oder aktueller in (Soltani et al. 2012) immer noch eine zentrale Steuerungsumsetzung zu Einsatz.

**Microgrids** In Verteilnetzen mit hohem Anteil dezentraler Erzeugung und größenordnungsmäßig vergleichbarem Bedarf können die Anwendungsfälle für Virtuelle Kraftwerke auch auf einen lokalen Ausgleich der Energie abzielen (Kok et al. 2008; Weidelt et al. 2008). Beispiele hierfür sind (Akella et al. 2010; Kamphuis et al. 2007b). Ziel hierbei wäre es somit den Im- bzw. Export elektrischer Energie aus/ an übergelagerte(n) Netzschichten zu minimieren. Man spricht bei einem zumindest auf teilweise Autarkie abzielende Betriebsweise auch von Microgrids (Schwaegerl & Tao 2014; Schlebusch et al. 2006; Woldt 2008). Der Inselbetrieb (z. B. Kalenda 2003) als Spezialfall des Microgrid erfordert zudem auch die Übernahme von Systemdienstleistungen durch die vorhandenen dezentralen Anlagen (Nieße et al. 2012a). Erfolgt eine Umsetzung noch oft mit zentralen Ansätzen, so sind in diesem Bereich doch bereits auch verteilte oder hierarchische Architekturen als Lösungsansätze zu finden (Tröschel 2010; Kamphuis et al. 2007b; Wedde et al. 2007). Meist arbeiten derartige VK beim Ausgleich (Supply-Demand-Matching) eher reaktiv als planerisch.

**Regelleistung** Soll konventionelle Kraftwerkskapazität durch dezentrale Erzeuger ersetzt werden, so muss dies insbesondere auch die Bereitstellung von Systemdienstleistungen einbeziehen (Tröschel 2010). Hierzu gehört auch die Bereitstellung von Regelleistung (Buchholz et al. 2008; Strese 2006). Der Fokus solcher Virtuellen Regelkraftwerke liegt auf der Planung einer ausreichenden Vorhaltung von Energie für den Fall des Abrufes. Hauptgrund für die Aggregation von Anlagenleistung liegt in der notwendigen Präqualifikation bei der der sichere Abruf gewisser Mindestmengen verlangt wird (Kurscheid 2009).

Abgesehen von diesen Spezialaufgaben gehört aber auch die Übernahme klassischer Kraftwerksaufgaben bei der Bereitstellung von Wirkleistung nach Fahrplan für (beispielsweise) den folgenden Tag zum Aufgabenportfolio des Virtuellen Kraftwerks (Lukovic et al. 2010). Beispiele für diese Planungsaufgabe finden sich in (Nielsen et al. 2001; Pantoja et al. 2011; Tröschel 2010). Neuere Ansätze berücksichtigen hier bereits mehr und mehr die geforderte Integration von Marktstrukturen (vgl. Nieße et al. 2014). Äquivalent hierzu kann auch die prinzipiell gleiche (bis auf Vorzeichen) Frage nach dem Scheduling einer Gruppe von Verbrauchern betrachtet werden, die ihren Verbrauch auf eine gegebenen Erzeugungskurve (z. B. prognostizierte PV-Einspeisung) einstellt. Beispiele mit verschiedenen Lösungsansätzen (zentral und verteilt) geben (Caron & Kesidis 2010a) sowie (Jiang & Low 2011).

Während VK eher statische Konstrukte darstellen werden in der Forschung auch Ansätze diskutiert, die Organisation dezentraler Energieanlagen selbstorganisiert beispielsweise durch Multiagentensysteme vorzunehmen. Diese werden im nächsten Abschnitt eingehender betrachtet.

### 3.2.3. Verteilte Ansätze

Mit dem kommunikationstechnischen Zusammenschluss verteilter Anlagen allein ist zwar eine Grundlage für Koordination vorhanden; diese benötigt aber noch einen entsprechenden Mechanismus (Meister 2012). Hierzu bedarf es entweder einer zentralen Instanz, welche alle Informationen sammelt und zusammenführt oder aber Entscheidungen dürfen dezentral getroffen werden. In Falle eines zentralen Konstruktes stellt sich einerseits die Frage, wo dieses angesiedelt sein soll und wie der entstehende Mehrwert aufzuteilen ist. In dezentralen Fall wird jedoch ein funktionierender Markt benötigt (Meister 2012).

Üblicherweise adressieren Virtuelle Kraftwerke die längerfristige Bündelung einer mehr oder weniger fixen Menge verteilter Anlagen. Oft werden diese auch nach wie vor zentral gesteuert (Nieße et al. 2014). Seit einiger Zeit jedoch stehen auch verteilte Steuerungsmechanismen im Fokus der Forschung. So untersuchen beispielsweise McArthur et al. (2007) sowie Ramchurn et al. (2012) die Verwendung Agenten-basierter Kontrollmechanismen in der Energieversorgung. Exemplarische Anwendungen finden sich u. a. in (Negenborn et al. 2010; Ramchurn et al. 2011; Anders et al. 2010).

Für große Problemgrößen werden derzeit verteilte (meist Agenten-basierte) Verfahren aus Gründen der Berechenbarkeit, aber auch aus anderen Gründen diskutiert. Ein Beispiel hierfür sind Privacy Aspekte. Einige aktuelle Implementierungen verteilter Kontrollansätze finden sich beispielsweise in (Hinrichs et al. 2013b; Ramchurn et al. 2011; Kamphuis et al. 2007b). Ursprüngliche Arbeiten in dieser Richtung die lediglich zu einem gewissen Grad das Merkmal der Verteiltheit erfüllen sind Ansätze, welche eine hierarchische Organisation der Koordination vornehmen (Hommelberg et al. 2008). Die zwei bekanntesten Beispiele hierfür sind zum einem der PowerMatcher-Ansatz (Kamphuis et al. 2007b; Kamphuis et al. 2007a) und zum anderen der Ansatz DEZENT (Wedde et al. 2007). Beide Ansätze setzen auf Markt- bzw. Auktions-basierte Verhandlungen in einem Multi-Agentensystem. Beide Verfahren arbeiten ohne Vorausplanung und agieren lediglich reaktiv im laufenden Betrieb um sich auf die jeweilige Situation einzustellen. PowerMatcher ist zudem auf den Anwendungsfall des Supply-Demand-Matchings spezialisiert (Kok et al. 2008; Lehnhoff 2010).

Im Falle des PowerMatchers ist für die Steuerung der verschiedenen Anlagen in einem Verbund ist ein spezieller Agententyp zuständig. Dieser Local Device Agent ist für verschiedene Anlagentypen (je nach Steuerbarkeit) speziell und spezifisch ausgestaltet. Eine Planung des Betriebes dieser Anlagen findet nicht statt, der Agent gibt auf Basis seines anlagenspezifischen Wissens Gebote ab und steuert abhängig vom jeweils aktuellen Preis. Eine Modellierung der Flexibilitäten der Anlagen als Menge von Handlungsoptionen ist daher bei diesen Ansätzen nicht vorhanden. Es gibt jedoch Ansätze, aus dem beobachteten Verhalten einer Anlage (konkret eines Kühlhauses) ein Modell mittels einer Supportvektormaschine zu lernen, um dem Local Device Agenten als Entscheidungsgrundlage zu dienen ohne eine explizite Modellierung erstellen zu müssen (Platt 2007). Diese Modelle erlauben aber lediglich die Vorhersage von Verhalten (für geringe Zeithorizont) und nicht die Modellierung von Flexibilitäten. DEZENT modelliert ebenfalls keine Fahrpläne zur Vorausplanung, was gegebenenfalls eine hohe Unruhe im Betrieb der Anlagen durch häufige, reagierende Betriebsänderungen zur Folge hat (Tröschel 2010). Eine ausführliche Diskussion beider Ansätze findet sich beispielsweise in (Tröschel 2010).

### 3. Verwandte Arbeiten

---

Vorteile einer dezentralen Ausführung der Planung und Kontrolle in einem Virtuellen Kraftwerk liegen in erster Linie in der Skalierbarkeit. Ein weiterer wichtiger Vorteil liegt in der Möglichkeit, Entscheidungen für oder gegen eine bestimmte Betriebsweise ebenfalls lokal treffen zu können. Die Begründung (bzw. die formale Modellierung hinter einer solchen Entscheidung) muss nicht öffentlich gemacht werden.

Nachteilig hingegen ist der Umstand, dass dezentrale Ansätze meist (aufgrund unvollständiger Information) einen global gesehen schlechteren Erlös für die einzelnen Akteure erzielen (Gatterbauer 2010). Verteilte Koordination erfordert immer Abstimmung durch Kommunikation. Der Aufwand hierfür ist höher als bei zentralen Ansätzen. Ein entscheidender Nachteil speziell bei den sukzessiv (abschnittsweise) planenden Verfahren wie beispielsweise den rein reaktiven hierarchischen Ansätze dürfte jedoch sein, dass ein Teil der Möglichkeiten gar nicht berücksichtigt wird, da eine Modellierung mit Projektion für nachfolgende Zeiträume nicht erfolgt. Entscheidungen, die zu einem gegebenen Zeitpunkt für die Betriebsweise getroffen werden (wenn beispielsweise ein Pufferspeicher gefüllt wird) haben Einfluss auf die Möglichkeiten in späteren Zeitabschnitten. Wird dieser wechselseitige Einfluss nicht in der Modellierung berücksichtigt, so wird ein Teil der Möglichkeiten nicht in der Planung berücksichtigt. Legt sich ein sukzessiv planendes Verfahren zu einem Zeitpunkt  $t_i$  für eine bestimmte Betriebsweise fest, so schränkt es die Möglichkeiten für den Zeitpunkt  $t_{i+x}$  ein. Wird anschließend für  $t_{i+x}$  geplant, so können keine Alternativvarianten aus  $t_i$  mehr einbezogen werden. Auch für ein verteiltes Verfahren ist es also vorteilhaft, eine vorausschauende statt einer reaktiven Planung durchzuführen.

Verteilte Organisation und Kontrolle ist insbesondere auch ein Merkmal dynamischer Virtueller Kraftwerke (Nieße et al. 2014) wie sie aktuell im Forschungskontext des Forschungsverbundes Smart Nord (<http://smarnord.de>) stehen, in den auch diese Arbeit eingegliedert ist. In derartigen VK finden sich die verteilten Anlagen dynamisch zu Verbänden zusammen, die speziell auf die Erfüllung eines Stromproduktes ausgelegt sind. Nach Lieferung des Stromproduktes zerfallen diese Verbände sofort wieder. Hierdurch wird speziell der geforderten Marktintegration Rechnung getragen.

In dynamischen VK ergibt sich für die Wirkleistungsplanung eine Einteilung in verschiedene Prozessphasen (Nieße et al. 2014). In der ersten Phase muss sich die Menge aller Anlagen selbsttätig in Gruppen einteilen, je nach aktuell am Markt nachgefragten Stromprodukten. Nach dem drauffolgenden Marketclearing steht für jede Gruppe eine Verpflichtung zu einem bestimmten Stromprodukt fest. Für dieses Stromprodukt muss je Gruppe eine interne Planung erfolgen (Phase 3), welche einen bezüglich mehrerer Ziele optimalen Einsatzplan für die Lieferung erstellt. Im letzten Schritt (Phase 4) schließlich müssen die Anlagen kontinuierlich (bis zum Abschluss der Stromproduktlieferung) überwachen, ob die ursprüngliche Planung eingehalten werden kann oder ob gegebenenfalls Neuplanungen stattfinden müssen.

In Smart Nord findet die Wirkleistungsplanung (wie auch die Verbundbildung) völlig dezentral statt. Als Voraussetzung wird eine Möglichkeit benötigt, auf zulässige Fahrpläne zugreifen zu können. Da ein und dieselbe Agentenimplementierung mit beliebigen Anlagen zusammenarbeiten können soll, müssen diese zulässigen Fahrplanmengen abstrakt formuliert sein. Eine zentrale Constraintfassung (bezüglich der Anlagenmöglichkeiten) müsste anlagenspezifisch erfolgen und scheidet als Ansatz aus. Für die Phasen 1,3 und 4 wird somit eine Modellierung der Flexibilitäten der Anlagen benötigt, welche Mengen von

Fahrplänen und ihre Bewertung hinsichtlich verschiedener Kriterien so kodiert, dass verschiedene Planungsverfahren hieraus einen Fundus an umsetzbaren Fahrplanalternativen schöpfen können ohne (aufgrund der dynamischen Zusammensetzung) die zugrundeliegenden Formalia kennen zu müssen.

### 3.3. Kritische Diskussion verwandter Arbeiten

Abschließend sollen in diesem Abschnitt einige konkrete Ansätze für die Planung und insbesondere für die zugrundeliegende Modellierung kritisch diskutiert werden.

Zwei Merkmale sind bei den verschiedenen Ansätzen näher zu betrachten: Zum einen stellt sich die Frage nach einer möglichst individuellen Modellierung der verschiedenen Anlagen sowie ob und wie die individuellen Möglichkeiten mit denen eine Anlage betrieben werden kann erfasst und dem Planungsverfahren zugänglich gemacht werden können. Hierzu gehört insbesondere auch die Integration einer Modellierung individueller Constraints, welche die möglichen Betriebsweisen einer Anlage einschränken. Zum anderen stellt sich die Frage nach der Möglichkeit ein Modell des Optimierungsproblems für die Wirkleistungsplanung für die Gruppe als ganzes dynamisch und automatisiert erstellen zu können ohne a priori eine statische Modellierung vorgeben zu müssen.

Ein Ansatz zum Finden optimaler (im Hinblick auf Preissignale) Startzeitpunkte für eine Reihe verschiebbarer Tasks wird in (Goudarzi et al. 2011) vorgestellt. In diesem Paper werden eine ganze Reihe vereinfachender Annahmen gemacht. Es gibt nur verschiebbare, nicht unterbrechbare, voneinander unabhängige Jobs. Als Preissignal wird ein so genannter Time of Use Tarif angenommen, bei dem zu fest eingeteilten Zeitzonen über den Tag verschiedene Preise gelten. Als einziger Constraint kann eine Deckelung der Gesamtlast für verschiedene Zeitpunkte angegeben werden. Als eine zweite Zielgröße wird die Minimierung einer artifiziellen Inkonvenienz-Größe angestrebt, welche sich als Abstand zu einem vorgegebenen präferierten Startzeitpunkt bestimmt. In einer zweiten Variante wird statt der Deckelung ein variabler Strompreis (abhängig von der Gesamtleistung) als NLP-Formulierung untersucht. Für die Lösung der Problemstellungen werden verschiedene Standardverfahren betrachtet. Eine Integration anderer Anlagentypen ist nicht möglich. Alle Jobs sind a priori bekannt. Solche Job-Scheduling Ansätze bei denen sich die verschiedenen Fahrpläne im Suchraum einer Anlage (Erzeuger können prinzipiell genau wie Verbraucher behandelt werden) durch Verschieben des Startzeitpunktes eines fest vorgegebenen Lastgangs entlang der Zeitachse ergeben sind ein häufiger betrachteter Ansatz (Caron & Kesidis 2010b; Shah 2012a; Lee et al. 2012). Wenngleich hierbei alle Jobs (es gibt nur einen Typ) gleich modelliert und behandelt werden können, erfolgt keine integrierte Betrachtung individueller Kosten. Caron & Kesidis (2010a) geben ein Beispiel für eine verteilte Variante dieses Ansatzes jedoch mit den gleichen Einschränkungen. Eine Erweiterung auf andere Flexibilisierungsmodelle gestaltet sich hier ebenfalls schwierig.

Im Bereich der Modellierung von dezentralen Energieanlagen finden sich meist sehr spezifische Arbeiten, welche genau einen Anlagentyp hochgenau modellieren. Beispiele hierfür finden sich in (Chassin et al. 2008; Milano 2010). Ziel solcher Simulationsmodelle ist aber meist nicht die Integration in Optimierungsmodelle. Derartige Anlagenmodelle bilden in der Regel ausgehend von äußeren Randbedingungen das zu erwartende Verhalten (also

### 3. Verwandte Arbeiten

---

den einen zu erwartenden Fahrplan) nach. Aber auch die Analyse von Was-Wäre-Wenn-Szenarien gehört zu den Aufgaben solcher Modelle (z. B. Treur 2013). Für den in dieser Arbeit betrachteten Anwendungsfall sind derartige Modelle nicht geeignet, da sie nicht in der Lage sind eine systematische (gegebenenfalls auch unvollständige) Enumeration zulässiger Fahrpläne zu liefern. Dies ist aber Voraussetzung für das Durchsuchen der Möglichkeiten einer Anlage.

Modelle, die für die Integration in Optimierungsansätze vorgesehen sind finden sich beispielsweise in (Makkonen & Lahdelma 2006). Diese Arbeit ist beschränkt auf BHKW-Modelle. Die Modellierung der Flexibilität eines BHKW erfolgt als Darstellung des zulässigen Bereichs der Erzeugung in der komplexen Ebene (Wirk- und Blindleistung). Die Möglichkeiten der Erzeugung werden als konvexe Fläche dargestellt und beschränken sich auf die jeweils mittlere Leistung in *einem* 1-Stunden-Intervall. Komplexere Anlagenmodelle mit nicht konvexem Zulässigkeitsbereich werden als Dekomposition in mehrere konvexe Teilflächen dargestellt. Die Darstellung der Flächen erfolgt linear über die Angabe der Extrempunkte. Das Polytop stellt immer nur ein Zeitintervall dar, eine Darstellung ganzer Fahrpläne ist nicht vorgesehen. Ein vergleichbarer Ansatz findet sich auch in (Rong & Lahdelma 2004). Eine Modellierung mit polyedrischen Zulässigkeitsbereichen für BHKW in der durch produzierte Wärmemenge und elektrischer Energie aufgespannten Ebene finden sich in (Eichhorn et al. 2005).

Ein Beispiel für die Abstimmung einer Gruppe von Anlagen in einer Siedlung für einen kostenminimalen Betrieb auf Fahrplanbasis gibt (Thoma 2007). Allerdings sind hier nur genau ein BHKW, eine Batterie und wenige PV-Anlagen (als Störgröße) zu betrachten. Eine abstrakte Modellierung erfolgt nicht, der Lösungsansatz ist statisch auf die Problemstellung zugeschnitten.

Ein Ansatz bei dem die Umsetzbarkeit eines Fahrplans während der Optimierung durch direkte Einbindung eines Anlagensimulationsmodells überprüft wird findet sich in (Bremer 2006). Eine abstrakte Modellierung findet nicht statt, lediglich auf Modell-Implementierungsebene ist ein Interface als Abstraktion für die Integration beliebiger Modelle vorgesehen. Der verteilte Koordinationsmechanismus von Hinrichs (2014) hingegen beachtet die Zulässigkeit von erstellten Einsatzplänen auch bereits auf abstrakter Ebene, integriert hierzu jedoch u. a. in dieser Arbeit erstellte Ansätze (Hinrichs et al. 2013b) zur Suchraummodellierung bzw. zum Sampling (vgl. Abschnitt 8.2 auf Seite 102 bzw. Kapitel 13 ab Seite 199).

Palensky (2001) betrachtet ein verteiltes Verfahren, bei dem jeder Agent lediglich lokale Entscheidungen trifft. Aus einer endlichen Menge möglicher eigener Fahrpläne (gegeben als Abfolge von Anlagenbetriebszuständen) wird derjenige ausgewählt, der den Einsatzplan der Gruppe optimiert. Anschließend wird der neue Fahrplan kommuniziert (Broadcast, oder gegebenenfalls auch in zufälliger Abfolge), woraufhin andere Agenten das eigene Verhalten ebenfalls anpassen durch Auswahl aus einem lokalen Fundus von Möglichkeiten. Wie diese (fest hinterlegte) endliche Menge von lokalen Zustandsabfolgen jedoch für beliebige Anlagen erzeugt werden kann, bleibt offen. Ebenso wenig ist eine Berücksichtigung aktueller Gegebenheiten (Wetter, Speicherfüllstand, etc.) vorgesehen.

## 3.4. Zusammenfassung

Der Ansatz des Virtuellen Kraftwerks als Koordinationsinstrument für die Integration kleiner, verteilter Energieerzeuger hat in der Vergangenheit bereits eine gewisse Aufmerksamkeit in der Forschung erfahren. Diverse Koordinationsansätze – sowohl zentral als auch dezentral – werden bereits diskutiert.

Wenn alle Aufgaben der Netzführung auf dezentrale Anlagen übertragen werden sollen – wovon auszugehen ist, da es nicht zuletzt politisches Ziel ist vollständig auf regenerative Versorgung umzusteigen – so gehört hierzu auch die vorausschauende Wirkleistungsplanung, um Erzeugung und Verbrauch soweit wie absehbar aufeinander abzustimmen, sodass notwendige reaktive Maßnahmen nur noch unvorhergesehene Abweichungen ausregeln müssen. Soll eine Koordination auf Basis vorausschauender Einsatzplanung erfolgen, so ist eine Modellierung von Flexibilitäten, also Mengen von durch die individuellen Anlagen tatsächlich realisierbaren Fahrplänen unabdingbar. Eine solche Modellierung wird bei den heutigen Ansätzen nicht oder nur unvollständig vorgenommen. Zentrale Ansätze modellieren zwar vollständig, sind aber unflexibel da ihre Modellierung immer nur für einen fest gegebenen Verbund gültig ist. Es fehlt eine Abstraktionsebene mit der eine solche Modellierung dergestalt vorgenommen werden kann, dass jederzeit prinzipiell beliebige Suchräume, welche nur genau die zulässigen alternativen Fahrplanmöglichkeiten einer Anlage modellieren, zu einem Optimierungsmodell kombiniert werden können ohne händisches Zutun.

Wie in diesem Kapitel analysiert, findet in den meisten Fällen lediglich eine reaktive Einsatzplanung für den jeweils aktuellen Moment statt oder aber die Planungsmodelle sind so spezialisiert, dass bei Änderung der Verbundzusammensetzung oder einer Änderung irgendwelcher (äußeren) Rahmenbedingungen des Betriebes von Hand ein neues (oder zumindest aktualisiertes) Optimierungsmodell aufgestellt werden müsste. Praktisch alle Ansätze integrieren direkt das spezialisierte Anlagenmodell in den Planungsvorgang.

Der Fahrplanraum in dem nach zulässigen Lösungen zu suchen ist, ergibt sich aus den Constraints die in diesen Arbeiten bisher explizit zu statischen Optimierungsmodellen integriert sind. Ein derartiges Vorgehen ist unflexibel gegenüber Änderungen in der Zusammensetzung eines Verbundes sowie gegenüber Änderungen im Betriebsprofil einzelner Anlagen. Für dynamische Virtuelle Kraftwerke ist ein solches Vorgehen unbrauchbar.

Der in dieser Arbeit entwickelte Ansatz hingegen wird die Modellierung des Fahrplanraums (und damit implizit auch die Modellierung der Constraints die diesen definieren) auf einer abstrakten Ebene ermöglichen. Hierdurch können alle Anlagen aus Sicht des Optimierungsverfahrens (zentral oder dezentral) gleichbehandelt werden. Der Algorithmus für die eigentliche Optimierung benötigt keinerlei Kenntnis der jeweiligen Anlagen und kann alle Modelle standardisiert integrieren und einheitlich ansprechen.

Zentrale Aufgabe beinahe aller Typen von Virtuellem Kraftwerk ist das Lösen der Frage nach der (unter verschiedenen Gesichtspunkten) optimalen Aufteilung der Gesamtaufgabe auf die einzelnen Anlagen in dem Verbund. Diese Planung muss unter Beachtung der individuellen Constraint unter denen die verschiedenen Anlagen operieren können erfolgen. Daher wird im nachfolgenden Grundlagenkapitel das Thema der Constraint-basierten Optimierung einer näheren Betrachtung unterzogen.



## 4. Constraint-basierte Optimierung

*Es gibt drei Gründe für das Leiden:  
Schmerz, Wandel und  
die allumfassende Abhängigkeit von Bedingungen.*

---

*(s. Hl. Tenzin Gyatsho, 14. Dalai Lama)*

### 4.1. Einleitung

Optimierung – oder genauer: die Suche nach einer Parameterbelegung bei der eine gegebene Funktion einen optimalen Wert (Minimum oder Maximum) annimmt – spielt im Rahmen dieser Arbeit eine zentrale Rolle. Primäres Ziel dieser Arbeit ist die Entwicklung und Umsetzung eines neuen Konzepts zur Behandlung von Constraints während einer verteilten Optimierung, so wie es in dem eingangs skizzierten Anwendungsfall für die Wirkleistungsplanung in Smart Grid Szenarien benötigt wird. Aus diesem Grund sollen an dieser Stelle zunächst die wichtigsten Grundlagen zur Optimierung und Constraintbehandlung rekapituliert werden.

Dabei wird die Optimierung selbst jedoch nicht Schwerpunkt sein. Der Fokus dieses Kapitels wird auf bereits bestehenden und erprobten Verfahren zur Behandlung von Constraints während der Optimierung liegen, um die Einordnung dieser Arbeit in den aktuellen Forschungskontext zu ermöglichen.

### 4.2. Grundlagen

Die Schwierigkeiten und die Komplexität praktischer Optimierungsprobleme aus den Bereichen Telekommunikation, Logistik, aus der Finanzwelt, der Produktion und nicht zuletzt auch aus der Energieversorgung haben zur Entwicklung von Optimierungstechniken geführt, die nicht mehr unbedingt auf die völlige Exaktheit des Ergebnisses abzielen sondern sich stattdessen mit der Bestimmung von guten Ergebnissen begnügen. Diese sind dafür aber in akzeptabler Zeit bestimmbar. Solche Heuristiken verwenden häufig Techniken, Ideen und Ansätze, welche aus einer Vielzahl von Forschungsbereichen entlehnt sind. So nimmt beispielsweise das Verfahren Simulated Annealing (Kirkpatrick et al. 1983) Anleihen aus der statistischen Mechanik und der Metallurgie und genetische Algorithmen imitieren Prozesse der Evolutionsbiologie. Ant Colony Optimization (Dorigo & Stützle 2004) – wie es beispielsweise in (Bremer 2006) bzw. in (Dethlefs et al. 2014) zum Finden optimaler Fahrpläne eingesetzt wurde – ist andererseits ein auf Stigmergie basierender Ansatz und imitiert das kooperative Verhalten einer Ameisenkolonie.

#### 4. Constraint-basierte Optimierung

---

Das Gebiet der Optimierung befasst sich im Allgemeinen damit, optimale Parameter für ein (mathematisch beschreibbares) System zu finden. Der Terminus optimal bezieht sich hierbei auf das Minimum (oder äquivalent das Maximum) einer Bewertungsfunktion  $f$ : Finde

$$\mathbf{x} \in \mathbb{R}^n$$

sodass

$$f(\mathbf{x}) \succeq f(\mathbf{x}^*) \quad \forall \mathbf{x}^* \in \mathcal{D}_f,$$

mit der Relation  $\succeq$  in der Bedeutung: besser oder gleich gut.

Alle grundlegenden Ansätze der Optimierung beinhalten im Wesentlichen Konzepte für die Lösungsrepräsentation, das Optimierungsziel und eine Bewertungsfunktion (Michalewicz & Fogel 2004) zum Vergleich und zum gegeneinander Abwägen von Lösungskandidaten.

Es ist sinnvoll, während des Vorgangs der Optimierung zu unterscheiden zwischen der eigentlichen Lösung (diese liegt oft erst im Nachgang vor) und der Lösungskonfiguration, die während der eigentlichen Optimierung eine Lösung repräsentiert. Letztere kodiert für den Algorithmus alternative Lösungskandidaten in einer geeigneten Form und spannt so auch den Suchraum  $S$  auf. Beispielsweise können Lösungen für einfache Lastverschiebungsprobleme im Smart Grid repräsentiert werden als Binärstring (Bremer et al. 2009), wobei jede 1 kodiert, dass ein Gerät eingeschaltet ist und eine 0 ein Ausgeschaltetsein des Gerätes bedeutet. Als Suchraum *einer* Anlage ergibt sich in diesem Fall also  $\{0, 1\}^n$  für  $n$  zu betrachtende Zeitintervalle. Die eigentliche Lösung, welche ja ein Fahrplan ist, mit dem die entsprechende elektrische Anlage gefahren werden soll, entsteht mithilfe der EIN/ AUS Informationen aus der Lösungskonfiguration und zusätzlichen Informationen zur konkreten Leistungsaufnahme bzw. -abgabe. Man spricht hierbei auch von Genotyp und Phänotyp einer Lösung. Zudem lässt sich an diesem Beispiel bereits erkennen, dass in Szenarien mit mehreren zu optimierenden Einheiten (hier eine Menge von elektrischen Anlagen) oft auch die Notwendigkeit besteht, eine Menge von Lösungsrepräsentanten zu verwenden. Jedes Gerät hat aufgrund sehr individueller Eigenschaften einen eigenen Lösungsraum, der aber für alle Geräte einheitlich kodiert wird. Die Lösungskonfiguration während der Optimierung besteht somit aus einer Menge von Lösungsrepräsentanten für die einzelnen Einheiten.

Die Güte einer Lösung (ihre Fitness) wird üblicherweise anhand einer gegebenen Zielfunktion  $f : \mathcal{D} \rightarrow \mathbb{R}$  bewertet. In der Regel wird von dieser Zielfunktion ein Minimum gesucht. Maximierungsprobleme lassen sich jederzeit in ein Minimierungsproblem umwandeln durch Verwendung von  $-f$ .

Die Bewertungsfunktion (auch Fitnessfunktion, Zielfunktion, Hamiltonian oder Loss) schließlich ermöglicht den Vergleich mehrerer Lösungskandidaten indem sie jedem Kandidaten einen Wert in Bezug auf das Optimierungsziel zuweist. Üblicherweise bildet eine solche Bewertungsfunktion einen Lösungskandidaten (bzw. eine Lösungskonfiguration) nach  $\mathbb{R}$  ab und ermöglicht so nicht nur den Vergleich, welcher der beiden Kandidaten besser ist, sondern auch die quantitative Bewertung um wie viel besser der Kandidat ist. Andererseits mag es hilfreich sein, lediglich einen quantitativen Vergleich zwischen zwei Kandidaten durchzuführen, wenn die Bewertung verhältnismäßig zeitaufwändig ist. In einem solchen Fall könnte eine Bewertungsfunktion zwei Kandidaten entgegennehmen

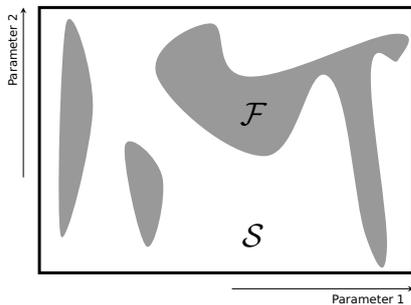


Abbildung 4.1: Lösungsraum der gültigen Lösungen  $\mathcal{F}$  als Unterraum des durch die Parameter aufgespannten Suchraums  $S$ .

und lediglich einen (schneller berechenbaren) Hinweis geben, welcher Kandidat der besserer ist. Eine quantitative Bewertung würde nur von Zeit zu Zeit durchgeführt, um den gesamten Prozess zu beschleunigen.

### 4.3. Evolutionäre Lösungsansätze

Existierende Ansätze zur Optimierung sind vielfältig. Diese können hier nicht erschöpfend behandelt werden. Der Fokus dieser Arbeit wird auf der Integration evolutionärer Optimierungsverfahren liegen. Wie bereits im vorherigen Kapitel dargelegt, erfordert die Komplexität bei der Planung einer optimierten Wirkleistungsbereitstellung durch eine Gruppe von Anlagen den Einsatz spezieller Optimierungsansätze. Heuristiken – Verfahren welche nicht nach dem absoluten Optimum sondern lediglich nach einer guten Lösung suchen, dies aber sehr schnell können – stellen eine naheliegende Alternative dar, um die stetig wachsende Komplexität durch steigende Anlagenzahlen zu beherrschen. Darüber hinaus ist damit zu rechnen, dass aufgrund gegebener Realzeitanforderungen an das System jederzeit ein Abbruch des Optimierungsvorgangs mit einer zumindest gültigen Lösung möglich sein muss. Die Anytime-Eigenschaft (vgl. z. B. Dean & Boddy 1988; Radulescu et al. 2013) von beispielsweise Evolutionären Algorithmen (EA) ermöglicht dies, da bei diesem Typ von Optimierungsverfahren jederzeit eine gültige Lösung vorliegt, die sukzessive weiter verbessert wird. Auf diese Weise ist jederzeit die bis dato beste gefundene Lösung bekannt.

Abgesehen von ihrer praktischen Eignung für den in dieser Arbeit betrachteten Anwendungsfall, sind Evolutionäre Algorithmen auch aus Forschungssicht interessant, da bisher eher wenig Anstrengung in die Integration von Constraintbehandlungstechniken in EA gesteckt wurden (Kramer 2010). Insgesamt stellen Evolutionäre Algorithmen also eine geeignete, interessante Gruppe von Algorithmen dar, um für die Wirkleistungsplanung weiter betrachtet zu werden.

Unter dem Begriff evolutionäre Optimierungsverfahren wird eine Gruppe verschiedener Heuristiken zusammengefasst. Allen gemein ist, dass es sich um stochastische Optimierungsverfahren handelt, deren Funktionsweise inspiriert ist vom Prozess der Evolution. Eine Menge von Lösungen (Population) wird in einem sich wiederholenden Prozess sukzessive verbessert, indem aus der gegebenen Menge von Lösungskandidaten eine nachfolgende Generation erzeugt wird. Allgemein werden hierzu verschiedene Lösungen zu neuen kombiniert (Rekombination). Hierbei werden gegebenenfalls zufällige Verände-

rungen (Mutation) hinzugefügt. Schließlich erfolgt für jede Generation eine Auswahl der Besten durch Evaluation anhand einer Zielfunktion (Selektion).

Ausgehend von ersten frühen Vorarbeiten in den 50er Jahren bei denen evolutionäre Prozesse für die automatische Programmierung erwogen wurden (Friedberg 1958, und Referenzen darin), haben sich in der Folge drei Stränge – im Wesentlichen nebeneinander – entwickelt (De Jong et al. 1997): Evolutionäre Programmierung (EP) (Fogel et al. 1966), Genetische Algorithmen (GA) (Holland 1967) und Evolutionsstrategien (EA) (Rechenberg 1965). Prominente Vertreter sind:

**Genetische Algorithmen:** ~ (Holland 1975) sind die vermutlich populärsten Vertreter evolutionärer Ansätze. Grundgedanke hier ist eine direkte Adaption von Konzepten der biologischen Genetik. Die Evolutionstheorie dient somit als Vorbild. Lösungen eines komplexen Optimierungsproblems werden mit der Fähigkeit zur Selbstmodifikation ausgestattet, um sich im Wettbewerb mit anderen Lösungskandidaten adaptiv so anzupassen, dass sie gemessen an einer Zielfunktion optimal funktionieren (De Jong et al. 1997). Gene kodieren die jeweilige Lösung und spezielle Operatoren kombinieren verschiedene Lösungen zu Nachfolgelösungen. Innovation als Motor des Adaptionsprozesses kommt in Form von Mutationsoperatoren ins Spiel.

**CMA-ES** Die Covariance Matrix Adaptation Evolution Strategy (CMA-ES) (Ostermeier et al. 1994; Hansen 2006) zielt darauf ab, aus vormals erfolgreichen Suchschritten Lehren zu ziehen für das weitere Vorgehen. Die neue Population von Lösungen entsteht bei diesem Verfahren durch Sampeln einer multivariaten Normalverteilung  $N(0, C)$  mit Kovarianzmatrix  $C$ , welche in jeder Iteration basierend auf den vorherigen Sucherfolgen angepasst wird, sodass die Wahrscheinlichkeit von in der Vergangenheit guten Verbesserungsschritten erhöht wird. Das Sampling erfolgt um den Mittelwert (jeweils angepasst durch gewichtetes Mittel einer Auswahl von Lösungspunkten) der Verteilung der vorherigen Generation. Das Verfahren lernt gewissermaßen ein Modell zweiter Ordnung der zugrunde liegenden Zielfunktion, um die Struktur für das weitere Sampling auszunutzen ohne dass der Nutzer vorab eine Parameteranpassung basierend auf Problemwissen vornehmen muss. Eine Beschreibung des Verfahrens findet sich beispielsweise in (Hansen 2011). Das Verfahren ist speziell geeignet für nicht lineare, nicht konvexe Blackbox-Probleme, d. h. der Funktionswert ist nur an einzelnen, evaluierten Punkten bekannt (Hansen 2000). CMA-ES ist für Probleme ohne Constraints oder für Probleme geeignet, die auf so genannte Boxconstraints (für die Parameter ist jeweils ein Intervall als zulässiger Bereich gegeben) beschränkt sind.

**Partikelschwarmoptimierung** ~ (Kennedy & Eberhart 1995) ist das vermutlich bekannteste der populationsbasierten Verfahren, welches das Verhalten im Schwarm als Analogon hat. Die Nachfolgeneration wird erzeugt durch Bewegung der Ausgangspopulation basierend auf einer gewissen Anziehungskraft von guten Lösungen. Vergleichbare Algorithmen dirigieren den Schwarm beispielsweise analog zur Nahrungssuche von Bienen (Karaboga & Basturk 2007).

**Lokale Suche** Bei den Verfahren der lokalen Suche (siehe z. B. Michalewicz & Fogel 2004) wird ausgehend von einer gegebenen Startlösung diese sukzessive verbessert durch betrachten aller Lösungen einer definierten Nachbarschaft und Selektion einer

(im Trivialfall der besten) Nachfolgelösung aus dieser. Verschiedene Strategien zum entkommen aus lokalen Minima kommen hierbei zum Einsatz. Bekannte Vertreter sind die Tabusuche (Glover & Laguna 1997) oder Simulated Annealing (Kirkpatrick et al. 1983).

Jedes dieser Verfahren (wie auch die nicht aufgeführten) hat seine individuellen Stärken und Schwächen, die problemspezifisch bewertet werden müssen (siehe auch Wolpert & Macready 1997). Diese sind an dieser Stelle zunächst nicht von Bedeutung, da im Rahmen dieser Arbeit kein spezielles Verfahren betrachtet werden soll. Allen gemeinsam ist (bis auf spezifische Details und Erweiterungen) der folgende allgemeine Ablauf (Bäck 1997), in den später das in dieser Arbeit entwickelte Verfahren zur Constraintbehandlung eingegliedert werden soll:

- Zunächst wird eine Population aus Lösungsrepräsentanten initialisiert (gegebenenfalls auch nur ein Individuum), wobei jedes Individuum einen Lösungskandidaten repräsentiert.
- Die einzelnen Lösungskandidaten werden bewertet.
- Anschließend wird der Evolutionsprozess wiederholt durchgeführt:
  1. Erzeugung einer neuen Population aus der alten durch Rekombination.
  2. Erzeugung einer Variation der neuen Population durch Mutation.
  3. Bewertung der einzelnen Lösungskandidaten der neuen Population.
  4. Selektion einer Bestenauswahl aus der neuen und der alten Population.

Die Begrifflichkeiten als metaphorische Umschreibungen variieren stark je nach verwendetem Algorithmus und dessen Vorbild in der Natur, die Grundfunktionen sind jedoch immer gegeben (Sörensen 2013).

Eine Beschreibung der einzelnen Verfahren im Detail soll an dieser Stelle nicht erfolgen. Die später für die Evaluation ausgewählten Repräsentanten sind dort (Kapitel 12 ab Seite 177) näher beschrieben. Eine gute Übersicht und Diskussion über existierende Ansätze, ihrer Vor- und Nachteile findet sich beispielsweise in (Bäck et al. 1997; De Jong 2006).

Auch im Bereich des Energiemanagements wurden evolutionäre Algorithmen bereits zahlreich eingesetzt. Einige Beispiele seien hier exemplarisch aufgeführt. So nutzt Palensky (2001) u. a. genetische Algorithmen zur Bestimmung von Fahrplänen für eine Gruppe von Geräten in einem Verbund. Dieser arbeitet auf einem vorab zu erstellenden Zustandsgraphen (Anlagenbetriebszustände), um die Zulässigkeit der Lösung einzubeziehen, nicht auf elektrischen Fahrplänen direkt (Genotyp und Phänotyp). Wie dies automatisiert und Geräte-abstrakt geschehen kann wird nicht diskutiert.

Populationsbasierte Verfahren für das Scheduling von Elektroautos kommen zum Einsatz in (Soares et al. 2013) und in (Faria et al. 2013) für ein Demand Response Management, allerdings ohne die Beachtung irgendwelcher Constraints. Aber auch ein Einsatz als sekundäres Mittel ist denkbar, beispielsweise bei der Parameteroptimierung eines Nadaraya-Watson-Schätzers (Nadaraya 1964) mittels CMA-ES (Kramer et al. 2010) in der Windprognose. Weitere Verfahren sind im Bereich Energiemanagement im Einsatz. In ihrer vollen Breite können diese im Rahmen eines Überblicks hier nicht dargestellt werden. Da-

her wurde hier lediglich eine Auswahl getroffen, um einen ersten Eindruck vom möglichen späteren algorithmischen Einsatzgebiet des hier entwickelten Ansatzes zu geben. Einen guten Überblick über evolutionäre Optimierung im Allgemeinen gibt beispielsweise (Bäck et al. 1997), Einblicke über den Einsatz im Energiebereich finden sich in (Saravanan et al. 2013; Ramchurn et al. 2012; Bremer et al. 2009). Nachfolgend sollen nun Constraints und ihre Integration in die Optimierung betrachtet werden.

### 4.4. Constraints im Kontext Smart Grid

Für viele Probleme ist die Suche nach einer optimalen Lösung bereits ausgesprochen schwierig. Die meisten Problemstellungen aus dem praktischen Bereich bringen aber häufig noch weitere Schwierigkeiten mit sich. Üblicherweise gibt es gewisse Einschränkungen (Constraints) was die Auswahl von Lösungskandidaten betrifft. Wenn eine Lösung diese Einschränkungen nicht beachtet, ist die Folge, dass die gefundene Lösung keine praktische Bedeutung hat, da sie sich nicht umsetzen lässt (Michalewicz & Fogel 2004). Um ein Beispiel aus dem Smart Grid Umfeld anzuführen: Eine Lösung für ein Wirkleistungsplanungsproblem, welche dazu führt, dass eine oder mehrere KWK-Anlagen ihre nebenläufig produzierte Wärme nicht abgeben oder zwischenspeichern können, kann durch die Anlagen nicht umgesetzt werden, da sich die Anlage aus Sicherheitsgründen abschalten würden. Somit würde nicht der durch die Lösung vorgegebene Wirkleistungslastgang produziert werden.

Dies bedeutet zunächst einmal, dass der Raum der zu durchsuchenden, möglichen Lösungen eingeschränkt und somit kleiner wird. Die Folge ist aber nicht, dass die Suche hierdurch vereinfacht wird, insbesondere nicht bei evolutionären Lösungsstrategien. Auf der Suche nach neuen, besseren Lösungen muss sich ein Verfahren durch den Suchraum von einer gültigen Lösung zur nächsten bewegen können. Hier kann die durch Constraints induzierte Geometrie des Teilraums gültiger Lösungen (vgl. Abbildung 4.1) unter Umständen Schwierigkeiten bereiten.

Zunächst sollen aber die verschiedenen Arten von Constraints sowie deren Modellierung betrachtet werden.

#### 4.4.1. Hard Constraints

Constraints werden üblicherweise in zwei Kategorien unterteilt: harte und weiche Constraints (Hard Constraints und Soft Constraints). Die zuvor betrachtete Charakterisierung nach der Constraints dazu führen, dass eine Lösung in der Praxis gar nicht umsetzbar ist, bezieht sich streng genommen auf die so genannten harten Constraints. Der Begriff Hard Constraint bezieht sich in der Regel auf technische Erfordernisse oder auf Sicherheitsaspekte. Dieser Umstand macht ebenfalls deutlich, dass Hard Constraints unter allen Umständen einzuhalten sind.

Beispiele für Hard Constraints aus der Praxis (speziell mit Smart Grid Bezug) sind:

- Zu einer KWK-Anlage (beispielsweise im Heimgebrauch) gehört meist ein thermischer Pufferspeicher zur Zwischenspeicherung aktuell nicht benötigter Wärme (welche bei

der Stromerzeugung zeitgleich immer anfällt). Die Kapazität dieses Speichers (zusammen mit dem aktuellen Füllstand) gibt eine Begrenzung für zukünftige Betriebsdauer (max. bis Völlladung) oder Standzeit (kein Absinken unterhalb Minimalfüllstand) vor.

- Für viele Anlagen gelten gewisse Mindestzeiten, für die sie nach einem (Kalt-)Start in Betrieb sein müssen. Dies gilt beispielsweise für Anlagen mit durch Kompressor betriebenen Wärmetauscher wie Kühlturme oder Wärmepumpen.
- Viele Anlagen können aus dem Stand zunächst nur mit einer begrenzten Beschleunigung ihre Leistung steigern. Solche Anfahrtrampen (ggf. abgestuft) gelten z. B. für BHKW. Aber auch Prozesse wie etwa die Elektrolyse zur Wasserstoffgewinnung, wie sie beispielsweise in der neuen Audi e-Gas Anlage (Audi AG 2013) auch zur Bereitstellung von Regelleistung eingesetzt werden soll, müssen Anfahrtrampen (hier etwa  $30 \text{ As}^{-1}$ ) beachten.

Dies sind nur einige Beispiele, aus anderen Anwendungsgebieten sind zahlreiche weitere Formen von Constraints denkbar, jedoch werden in dieser Arbeit vordringlich Anlageninterne Constraints behandelt. Diese können (z. B. im Rahmen des Job-Scheduling) auch die Ausführungsreihenfolge betreffen. In (Eysholdt et al. 2008) wurden beispielsweise auch Constraints bezüglich der Ausführungsreihenfolge elektrischer Geräte im Rahmen eines Inhouse-Energiemanagements behandelt. In dieser Arbeit werden zunächst Constraints betrachtet werden, die die möglichen Betriebsweisen einer einzelnen Anlage betreffen.

#### 4.4.2. Soft Constraints

Weiche (Soft) Constraints hingegen legen Eigenschaften einer Lösung fest, die man gerne erreichen möchte. Sollte dies nicht oder nicht vollständig gelingen, so wird die Lösung dadurch aber nicht vollständig wertlos, da sie sich immer noch implementieren lässt.

Soft Constraint repräsentieren häufig Benutzervorlieben oder wirtschaftliche Aspekte (eine teurere Lösung – solange sie noch Gewinn abwirft – ist zwar nicht gewollt aber immer noch möglich). Aber auch die Integration von Wahrscheinlichkeiten (z. B. die Ausfallwahrscheinlichkeit einer Komponente) kann über Soft Constraints modelliert sein (Sachenbacher & Williams 2005). Beispiele aus der Praxis für Soft Constraints im Bereich der Smart Grid Lastplanung sind:

- Monetäre Kosten für die Stromerzeugung
- Vorhaltung von Reserven für die Erhaltung von Flexibilität für spätere Planungszeiträume
- Primärenergieverbrauch und Emissionen
- Ähnlichkeit eines Plans für die verteilte Erzeugung zu einem vorherigen Plan (um im Falle einer Umplanung gegebene Netzrestriktionen nicht zu verletzen)

Soft Constraints lassen sich gegebenenfalls als zusätzliche Optimierungsziele integrieren, wodurch allerdings die Verwendung von Verfahren der Mehrzieloptimierung notwendig wird. Die Rolle von Soft und Hard Constraints bei der Optimierung lässt sich folgendermaßen interpretieren: Jede Lösung, welche keines der Hard Constraints verletzt, ist eine

gültige Lösung; diese ist jedoch nicht notwendigerweise optimal bezogen auf Soft Constraints und insbesondere die Zielfunktion.

#### 4.4.3. Globale Constraints

Der Vollständigkeit halber seien an dieser Stelle noch globale Constraints aufgeführt, die eine Gruppe dezentraler Anlagen als Ganzes und nicht nur einzelne Anlagen betreffen. Zwei Beispiele hierfür sind:

- Netzrestriktionen wie Spannungsbänder, Netzfrequenz, thermische Belastbarkeit von Leitungen und vieles mehr schränken die Möglichkeiten für die Umsetzbarkeit eines Einsatzplans für ein Virtuelles Kraftwerk als Ganzes ein. Viele dieser Einschränkungen betreffen nicht zwingend die ganze Gruppe von Anlagen sondern möglicherweise nur Teilmengen, beispielsweise nur einzelne Knoten im Verteilnetz. Dennoch muss die Auswertung im Kontext des gemeinsam genutzten Betriebsmittels Stromnetz evaluiert werden.
- Zu hohe Gradienten – z. B. beim gleichzeitigen Einschalten mehrerer Erzeuger – können ebenfalls zu unzulässigen Überlastungen führen. Die Beachtung des so genannten Gleichzeitigkeitsfaktors ist somit auch ein globaler Constraint, der immer anlagenübergreifend zu betrachten ist.

Diese Constraints lassen sich nicht in die Modellierung einzelner Anlagen unabhängig integrieren, da immer mehrere Anlagen betrachtet werden müssen. Derartige Constraints sind daher hier nicht im Fokus dieser Arbeit.

### 4.5. Constraintbehandlung während der Optimierung

Die Integration von zu beachtenden Constraints wirft eine Reihe von Fragen an das Design eines Lösungsansatzes für ein gegebenes Optimierungsproblem auf. Von Bedeutung für eine effiziente Lösungsstrategie sind u. a. (Michalewicz 1995):

- Wie sollen ungültige Lösungen miteinander verglichen werden? Muss hierfür eine andere Bewertungsfunktion verwendet werden als für den Vergleich zweier gültiger Lösungen?
- Sei  $c_{\mathcal{F}}$  eine Bewertungsfunktion für zulässige Lösungen und  $c_{\mathcal{I}}$  eine Bewertungsfunktion für unzulässige Lösungen. Gilt  $c_{\mathcal{F}}(x) \succ c_{\mathcal{I}}(y) \forall x \in \mathcal{F}$  und  $y \in \mathcal{I}$ , mit der Relation  $a \succ b$  interpretiert als *a besser als b*? Ist also eine gültige Lösung im Sinne der Bewertungsfunktion immer besser als eine beliebige nicht gültige? Die Problematik wird klar, wenn eine dicht neben dem Optimum gelegene aber ungültige mit einer weit entfernt gelegenen gültigen verglichen werden soll.
- Sind ungültige Lösungen ungeeignet als Lösungskandidaten oder führen sie möglicherweise schneller zum Optimum, da sie als Abkürzung oder zum Überbrücken ungültiger Regionen des Suchraums dienen?

- Dies führt direkt zu der Fragen, ob und wenn ja wie ungültige Lösungen repariert werden sollten. Oder sollten ungültige Lösungen lediglich abgewertet werden?

Eine Reihe von Verfahren zur Handhabung von Constraints während der Optimierung sind mittlerweile bekannt und auch eingehender untersucht worden. Überraschend wenig Bemühungen wurden bisher allerdings in die (standardisierte) Integration von Constraintverfahren in evolutionäre Optimierungsstrategien gesteckt (Kramer 2010; Coello Coello 2002), obwohl gerade diese ein hohes Potenzial auch für die Lösung komplexer Probleme aufweisen.

Nachfolgend sollen einige Vertreter etablierter Lösungsstrategien eingehender diskutiert werden. Diese haben teilweise vorbereitenden Charakter für den in dieser Arbeit entwickelten Ansatz bzw. werden später zum Vergleich während der Evaluation dienen. Ein guter Überblick über diverse Strategien zur Constraintbehandlung im Allgemeinen und für evolutionäre Ansätze findet sich u. a. in (Coello Coello 2002; Michalewicz & Schoenauer 1996; Richardson et al. 1989; Kramer 2010; Michalewicz 1996; Michalewicz 1995; Gen & Cheng 1997).

In rein mathematischen Fragestellungen sind Constraints in einem Optimierungsproblem häufig durch einen Satz von Gleichungen oder Ungleichungen gegeben (siehe beispielsweise Michalewicz & Schoenauer 1996; Kramer 2010; Helwig & Wanka 2007). Die Constraint-Gleichungen müssen dabei keineswegs linear sein (Avriel 2003).

Aber nicht alle Constraints lassen sich dergestalt über eine Gleichung beschreiben. Aus der Logik beispielsweise sind Ansätze zur Beschreibung der Beziehungen verschiedener Parameter einer Lösung untereinander bekannt. Bei diesem Ansatz der Constraint-Programmierung definieren Relationen zwischen den Variablen eine Eigenschaft, die eine Lösung erfüllen muss, um zulässig zu sein (siehe z. B. Rossi et al. 2006).

Zu unterscheiden sind Constraint Satisfaction Probleme (CSP), bei denen es vorrangig darum geht, alle Lösungen (zumindest eine) zu finden, welche alle Constraints erfüllt und Constrained Optimization Probleme (COP) mit dem vorrangigen Ziel die optimale (nach separat gegebener Zielfunktion) aller zulässigen Lösungen zu finden. Für CSP erfüllt eine Logik-basierte Darstellung ihren Zweck, welche gleichzeitig formal spezifiziert als Input für eine Reihe spezieller Constraint-Solver dient. Dieses bis auf Spezialfälle NP-harte Problem wird hier nicht eingehender betrachtet.

Für durch Constraints beschränkte Probleme sollen an dieser Stelle zunächst einige Definitionen begrifflich gefasst werden. Die Definitionen richten sich nach (Eiben 2001; Eiben & Ruttkay 1997).

**Definition 4.1.** Seien  $\mathcal{D}_1, \dots, \mathcal{D}_n \subseteq \mathbb{R}$  die Definitionsmengen einer Funktion  $f(x_1, \dots, x_n)$  deren Optimum gesucht ist. Ein Operator (Mutation, Rekombination, ...) heißt frei, wenn die Ausführung dieses Operators auf einem Lösungskandidaten (oder einer Menge von Lösungskandidaten) wieder zu einer gültigen Lösung führt. Ein kartesisches Produkt von Mengen  $S = \mathcal{D}_1 \times \dots \times \mathcal{D}_n$  ist ein freier Suchraum in Bezug auf übliche Operatoren eines GA (Eiben 2001) falls alle  $\mathcal{D}_i$  diskret oder kontinuierlich und konvex sind.

#### 4. Constraint-basierte Optimierung

---

Alle üblichen Operatoren eines EA sind frei in einem freien Suchraum. Dies gilt aber nicht für beliebige Operatoren!

**Definition 4.2.** Ein freies Optimierungsproblem (FOP) ist definiert als Paar  $(S, f)$ , wobei  $S$  ein freier Suchraum und  $f : S \rightarrow \mathbb{R}$  eine reellwertige Zielfunktion sind. Lösung eines freien Optimierungsproblems ist ein  $x \in S$  mit optimalem Wert  $f(x)$ .

**Definition 4.3.** Ein Constraint Satisfaction Problem ist ein Paar  $(S, \phi)$  mit einem freien Suchraum  $S$  und einer Boole'schen Funktion  $\phi$ , der Zulässigkeitsbedingung. Eine Lösung  $x$  ist zulässig, falls  $\phi(x) = \text{true}$ . Die Menge  $\mathcal{F} = \{x | \phi(x) = \text{true}\}$  bezeichne den zulässigen Suchraum (engl. feasible search space). Dieser sei auch als Lösungsraum bezeichnet.

**Definition 4.4.** Ein Constraint Optimization Problem ist als Tupel  $(S, f, \phi)$  eine Kombination aus beidem. Lösen eines CSP bedeutet, mindestens ein zulässiges Element der Lösungsmenge zu finden. Im Falle eines COP ist das Ergebnis eine gleichzeitig zulässige und optimale Lösung (Eiben 2001). Im Rahmen dieser Arbeit werden lediglich COP-Fragestellungen betrachtet.

In weiteren Verlauf dieser Arbeit soll der Begriff Constraint nicht eingeschränkt sein auf eine bestimmte Constraintformulierung sondern sich vielmehr allgemein auf jedwede Einschränkung beziehen die von einer Lösung erfüllt sein muss, unabhängig von der konkreten Modellierung.

Verschiedene Ansätze mit ihren individuellen Vor- und Nachteilen – auch zum Einsatz in EA – sind in der Vergangenheit entwickelt und auf verschiedene Anwendungsfälle angepasst worden. Constraint-Handling in evolutionären Algorithmen ist nicht unkompliziert integrierbar. Speziell die Operationen für die Rekombination zur Erzeugung der Nachfolgepopulation und die Mutation von Lösungsindividuen sind üblicherweise blind gegenüber Constraints. Das heißt, es gibt keine Garantie dass aus einer Population von zulässigen Lösungen auch zulässige Nachfolgelösungen hervorgehen. Somit sind evolutionäre Algorithmen eigentlich ungeeignet für constraintbehaftete Problemstellungen (Eiben 2001). Ein guter Überblick über bereits in der Forschung behandelte Ansätze findet sich beispielsweise in (Coello Coello 2002; Michalewicz & Schoenauer 1996; Kramer 2010). Die Verfahren von denen einige in den nachfolgenden Abschnitten kurz vorgestellt werden lassen grob wie folgt gliedern (nach Kramer 2010):

- Straffunktionen welche die Fitness einer ungültigen Lösung herabsetzen.
- Reparaturverfahren, welche eine unzulässige Lösung durch eine gültige ersetzen können oder eine solche zumindest für die Evaluation der Lösung nutzen.
- Dekoderansätze welche den Genotyp der Lösung auf einen zulässigen Phänotyp abbilden und somit eine Transformation des Problems in einen nicht eingeschränkten Suchraum ermöglichen.
- Zulässigkeit erhaltende Operatoren, welche dafür sorgen, dass alle Lösungskandidaten einer Nachfolgepopulation zulässig sind.

- Mehrzielansätze, welche Constraints als separate Optimierungsziele ansehen.

Die Grenzen zwischen diesen Klassen sind in der Regel fließend, sodass Verfahren Eigenschaften mehrerer Klassen aufweisen können. Im Rahmen dieser Arbeit sind solche Verfahren von Interesse, welche sich auch für Blackbox-Probleme eignen, also für solche bei denen keine Kenntnis der Zielfunktion vorausgesetzt wird und wo diese lediglich an den bereits evaluierten Stellen bekannt ist.

Als erstes soll nachfolgend jedoch die Methode der Lagrange-Multiplikatoren vorgestellt werden, welche als eigenständige, analytische Methode eigentlich nicht dieser Aufzählung zuzuordnen ist. Sie ist jedoch für das Verständnis einiger in den nachfolgenden Kapiteln behandelte Konzepte von Bedeutung und wird hier daher als Grundlage ebenfalls vorgestellt.

### 4.5.1. Lagrange'sche Multiplikatoren und die Karush-Kuhn-Tucker-Bedingungen

Die Lagrange'sche Methode (nach Joseph-Louis Lagrange) überführt gegebene Optimierungsprobleme mehrerer Variablen mit Constraints in eine constraintfreie Form, die anschließend im Idealfall analytisch oder beispielsweise mit einfachen Gradientenverfahren gelöst werden kann.

Ausgehend von einer Funktion  $f(\mathbf{x}) : \mathbb{R}^d \rightarrow \mathbb{R}$  deren Extremum unter einem Satz von  $n$  Nebenbedingungen  $g_1(\mathbf{x}), \dots, g_n(\mathbf{x})$  gesucht ist

$$f(x_1, \dots, x_i) \rightarrow \text{extrem}, \quad g_j(x_1, \dots, x_i) = 0, \quad 1 \leq j \leq n,$$

bildet man eine neue Funktion

$$F(\mathbf{x}, \boldsymbol{\lambda}) \equiv f(\mathbf{x}) + \sum_j \lambda_j \cdot g_j(\mathbf{x}).$$

(Lang & Pucker 1998). Hieraus ergeben sich die Extremalbedingungen

$$\begin{aligned} \frac{\partial F}{\partial x_i} &= 0, \quad 1 \leq i \leq d \\ \frac{\partial F}{\partial \lambda_j} &= 0, \quad 1 \leq j \leq n. \end{aligned} \tag{4.1}$$

Unter der Voraussetzung, dass die Zahl der Nebenbedingungen kleiner ist als die Zahl der Parameter von  $f$ , erfüllen genau die Extrempunkte das Gleichungssystem (4.1). Zusätzliche Bedingung für die Anwendbarkeit des Verfahrens sind die Differenzierbarkeit von  $f$  und die stetige Differenzierbarkeit aller  $g_j$  (Königsberger 1991).

Die Methode der Lagrange'schen Multiplikatoren erlaubt nur Gleichungen als Nebenbedingung, nicht aber Ungleichungen oder gar nicht-lineare Nebenbedingungen. Eine Verallgemeinerung des Lagrange-Ansatzes, welche diese Einschränkung aufhebt, stellen die Karush-Kuhn-Tucker (KKT) Bedingungen dar, welche die notwendigen Bedingungen für die Optimalität einer Lösung in der nicht-linearen Optimierung angeben (Kuhn & Tucker 1950).

#### 4. Constraint-basierte Optimierung

---

Liegen weitere Nebenbedingungen in Form von Ungleichungen  $h_k \leq 0$  vor, so ergibt sich die Lagrange-Form (Boyd & Vandenberghe 2004)

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \sum_j \lambda_j \cdot g_j(\mathbf{x}) + \sum_k \mu_k \cdot h_k(\mathbf{x}). \quad (4.2)$$

Seien  $\mathbf{x}^*$  und  $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$  Extremstellen von  $L$  mit gleichem Wert für primales und duales Optimierungsproblem (zero duality gap). Die KKT-Bedingungen für die Optimalität einer Lösung  $\mathbf{x}^*$  welche für die Existenz von  $\boldsymbol{\lambda}^*$  und  $\boldsymbol{\mu}^*$  notwendigerweise erfüllt sein müssen sind (Boyd & Vandenberghe 2004):

##### **Definition 4.5.**

**Stationarität** (der Gradient verschwindet im Extrempunkt):

$$\nabla L(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = \nabla f(\mathbf{x}^*) + \sum_j \lambda_j^* \cdot \nabla g_j(\mathbf{x}^*) + \sum_k \mu_k^* \cdot \nabla h_k(\mathbf{x}^*) = 0, \quad (4.3)$$

**Primärbedingungen** (Gültigkeit der Lösung):

$$\begin{aligned} h_j(\mathbf{x}^*) &\leq 0 \quad \forall j \\ g_k(\mathbf{x}^*) &= 0 \quad \forall k, \end{aligned} \quad (4.4)$$

**Dualität** (duale Zulässigkeitsbedingung):

$$\mu_k \geq 0 \quad \forall k, \quad (4.5)$$

**Komplementärer Schlupf** (Nachweis der Optimalität nach dem Satz vom komplementären Schlupf (z. B. Grötschel 2013)):

$$\mu_k \cdot h_k(\mathbf{x}^*) = 0 \quad \forall k. \quad (4.6)$$

Für konkave Funktionen sind diese notwendigen Bedingungen bereits hinreichend, falls alle Ungleichungsnebenbedingungen konvex und alle Gleichungsnebenbedingungen affin sind. In anderen Fällen existieren zusätzliche Qualifikationsbedingungen an eine Stelle  $\mathbf{x}^*$ , um zum Nachweis eines globalen Optimums zu dienen (z. B. in Merz & Wüthrich 2013; Boyd & Vandenberghe 2004; Schild 2014).

In jedem Fall wird zur Anwendung die Differenzierbarkeit von  $f$ ,  $g$  und  $h$  benötigt. Diese Vorbedingung ist in den in dieser Arbeit im Fokus stehenden Optimierungsproblemen für die Wirkleistungsplanung im Smart Grid nicht garantiert (wodurch die Anwendbarkeit hier nicht gegeben ist), jedoch sind die hier reflektierten Grundlagen hilfreich für das Verständnis der Grundlagen der in den nachfolgenden Kapitel beschriebenen Supportvektor-Verfahren und wurden deshalb hier dargelegt.

### 4.5.2. Mehrzielansätze

Eine Möglichkeit zur Integration von Constraints in evolutionäre Algorithmen besteht darin, ein Einzelproblem mit  $m$  Nebenbedingungen umzuwandeln in eine Mehrzielformulierung mit  $m + 1$  Optimierungszielen. Constraints werden hierbei nicht mehr als Hard Constraints sondern als zu optimierendes Ziel behandelt. Prinzipiell kann somit jedes Standard-Mehrzieloptimierungsverfahren (Fonseca & Fleming 1995) benutzt werden (beispielsweise Beume et al. 2008; Surry & Radcliffe 1997, o. a. ). Coello Coello (2002) listet eine Reihe von durchgeführten Studien in denen dieses Verfahren mit anderen Verfahren verglichen wird mit dem Ergebnis, dass sich dieses Verfahren am ehesten in hochgradig eingeschränkten Suchräumen eignet, um den zulässigen Bereich überhaupt zu erreichen. In Bezug auf das eigentliche Optimierungsziel sind die Ergebnisse häufig nicht so überzeugend (Coello Coello 2002). Allerdings würde z. B. der zusätzliche Einbezug von a priori Wissen über die Suchraumstruktur sowie ein Ranking der Lösungskandidaten – wie beispielsweise in (Ray et al. 2000; Runarsson & Yao 2000) – eine vielversprechende Forschungsrichtung aufweisen. Am ehesten eignet sich dieser Ansatz für Soft Constraints da eine Erfüllung aller Constraints durch die gefundene Lösung (insbesondere bei Vorliegen einer Vielzahl von Hard Constraints) nicht garantiert ist.

### 4.5.3. Penaltyansatz

Ein seit langem etablierter und häufig genutzter Ansatz ist die Bestrafung von unzulässigen Lösungen durch Beaufschlagung von Strafwerten mithilfe einer Penaltyfunktion mit welcher die Zielfunktion erweitert wird (Carroll 1961; Fiacco et al. 1965, u. a. ).

Die häufigste Formulierung ergänzt eine Bestrafung ungültiger Lösungen (Michalewicz & Schoenauer 1996):

$$f^*(\mathbf{x}) = \begin{cases} f(\mathbf{x}), & \text{falls } \mathbf{x} \in \mathcal{F} \\ f(\mathbf{x}) + p(\mathbf{x}), & \text{bei ungültigem } \mathbf{x}. \end{cases} \quad (4.7)$$

Gebräuchlich sind aber im Wesentlichen drei Ansätze mit unterschiedlicher Reichweite der Straffunktion (Schwefel 1995):

- Als Extremfall kommen so genannte Barrierefunktionen zum Einsatz welche allen ungültigen Lösungen einen Wert von  $f(\mathbf{x}) = +\infty$  (im Minimierungsfall,  $-\infty$  sonst) zuweisen. Auf diese Weise soll verhindert werden, dass ungültige Lösungen überhaupt betrachtet werden. Für evolutionäre Strategien ist als Startlösung unbedingt eine gültige Lösung erforderlich.
- Bei der partiellen Penaltyfunktion werden Strafwerte nur für Lösungsvektoren nahe der Grenze und auch nur innerhalb des zulässigen Lösungsraums vergeben. Ziel ist es die Suche vom Rand fern und damit innerhalb des zulässigen Bereichs zu halten. Derartige Strategien werden auch als interior Penalties bezeichnet. Die gegensätzliche (exterior) Strategie versucht unzulässige Lösungen in den gültigen Raum hineinzuziehen (Coello Coello 2002).
- Globale Penaltyansätze schließlich variieren die Zielfunktion durch zusätzliche Straf-

werte über den gesamten Definitionsbereich von  $f$ . Diese Form wird meist für Constraints in Form von Ungleichungen verwendet.

Zudem gibt es Unterschiede in der möglichen Bestimmung der Strafwerte. Die einfachste Methode ist die Verwendung statischer Penalties, bei der im Falle der Ungültigkeit einer Lösung ein konstanter Strafwert auf den Funktionswert beaufschlagt wird (im Minimierungsfall). Im Falle mehrerer Constraints kann hier auch eine Metrik zum Einsatz kommen, die den Strafwert abhängig von der Anzahl der verletzten Constraints bestimmt (Smith & Coit 1997). Alternativ sind auch dynamische oder adaptive Penalty-Ansätze üblich, welche beispielsweise eine angepasste Bestimmung des Strafwertes in Abhängigkeit vom Abstand vom zulässigen Lösungsbereich oder von der bereits zurückgelegten Länge des Suchpfades abhängen, bzw. welche sich adaptiv an den Verlauf der Suche (und damit an die vorgefundenen Eigenschaften der Zielfunktion) anpassen. In beiden Fällen jedoch sind vom Nutzer vorab die entsprechenden Parameter spezifisch für das jeweilige Problem einzustellen bzw. die entsprechenden Ansätze für das Problem speziell zu implementieren. Für allgemein verwendbare Vorgehensweisen besteht noch weiterer Forschungsbedarf (Smith & Coit 1997; Dadios & Ashraf 2006).

Für die spätere Evaluation (vgl. Abschnitt 12.4.2 auf Seite 191) des in dieser Arbeit zu entwickelnden Ansatzes wird der Penaltyansatz im Vergleich betrachtet werden.

#### 4.5.4. Dekoderansatz

Ein weiterer möglicher Ansatz ist, eine spezielle Repräsentation der Lösungskandidaten zu wählen. Hierdurch wird das Problem so formuliert, dass alle Lösungen automatisch alle Constraints erfüllen (Simon 2013). Ergänzend hierzu können auch spezielle Operatoren definiert werden, die dafür sorgen, dass bei Rekombination und Mutation (z. B. bei Verwendung eines EA) einer Lösung ausschließlich gültige Nachfolgelösungen entstehen. Im Folgenden soll das Konzept eines Dekoders für die Constraintbehandlung eingeführt und näher betrachtet werden. Hierzu sei zunächst folgendes Beispiel für die spezielle Repräsentation eines Problems gegeben: Sei in Abbildung 4.2 die graue Fläche  $\mathcal{F} \subset [0, 1]^2$  der Lösungsraum der zulässigen Lösungen. Der Lösungsraum lässt sich beschreiben über folgende Parametergleichung:  $P(a, b) = \vec{o} + a \cdot \vec{r} + b \cdot \vec{s}$  mit  $a, b \in [0, 1]$ . Die Parameter  $a$  und  $b$  sind lediglich beschränkt durch ihren Wertebereich, durch die Instruktion  $P(a, b)$  wird dem Punkt  $(a, b) \in [0, 1]^2$  eindeutig eine gültige Lösung aus  $\mathcal{F} \subset [0, 1]^2$  zugeordnet.

Spezielle Repräsentationen (wie auch die nicht näher betrachteten speziellen Operatoren) sind weitgehend problemspezifisch. Es ist bisher nicht möglich, spezielle Repräsentationen zu schreiben, die für eine breite Klasse von Problemen anwendbar sind (Simon 2013). Dennoch hat dieses Konzept großes Potenzial, da trotz des Mehraufwandes durch die Erstellung einer solchen Repräsentation im Ergebnis in der Regel eine deutlich bessere Performanz als mit allgemein einsetzbaren EA erreicht wird (Simon 2013).

Eine Möglichkeit für die Umsetzung einer solchen speziellen Repräsentation stellen die so genannten Dekoder dar. Hierzu wird ein Satz an Instruktionen erstellt der es erlaubt, gültige Lösungskandidaten zu bestimmen sodass alle Individuen der Lösungspopulation (im Falle eines EA) gültige Lösungen darstellen. Anstatt direkt eine Lösung des Problems zu kodieren, wird eine Konstruktionsvorschrift für eine gültige Lösung kodiert (vgl. Si-

mon 2013). Eine solche Konstruktionsvorschrift, die das Erzeugen einer gültigen Lösung erlaubt wird Dekoder genannt (Koziel & Michalewicz 1998; Koziel & Michalewicz 1999; Palmer & Kershenbaum 1994). Ein Dekoder stellt eine Beziehung zwischen der speziellen Repräsentation und einer gültigen Lösung her (Dasgupta & Michalewicz 1997).

Idealerweise sollte ein solcher Dekoder die folgenden Eigenschaften aufweisen (Palmer & Kershenbaum 1994; Simon 2013; Coello Coello 2002):

- Jede gültige Lösung sollte von einem Dekoder erzeugt werden können.
- Eine gültige Lösung sollte nicht mehrfach durch verschiedene Repräsentationen erzeugt werden, um die Suche nicht voreingenommen auf bestimmte Regionen zu konzentrieren.
- Ein Dekoder sollte keine ungültige Lösung erzeugen.
- Die Transformation sollte in Vergleich zur Zielfunktion schnell zu berechnen sein.
- Kleine Änderungen im Dekoder sollten mit kleinen Änderungen der Lösungskandidaten korrespondieren.

Diese Regeln stellen Richtlinien, keine strengen Anforderungen für ein generelles Funktionieren dar. Bei Bedarf können diese Richtlinien auch gelockert werden, wenn sie je nach Anwendungsfall schwer zu erfüllen sind (Koziel & Michalewicz 1999). Beispielsweise kann die Forderung nach kleinen Änderungen im Falle disjunkter Teilbereiche mit gültigen Lösungen nicht aufrecht gehalten werden. In der speziellen Repräsentationen werden zwei gültige Lösungen, die in verschiedenen Teilregionen (durch einen Bereich ungültiger Lösungen getrennt) am jeweiligen Rand liegen, als Dekoderlösungen direkt benachbart sein, obwohl die Lösungen tatsächlich weiter auseinander liegen.

Eine praktische Umsetzung eines Dekoders finden sich in (Koziel & Michalewicz 1998; Koziel & Michalewicz 1999). Hierbei wird ausgehend von einem Referenzpunkt  $r_0$  im Inneren des zulässigen Bereichs ein Suchstrahl auf den Rand einer Boundingbox (Rand der Domäne der Parameter) gerichtet und im Zuge der Optimierung variiert. Sei  $b$  der Schnittpunkt mit der Boundingbox (vgl. Abbildung 4.3 auf der nächsten Seite). Für jeden Suchstrahl kann der entfernteste (vom Referenzpunkt) gerade noch zulässige Punkt bestimmt werden. Sei  $\nu_{max}(b)$  die maximale Entfernung. Ausgehend vom Referenzpunkt als Ortsvektor kann durch Addition eines Vektors (mit einem Skalierungsfaktor  $\nu \in [0, 1]$ ) der eine Länge aus

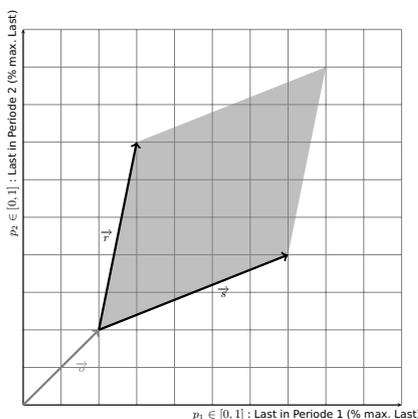
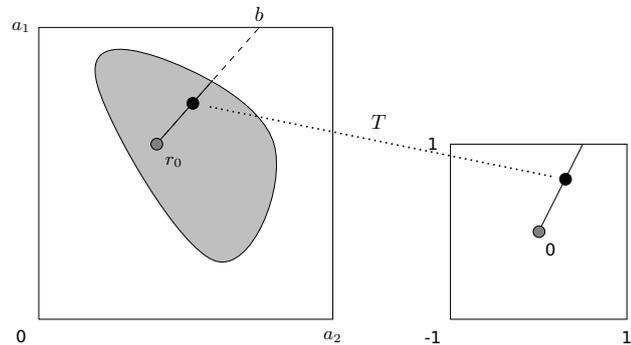


Abbildung 4.2: Vereinfachtes, lineares Dekoderbeispiel anhand der Parametergleichung der vektoriellen Darstellung eines Parallelogramms. Die Darstellung des Parallelogramms kann über die dargestellten Vektoren erfolgen. Alle Punkte in der Figur können erreicht werden durch Variation der Länge der beiden Spannvektoren. Beispielsweise könnte diese Fläche die zulässigen Wirkleistungswerte für zwei aufeinander folgende Zeitintervalle darstellen.

Abbildung 4.3: Dekoderansatz nach (Koziel & Michalewicz 1999); 2-dimensionales Beispiel aus (Koziel & Michalewicz 1998), verändert. Die Abbildung  $T$  bildet die zulässige Region (grau) in den Würfel  $[-1, 1]^2$  ab.



$[0, \nu_{max}(b)]$  erlaubt) eine gültige Lösung erzeugt werden. Für einen EA wird nunmehr ein Lösungsindividuum definiert über den variablen Randpunkt  $b$  und einen Skalierungsfaktor  $\nu \in [0, 1]$  sowie obiger Anweisung, um hieraus eine gültige Lösung zu erzeugen. Ist der zulässige Lösungsraum konkav oder besteht er aus mehreren disjunkten Teilbereichen, so muss dieses Verfahren noch angepasst werden indem auch der Suchstrahl in mehrere individuelle Teilbereiche unterteilt wird; je nach Randpunkt  $b$ .

Ein früherer Ansatz basierend auf Riemann'schen Abbildungen findet sich in (Kim 1998). Squire (1975) verwendet eine Schwarz-Christoffel Transformation für die winkeltreue Abbildung der komplexen Halbebene auf den durch ein Polygon beschränkten zulässigen Bereich. Beide Verfahren bedürfen der individuellen Anpassung an die konkrete Problemstellung und stellen somit keine universell wiederverwendbaren Ansätze dar.

Der hier eingeführte Dekoder-Ansatz wird in dieser Arbeit weiter ausgebaut werden. Ziel wird es sein, einen Dekoder automatisiert aus einem abstrakten Modell der zulässigen Region und der einschränkenden Constraints abzuleiten, ohne eine probleminstanzindividuelle formale Formulierung aufstellen zu müssen.

#### 4.5.5. Spacemapping und Surrogatmodelle

Eine in den Ingenieursdisziplinen gängige Technik zur Handhabung berechnungstechnisch schwieriger Optimierungsprobleme ist die Näherung des eigentlichen Modells durch eine grobe aber einfach (und vor allem schnell) zu berechnende Ersatzfunktion, dem so genannten Surrogatmodell. Basierend auf dem Konzept von Bandler et al. (1994) wurden als Techniken zunächst die Methode von Broyden (1965), eine Quasi-Newton-Methode zum finden von Nullstellen bei der die Jakobi-Matrix nur einmal berechnet und danach in jeder Iteration lediglich aktualisiert wird, das Trust-Region-Verfahren bei dem die Zielfunktion mit einem quadratischen Modell innerhalb eines definierten Vertrauensbereich genähert wird (Bakr et al. 1998), aber auch künstliche Neuronale Netze eingesetzt. Neuere Ansätze für Surrogatmodelle schließen auch Datascape, Kriging, Regression und Response Surfaces mit ein (Gano et al. 2006; Myers & Montgomery 1995).

Der Grundgedanke ist immer gleich: Zu jedem Zeitpunkt des Optimierungsprozesses existieren zwei Modelle. Ein hochgenaues, langsam zu berechnendes und ein vergrößertes auf dem die Optimierung durchgeführt wird. Das genaue Modell wird lediglich an neuralgischen Stellen zur Evaluation von Zwischenlösungen benötigt; wobei diese

Zwischenergebnisse auch zur Verbesserung des vergrößerten Modells genutzt werden. Hier ergeben sich auch Verwandtschaften zum einen zur modellbasierten prädiktiven Regelung (Grüne & Pannek 2011; Rau 2003) und zum anderen zum Active Learning – einer sukzessiv verbesserten Darstellung der Grenzen des Zulässigkeitsbereichs wie sie durch Constraints vorgegeben werden (Settles 2012).

Ursprünglich sind beim Konzept der Surrogatmodellierung die Constraints vollständig bekannt und bereits modelliert (Neddermeijer et al. 2000) im Gegensatz zur schwer zu berechnenden (oder auf aufwendigen Simulationen zur Evaluation beruhenden) Zielfunktion. Es gibt jedoch auch Ansätze, Surrogatmodelle zur schnelleren Überprüfung von Constraints einzusetzen (vgl. z. B. Kramer 2010).

## 4.6. Zusammenfassung

Ziel dieses Kapitels war es, einen kurzen Überblick über bekannte und etablierte Methoden zur Handhabung von Constraints während der Optimierung zu geben. Eine umfassende Einführung in die Optimierung im Allgemeinen ist im Rahmen einer solchen Überblicksdarstellung nicht möglich und war auch nicht das Ziel. Im weiteren Verlauf dieser Arbeit wird ein Ansatz für die Integration von Constraints basierend auf dem hier vorgestellten Konzept des Dekoderansatz bzw. dem des eng hiermit verwandten Ansatzes des Space-mappings entwickelt werden. Im Gegensatz zu den hier vorgestellten Methoden, welche auf einer explizit gegebenen (meist linearen) Darstellung der Constraints fußen, wird die in dieser Arbeit entwickelte Methode auch nicht explizit gegebene Constraints ad hoc integrieren können, um auch dynamische Optimierungsprobleme lösen zu können. Als weiterer, entscheidender Vorteil wird der Dekoder automatisiert erstellt werden können, ohne manuelle Anpassungen vornehmen zu müssen. Somit wird der neu entwickelte Dekoder auch in selbstorganisiert arbeitenden Algorithmen zum Einsatz kommen können.

Zudem wurde in diesem Abschnitt die Methode der Lagrange'schen Multiplikatoren sowie die KKT-Bedingungen rekapituliert, welche indirekt für das Verständnis der zugrundeliegenden Supportvektor-Ansätze hilfreich sind, die im nachfolgenden Kapitel vorgestellt werden.



# 5. Supportvektor Maschinen und Kern-basierte Verfahren

*Vector! That's me, because I commit crimes with both  
direction and magnitude.  
Oh yeah!*

---

*(Dispicable Me)*

## 5.1. Einleitung

Ein wesentlicher Bestandteil des in dieser Arbeit entwickelten Lösungsansatzes wird die Verwendung von Supportvektor Maschinen (SVM) sein. Daher wird deren Konzept in diesem Kapitel schon einmal einleitend betrachtet, um später deren Einbettung in den Lösungsansatz besser nachvollziehen zu können.

In der klassischen Klassifikationsfragestellung ist üblicherweise eine Menge von Paaren  $\{(x_1, y_1), \dots, (x_n, y_n)\}$  gegeben, bestehend jeweils aus einem Eingabevektor  $x$  und einem zugeordneten Label  $y$ . Aufgabe der Klassifikation ist es, ein Modell zu lernen, das für neue Eingabevektoren eine Vorhersage über das Label macht, bei möglichst minimalem Fehler. Verschiedene Verfahren wurden entwickelt, von denen viele eine lineare Separierbarkeit der Daten voraussetzen.

Bereits in den 60er Jahren stellten Minsky & Papert (1969) fest, dass die praktische Anwendung linearer Klassifikation recht begrenzt ist, da sich im Allgemeinen komplexe Real-World-Anwendungen nicht durch lineare Funktionen darstellen lassen und somit komplexere Repräsentationen erfordern.

Diese Frage wurde von Vapnik (1995); Vapnik & Chervonenkis (1974) gelöst durch Projektion der Eingabevektoren aus dem vergleichsweise eher niedrigdimensionalen Datenraum in einen deutlich höherdimensionalen Feature-Raum, wo eine lineare Trennung einfacher zu erreichen ist. Supportvektor Maschinen sind so genannte Maximum Margin Klassifikatoren, da sie eine trennende Hyperebene erzeugen, die einen möglichst großen Abstand zu den nächstgelegenen Datenpunkten hat. Hierdurch soll die Generalisierbarkeit des Modells verbessert werden. Die Supportvektor Maschinen in ihrer heutigen Form schließlich wurden von Boser et al. (1992) eingeführt. Sie werden im Anschluss kurz eingeführt.

Interessanter für diese Arbeit jedoch ist die Theorie der 1-Klassen SVM – genauer die der Support Vector Data Description – welche im zweiten Teil dieses Kapitels eingehender behandelt wird. Sie bildet die Grundlage für das später in dieser Arbeit konzipierte Modell der Fahrplansuchräume.

## 5.2. Supportvektor Maschinen

In diesem Abschnitt soll zunächst das Grundprinzip der Supportvektor Maschinen vorgestellt werden. Diese werden zwar in dieser Arbeit nicht direkt verwendet, bilden aber die Basis für das Verständnis der Supportvektor Domain Description, welche im nachfolgenden Abschnitt als Basistechnologie für diese Arbeit vorgestellt wird. Insgesamt stellt die Theorie der Supportvektor-Methoden ein sehr gut erforschtes und somit auch breites Feld dar, das hier nicht erschöpfend behandelt werden kann. An dieser Stelle sei daher für tiefer gehende Fragestellungen auch auf die Einführung in die Thematik beispielsweise in (Steinwart & Christmann 2008; Schölkopf et al. 1999a; Schölkopf 1997; Vapnik 1982; Gunn 1998; Cristianini & Shawe-Taylor 2000) verwiesen.

Vor der kurzen Einführung in die Funktionsweise der SVM sollen zunächst einige wichtige Begriffe definiert werden.

**Definition 5.1.** Ein Hilbertraum  $\mathcal{H}$  ist ein Vektorraum mit einem Skalarprodukt.

**Definition 5.2.** Ein Kern (engl.: kernel) ist nach (Schaback & Wendland 2006) zunächst einmal eine Funktion

$$k : X \times X \rightarrow \mathbb{R}, \quad (5.1)$$

wobei  $X$  eine beliebige, nichtleere Menge darstellt.

Mit  $k$  sind allerdings spezielle Eigenschaften verknüpft. Sei

$$\Phi : X \rightarrow \mathcal{H} \quad (5.2)$$

eine Feature Map, welche jedes Element  $x \in X$  auf eine möglichst große Menge  $\Phi(x)$  von Eigenschaften (Features) abbildet, die  $x$  möglichst gut charakterisieren. In  $\mathcal{H}$  existiert dann deutlich mehr Struktur zur Beschreibung der Daten als in  $X$  (vgl. Schaback & Wendland 2006). Zu einer solchen Feature Map gibt es dann einen Kern  $k$  mit

$$k(x, z) := \langle \Phi(x), \Phi(z) \rangle_{\mathcal{H}}. \quad (5.3)$$

Häufig erfolgt die Definition auch anders herum. Zu einem gegebenen Kern  $k$  gibt es eine Feature Map, sodass (5.3) erfüllt ist (vgl. z. B. Shawe-Taylor & Cristianini 2004). Als weitere Eigenschaft folgt aus (5.3), dass ein Kern symmetrisch ist.

**Definition 5.3.** Ein Kern  $k$  heißt reproduzierender Kern eines Hilbertraums  $\mathcal{H}$ , wenn

$$\forall f \in \mathcal{H} \bullet f(x) = \langle k(x, \cdot), f(x) \rangle. \quad (5.4)$$

Zu jedem positiv semidefiniten Kern  $k$  gehört ein assoziierter Reproducing Kernel Hilbert Space RKHS. Dies ist der durch die Feature Map erzeugte Feature Raum. Eine ausführliche Diskussion findet sich beispielsweise in (Schölkopf et al. 1999a; Shawe-Taylor & Cristianini 2004).

### 5.2.1. Grundprinzip

Auf konzeptioneller Ebene können Supportvektor Maschinen gesehen werden als Regularisierungsfragestellung (vgl. Gieseke et al. 2014; Steinwart & Christmann 2008; Tikhonov & Arsenin 1977):

$$\inf_{f \in \mathcal{H}} \lambda \|f\|_{\mathcal{H}}^2 + \frac{1}{2} \sum_{i=1}^n \mathcal{L}(y_i, f(x_i)), \quad (5.5)$$

mit  $\lambda \in \mathbb{R}^+$  und einer Fehlerfunktion (engl. Loss)  $\mathcal{L} : Y \times \mathbb{R} \rightarrow [0, \infty)$  zur Beurteilung der Klassifikatorvorhersagen.  $\|\cdot\|_{\mathcal{H}} = \langle \cdot, \cdot \rangle$  ist eine Norm im so genannten Reproducing Kernel Hilbert Space (RKHS)  $\mathcal{H} \subseteq \mathbb{R}^X = \{f : X \rightarrow \mathbb{R}\}$  (Aronszajn 1950). Jeder RKHS  $\mathcal{H}$  wird eindeutig durch einen gegebenen Kern  $k : X \times X \rightarrow \mathbb{R}$  erzeugt.  $\mathcal{H}$  kann im Rahmen von Machine-Learning-Ansätzen auch als Hypothesenraum gesehen werden. Der Hilbertraum ist ein Funktionenraum und jede dieser Funktionen stellt eine Hypothese über die Struktur der betrachteten Daten dar. Das Abschätzen von Ursachen für beobachtete Konsequenzen kann leicht zu schlecht konditionierten Problemen führen (Vapnik 1982). Die Tikhonov Regularisierung ist ein Ansatz, derartige Probleme anzugehen. Sie schränkt den Suchraum nach geeigneten Hypothesen mit gewissen gewünschten Eigenschaften ein (Garcke 2004). Während der zweite Term ein Maß für den Vorhersagefehler darstellt, bestraft der Norm-Term – der durch die Regularisierung von Tikhonov hinzu kommt (Garcke 2004) – zu komplexe Funktionen (vgl. auch Gieseke et al. 2014). Der Vorteil dieser Betrachtungsweise liegt in dem theoretischen Framework, dass eine Analyse und den Vergleich der SVM in Bezug auf Überanpassung und Generalisierbarkeit erlaubt – auch im Vergleich mit anderen Algorithmen für dieselbe Mustererkennungsfragestellung.

Geläufiger, häufiger verwendet und im Rahmen dieser Arbeit anschaulicher ist jedoch die ursprüngliche geometrische Interpretation der SVM, bei der es das Ziel ist, eine trennende Hyperebene zu konstruieren – im nicht linear trennbaren Fall unter Zulassung eines gewissen (allerdings zu minimierenden) Fehlers. Im Folgenden wird daher diese Interpretation verwendet.

Sei  $X \subset \mathbb{R}^d$  der Datenraum,  $Y = \{-1, +1\}$  und  $D = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)\}$  eine Trainingsmenge sodass ein Vektor  $\mathbf{w} \in \mathbb{R}^d$  mit  $\|\mathbf{w}\|_2 = 1$  und eine reelle Zahl  $b \in \mathbb{R}$  existieren, sodass gilt

$$\begin{aligned} \langle \mathbf{w}, \mathbf{x}_i \rangle + b &> 0 \quad \forall i \text{ mit } y_i = +1 \\ \langle \mathbf{w}, \mathbf{x}_i \rangle + b &< 0 \quad \forall i \text{ mit } y_i = -1. \end{aligned} \quad (5.6)$$

Dann werden die beiden Klassen perfekt durch die durch  $\langle \mathbf{w}, \mathbf{x}_i \rangle + b$  definierte Hyperebene getrennt. Mit dem Generalized Portrait Algorithm von Vapnik & Lerner (1963) wird die Hyperebene so konstruiert, dass sie einen maximalen Korridor (engl.: Margin) ohne enthaltene Punkte zwischen den beiden Klassen erzeugt. Dieses Vorgehen hat zwei Nachteile (Steinwart & Christmann 2008): für viele praktische Fälle ist eine lineare Entscheidungsgrenze nicht sinnvoll, weil sich die Klassen nicht linear trennen lassen und im Falle verrauschter Daten muss eine Fehlklassifikation gegebenenfalls zugelassen werden können, um eine Überanpassung zu verhindern.

Die Fragestellung nach der nicht-linearen Trennbarkeit wurde von (Boser et al. 1992) angegangen durch Verlagerung des Generalized Portrait Algorithm in einen Featureraum,

einen allgemeinen Hilbertraum  $\mathcal{H}_0$ . Die Trainingsdaten werden in den gegebenenfalls unendlichdimensionalen  $\mathcal{H}_0$  abgebildet durch eine Abbildung  $\Phi : X \rightarrow \mathcal{H}_0$  bevor die Hyperebenenbestimmung angewandt wird auf  $\{(\Phi(\mathbf{x}_1), y_1), \dots, (\Phi(\mathbf{x}_1), y_1)\}$ . Man spricht bei dieser Version von einer Hard Margin SVM und kann zeigen, dass es für alle widerspruchsfreien Trainingsdaten eine Hyperebene in  $\mathcal{H}_0$  gibt, welche beide Klassen trennt (Steinwart & Christmann 2008). Cortes & Vapnik (1995) führten später die Soft Margin SVM ein mit Schlupfvariablen, um gewisse Klassifikationsfehler zuzulassen. Die Kombination beider Erweiterungen führt zu folgendem Optimierungsproblem zur Bestimmung der SVM:

$$\begin{aligned} \frac{1}{2} \langle \mathbf{w}, \mathbf{w} \rangle + C \sum_{i=1}^n \xi_i \rightarrow \min \quad , \mathbf{w} \in \mathcal{H}_0, b \in \mathbb{R}, \boldsymbol{\xi} \in \mathbb{R}^n \\ \text{s.t. } y_i(\langle \mathbf{w}, \Phi(\mathbf{x}_i) \rangle + b) \geq 1 - \xi_i \quad , 1 \leq i \leq n \\ \xi_i \geq 0 \quad , 1 \leq i \leq n. \end{aligned} \quad (5.7)$$

Durch Einführung der Schlupfvariablen  $\boldsymbol{\xi}$  wird die Bedingung der Fehlerfreiheit an die Hyperebene  $y_i(\langle \mathbf{w}, \Phi(\mathbf{x}_i) \rangle + b) \geq 1$  abgeschwächt, sodass der spätere Klassifikator lediglich möglichst wenig Fehler macht. Hierdurch können aber eben auch eigentlich nicht trennbare Klassen separiert werden. Der Parameter  $C > 0$ , der die beiden Terme gegeneinander gewichtet kann frei gewählt werden. Statt Problem (5.7) direkt in  $\mathcal{H}_0$  zu lösen, wird üblicherweise das duale Problem gelöst, welches mit Hilfe der Lagrangemethode konstruiert wird:

$$\begin{aligned} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n y_i y_j \alpha_i \alpha_j \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle \rightarrow \max \quad , \alpha \in [0, C]^n \\ \text{s.t. } \sum_{i=1}^n y_i \alpha_i = 0, \end{aligned} \quad (5.8)$$

mit  $\alpha_i$  als Lagrange-Multiplikatoren. Nutzt man nun noch das Theorem von Mercer (Mercer 1909), (vgl. auch Schölkopf et al. 1999a) nach dem Skalarprodukte in  $\mathcal{H}$  im Datenraum  $X$  berechnet werden können ohne die Featuremap  $\Phi$  selbst kennen zu müssen, so können alle Berechnungen vollständig in  $\mathbb{R}^d$  durchgeführt werden. Bestimmte Kern-Funktionen  $k$  erfüllen nämlich die Eigenschaft

$$k(x, x') = \langle \Phi(x), \Phi(x') \rangle \quad ; x, x' \in X. \quad (5.9)$$

Somit können alle Ausdrücke in (5.8) in denen  $\Phi$  nur innerhalb eines Skalarproduktes vorkommt durch Ersetzen des Skalarproduktes durch  $k$  vereinfacht werden. Um den Kern-Trick (5.9) nutzen zu können, ist es notwendig, dass  $k$  positiv semi-definit ist (Aizerman et al. 1964). In diesem Fall ist die Gram-Matrix  $\mathbf{K}_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$  ebenfalls positiv semi-definit, d.h. es gibt Räume in denen  $\mathbf{K}$  ein Skalarprodukt darstellt (Cristianini & Shawe-Taylor 2000).

Nun definiert Mercers Theorem zwar welche Funktionen sich als Mercer-Kern für den Kern-Trick eignen. Die Frage ist aber noch wie gut die verschiedenen Funktionen jeweils sind. Einen guten Überblick über verschiedene, geeignete Kerne geben beispielsweise (Gunn 1998; Burges 1998).

Im Ergebnis steht nach Bestimmung der  $\alpha_i$  in Gleichung (5.8) eine Entscheidungsfunktion für die Klassifikation zur Verfügung, die wie folgt konstruiert werden kann, wenn  $\boldsymbol{\alpha}^* = (\alpha_1^*, \dots, \alpha_n^*)$  eine Lösung von Gleichung (5.8) darstellt:

$$f_D(\mathbf{x}) = \text{sgn}(\langle \mathbf{w}_D, \Phi(\mathbf{x}) \rangle + b^*). \quad (5.10)$$

Mit  $w_D = \sum_{i=1}^n \alpha_i^* y_i \Phi(\mathbf{x}_i)$  und Gleichung (5.9) ergibt sich

$$f_D(\mathbf{x}) = \text{sgn}\left(\sum_{i=1}^n \alpha_i^* y_i k(\mathbf{x}_i, \mathbf{x}) + b^*\right); \quad (5.11)$$

$b^* = y_j - \sum_{i=1}^n \alpha_i^* y_i k(\mathbf{x}_i, \mathbf{x}_j)$  lässt sich mit Hilfe eines der Trainingsvektoren bestimmen, deren zugehöriger  $\alpha$ -Wert die Bedingung  $0 \leq \alpha_j \leq C$  erfüllt. Die Ausgabe der Entscheidungsfunktion  $f_D(\mathbf{x})$  ist 1 oder -1 je nachdem auf welcher Seite der Hyperebene das Bild von  $\mathbf{x}$  in  $\mathcal{H}$  liegt. Es ist noch zu bemerken, dass die Berechnung von  $w_D$  lediglich von denjenigen Trainingsvektoren  $\mathbf{x}_j$  abhängt, die einen zugehörigen  $\alpha$ -Wert von  $\alpha_j \neq 0$  haben. Die Hyperebene wird somit alleinig von diesen  $\Phi(\mathbf{x}_j)$  gestützt (engl.: supported). Die zugehörigen Punkte  $(\mathbf{x}_j, y_j)$  heißen Supportvektoren (vgl. Steinwart & Christmann 2008).

### 5.2.2. Lösungsverfahren

Zur Bestimmung der Koeffizienten in Gleichung (5.8) ist ein Quadratisches Programm (QP) zu lösen. Prinzipiell kann hierfür ein beliebiger Standard-QP-Solver verwendet werden, wie er quasi in beinahe jeder Optimierungsbibliothek vorhanden ist. Allerdings ist dieser Weg nicht der schnellste. Die Bestimmung der SVM liegt in  $\mathcal{O}(n^3)$ , wobei  $n$  die Anzahl der Instanzen in der Trainingsmenge ist. Intuitiv lässt sich eine untere Grenze für den Rechenaufwand wie folgt sehen: hätte man ein Orakel, welches alle Nicht-Supportvektoren vorab identifizieren könnte, so müsste man zu Bestimmung der verbleibenden Koeffizienten  $\alpha_i$  für die  $r$  Supportvektoren ein lineares Gleichungssystem mit  $r$  Gleichungen lösen. Hierfür sind typischerweise  $r^3$  Operationen notwendig (Bottou et al. 2007). Verschiedene Verfahren wurden entwickelt, um insbesondere große SVMs schneller zu bestimmen. Einen Überblick gibt beispielsweise (Bottou et al. 2007). Ein Verfahren von (Platt 1999) soll an dieser Stelle kurz vorgestellt werden, da dieses für die Supportvektorbestimmung in dieser Arbeit adaptiert worden ist.

**Sequentielle Minimierung** Platt (1999) beschreibt ein Verfahren zum schnellen Training von Supportvektor Maschinen im Allgemeinen: die Sequential Minimal Optimization (SMO). Grundidee ist hier, das (große) QP-Problem aufzuspalten in eine Reihe von kleinstmöglichen QP-Problemen, die dann einzeln und vor allem analytisch gelöst werden können.

SMO beruht auf einer Dekomposition des großen QP-Problems in Teilprobleme, die nur noch jeweils zwei Lagrange-Multiplikatoren  $\alpha_1$  und  $\alpha_2$  betreffen. Auf diese Weise wird – im Gegensatz zu anderen Verfahren – kein numerisches Lösungsverfahren für die QPs mehr benötigt. Das Verfahren wählt in einem iterativen Vorgehen jeweils zwei Multiplikatoren aus und optimiert diese gemeinschaftlich. Dies ist zulässig, da die beiden so abgeändert werden, dass der Zielfunktionswert im Teilbereich dieser Multiplikatoren verbessert wird und alle Lagrange-Multiplikatoren stets zulässige Werte haben. Dadurch wird der Gesamtzielfunktionswert in jedem Schritt verbessert, sodass das Verfahren asymptotisch konvergiert (Platt 1999). Die Werte der jeweils gemeinschaftlich betrachteten zwei Multiplikatoren müssen zusammen lineare Constraints erfüllen. Für ein Problem mit zwei Multiplikatoren können Optima analytisch schnell gefunden werden. Allerdings ist für

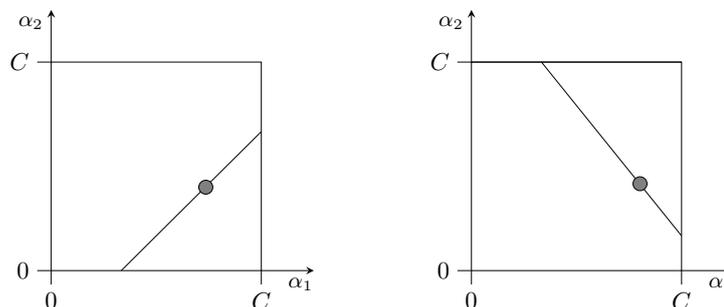


Abbildung 5.1.: Lage der möglichen Diagonalsegmente bei der SMO für den Fall dass beide Multiplikatoren zu Datenpunkten aus verschiedenen Klassen gehören ( $y_1 \neq y_2 w$ ) auf der linken Seite ( $\alpha_1^n - \alpha_2^n = \alpha_1^{n-1} + y_1 y_2 \alpha_2^{n-1}$ ) bzw. für den Fall, dass beide der selben Klasse angehören ( $\alpha_1^n + \alpha_2^n = \alpha_1^{n-1} + y_1 y_2 \alpha_2^{n-1}$ ) auf der rechten.

eine weitere Geschwindigkeitssteigerung eine heuristische Auswahl dieser Teilprobleme hilfreich, wie sie ebenfalls von Platt (1999) beschrieben wird.

Für lediglich zwei Lagrange-Multiplikatoren lässt sich die Situation leicht darstellen (vgl. Abbildung 5.1). Aufgrund der Constraints in Gleichung 5.8, welche die möglichen Werte für die Multiplikatoren einschränken. Die zulässigen Werte für  $\alpha_1$  und  $\alpha_2$  (des Teilproblems) müssen so gewählt sein, dass  $(\alpha_1, \alpha_2)$  einen Punkt auf einem der Liniensegmente in Abbildung 5.1 ergibt. Nur auf diesem braucht das Verfahren nach dem Optimum für die Zielfunktion in Gleichung 5.8 zu suchen. Das exakte Vorgehen hierfür ist beispielsweise ausführlich beschrieben in (Platt 1999) bzw. in (Platt 1998).

Bei diesem Vorgehen muss zwar insgesamt eine größere Zahl an Teilproblemen gelöst werden als dies bei anderen Zerlegungen der Fall ist, jedoch ist die analytische Lösung der Teilprobleme sehr viel schneller als die numerische Lösung größerer Teilprobleme. Insgesamt ergibt sich somit ein Geschwindigkeitsvorteil von bis zu 1 : 15 (Platt 1998).

SMO setzt sich zusammen aus zwei wesentlichen Teilen: dem soeben betrachteten Schritt zur Bestimmung der Werte für ein Paar von Multiplikatoren und zum anderen aus einer Heuristik für die geschickte Auswahl von Multiplikatorpaaren je Schritt.

### 5.2.3. Anwendungen Kern-basierter Verfahren

Die klassische SVM hat aufgrund ihrer guten Performanz bei praktischen Problemen in den letzten Jahren Anwendung auf vielfältigen Gebieten gefunden. Standardanwendung ist die Klassifikation basierend auf einem Modell gelernt auf Basis gelabelter oder teilweise gelabelter Daten (Gieseke et al. 2012). Anwendungsfelder im Bereich Smart Grid finden sich zum Beispiel in der Erkennung von Fehlerzuständen im Übertragungsnetz anhand von Messwerten (Shahid et al. 2012), in der Klassifikation zulässiger Betriebszustände (Blank et al. 2011) sowie im Bereich der Geräteerkennung mit Smart Metern (Mittelsdorf et al. 2013).

In den letzten Jahren wurde das Prinzip der Ausnutzung des Kern-Tricks zur Abbildung von nicht-linearen Problemfeldern auf wohl bekannte Lösungsansätze für eine Reihe weite-

rer Anwendungsfelder erschlossen. Einige dieser Algorithmen sollen hier kurz erläutert werden, da sie als verwandte Bereiche zumindest marginale Bedeutung für den in dieser Arbeit entwickelten Ansatz haben.

### Kern-basierte Regression

Für die Abschätzung eines Modells des funktionalen Zusammenhangs eines gegebenen Datensatzes werden Supportvektorverfahren seit einiger Zeit ebenfalls erfolgreich eingesetzt. In der Regression (oder Ausgleichsrechnung) werden die Parameter eines Modells so optimiert, dass eine bestmögliche Anpassung des Modells an eine Reihe gegebener Daten erreicht wird. Mithilfe eines Modells einer Funktion untersucht die Regression, wie sich Datenpunkte beschreiben und fortschreiben lassen. Als robustes nicht lineares Regressionsverfahren kann die Supportvektor Regression (SVR) (Smola & Schölkopf 2004) zum Einsatz kommen. Beispielsweise wird bei der  $\epsilon$ -SV Regression (Vapnik 1995) eine Funktion  $f(x)$  abgeschätzt, die zu einer gegebenen Menge von Datenpunkten  $x_i \in X$  einen Fehler der Größe von maximal  $\epsilon$  liefert und gleichzeitig möglichst flach ist. Für die SVR wird eine lineare Regression zur Abschätzung von

$$f(x) = \langle w, x \rangle + b, \text{ mit } w \in X, b \in \mathbb{R}$$

im RKHS vorgenommen durch Ersetzen des Skalarproduktes durch einen geeigneten Kern, welches schließlich einen Regressionsschätzer der Form

$$f(x) = \sum_{i=1}^N (\alpha_i^* - \alpha_i) k(x_i, x) + b \quad \text{mit } \alpha_i^*, \alpha_i \text{ Lagrange'sche Multiplikatoren}$$

liefert. Praktische Anwendung haben derartige Modellschätzer beispielsweise in der Modellierung von Umweltdaten (Pozdnoukhov & Kanevski 2007), bei der Modellierung komplexer geometrischer Gebilde in der Bildverarbeitung (Steinke et al. 2005) oder der Vorhersage der Windenergieintensität (Kramer & Gieseke 2011). In (Yan-bin 2009) werden auf SVR basierende Surrogat-Modelle für die Optimierung eines Parks von elektrischen Generatoren unter Umweltgesichtspunkten verwendet. Eine Preisvorhersage auf Energiemärkten versuchen (Gao et al. 2007), während (Wang et al. 2008) ein Beispiel für Kurzfrist-Lastprognosen darstellt.

### Kern-basierte Hauptkomponentenanalyse

Die Hauptkomponentenanalyse (engl. Principle Component Analysis, PCA) kommt zum Einsatz, um große, hochdimensionale Datensätze strukturell zu vereinfachen durch Ersetzen möglichst vieler Merkmale durch Linearkombination möglichst weniger aussagekräftiger Merkmale. Bei der Kern-PCA wird eine lineare PCA im Feature-Raum durchgeführt (Schölkopf et al. 1999a). Praktische Anwendungen finden sich im Bereich der Bildverbesserung durch Rauschreduktion (Mika et al. 1999a) oder in der Novelty Detection (Hoffmann 2007), bei der ein Modell des Raums gelernt wird, den ein Satz von Beispieldaten einnimmt, sodass neue Daten auf Zugehörigkeit (und damit auf Ähnlichkeit zu den bereits bekannte Daten) geprüft werden können.

### Independent Component Analysis

Die Unabhängigkeitsanalyse (engl. independent component analysis, ICA) dient der Zerlegung einer Mischung statistisch unabhängiger Zufallsvariablen und stammt aus der Signalverarbeitung. Im allgemeineren Fall der Blind Source Separation ist es das Ziel, die Überlagerung verschiedener Signale wieder in die ursprünglichen Signale zu trennen. Beispiele für Kern-basierte Varianten finden sich in (Bach & Jordan 2003; Harmeling et al. 2003). Während dieser Anwendungsfall nicht speziell interessant für diese Arbeit ist, so finden sich beispielsweise in (Harmeling et al. 2002) doch einige formal interessante mathematische Hinweise welche im Rahmen dieser Arbeit hilfreich sein werden (vgl. Kapitel 9 ab Seite 117).

### 5.3. Supportvektor Data Description

Die Detektion von Abnormalität bedarf anderer Verfahren als sie bei der klassischen Klassifikation oder Regression zum Einsatz kommen (Kim et al. 2008). Typische Anwendungsfälle sind die automatische Erkennung von Ausreißern, Novelty Detektion oder die Anomalieerkennung. Allen Fällen ist gemein, dass ein hierfür verwendetes Klassifikationsverfahren in der Lage sein muss, zwischen Zugehörigkeit und Nicht-Zugehörigkeit zu einer Zielklasse zu unterscheiden. Viele herkömmliche Klassifikationsverfahren (so auch die normalen SVM) haben hierbei Schwierigkeiten, da sie eine offene Entscheidungsgrenze (zwischen verschiedenen Klassen) verwenden (Moya et al. 1993).

In Einklassenproblemen gibt es nur genau eine Klasse (die Positivklasse), die sehr gut durch vorhandene Daten charakterisiert werden kann. Meist gibt es keine oder nur sehr wenige Gegenbeispiele. Ein Anwendungsbeispiel ist das Erkennen von Fehlfunktionen einer Anlage. Beobachtungen korrekter Betriebszustände liegen üblicherweise zahlreich vor. Gegenbeispiele gibt es erst bei Eintreten einer Fehlfunktion, die aber ja gerade erkannt werden soll.

Je nach Verfügbarkeit von Daten (nur positive Trainingsdaten, wenige Negativbeispiele oder positive und ungelabelte Daten vorhanden) wurden in der Vergangenheit verschiedene Ansätze für diese Problemstellung entwickelt (Khan & Madden 2010). Während letztere Datenlage in Wesentlichen im Bereich der Text- bzw. Dokumentenklassifikation von großer Bedeutung ist, entwickelten Schölkopf et al. (1999b) und Tax & Duin (2004) jeweils ein auf einem Supportvektoransatz beruhendes Verfahren für den Fall, dass ausschließlich Positivbeispiele als Daten vorhanden sind. Obgleich beide Verfahren einen unterschiedlichen Trainingsansatz und eine unterschiedliche Repräsentation der Positivklasse verwenden, existieren geometrische Beziehungen, zwischen den Verfahren, die sie vergleichbar machen. Diese Arbeit wird sich auf das Verfahren von Tax und Duin konzentrieren.

Tax & Duin (1999) entwickelten ein Verfahren zur Modellierung von Wertebereichen höher dimensionaler Daten (Data Domain Description) basierend auf der Idee, eine beliebig geformte Region im Datenraum (welche eine gegebene Datenmenge umschließt) darzustellen als hochdimensionale aber einfach zu beschreibende Kugel (genauer: als

Hypersphäre) im wie bei Supportvektor Maschinen durch eine Kern-Funktion induzierten Featureerraum  $\mathcal{H}$ .

Dieses Verfahren zur Support Vector Data Description oder Support Vector Domain Description (SVDD) basiert auf einer einfachen Grundidee. Gesucht wird eine Abbildung, welche die Daten der Trainingsmenge so in den Featureerraum abbildet, dass die Bilder der Trainingsvektoren von einer möglichst kleinen Kugel eingeschlossen werden. Ist diese Kugel gefunden, so dient das Urbild der Kugeloberfläche als geschlossene Begrenzungsfläche um den Trainingsdatensatz herum. Je kleiner die Kugel ist, desto enger schmiegt sich das Urbild dieser Kugel an die Form des Trainingsdatensatzes an. Da die spätere Modellierung der beschränkten Suchräume ganz wesentlich auf dieser Idee aufbauen wird, werden dieses Verfahren und seine mathematischen Grundlagen in diesem Kapitel nachfolgend zunächst eingehender betrachtet.

### 5.3.1. Grundprinzip

Die nachfolgenden Erläuterungen orientieren sich u. a. an (Tax & Duin 2004; Ben-Hur et al. 2001; Bremer et al. 2010).

Sei  $\{x_i\}_n \subseteq \mathcal{X} \subset \mathbb{R}^d$  eine Menge von  $n$  gegebenen Datenpunkten mit der Dimension  $d$ . Diese wird als Trainingsmenge für das Lernen des Datenraums  $\mathcal{X}$  verwendet. Sei ferner

$$\begin{aligned} \Phi : \mathcal{X} &\rightarrow \mathcal{H}, \\ \mathbf{x} &\mapsto \Phi(\mathbf{x}) \end{aligned}$$

eine nicht-lineare Abbildung von Datenpunkten  $x_i$  aus  $\mathbb{R}^d$  in einen hoch- (möglicherweise unendlich-) dimensionalen Featureerraum  $\mathcal{H}$ . Nun beschreibt

$$\|\Phi(\mathbf{x}_i) - \mathbf{a}\|^2 \leq R^2 \quad \forall i \tag{5.12}$$

eine Kugel in  $\mathcal{H}$  welche die Bilder aller Datenpunkte  $\{\Phi(x_i)\}_n$  enthält. Hierbei bezeichne  $\|\cdot\|$  den Euklidischen Abstand und  $\mathbf{a}$  den Mittelpunkt der Kugel. Auch wenn es sich aufgrund der Dimensionalität hierbei streng genommen um eine Hypersphäre handelt wird in den folgenden Ausführungen häufiger vereinfachend auch der Begriff Kugel verwendet.

Üblicherweise wird eine solche Kugel etwas unscharf formuliert durch Einführung von Schlupfvariablen  $\xi$  (siehe hierzu auch KKT-Generalisierung des Lagrangeverfahren, Abschnitt 4.5.1 auf Seite 45):

$$\|\Phi(\mathbf{x}_i) - \mathbf{a}\|^2 \leq R^2 + \xi_i \quad \forall i. \tag{5.13}$$

Abhängig von der genauen Ausgestaltung von  $\Phi$  sind auch entsprechend viele verschiedene Kugeln vorhanden, die  $\{\Phi(\mathbf{x}_i)\}_n$  jeweils einschließen. Im nächsten Schritt stellt sich nun die Frage nach der kleinsten dieser verschiedenen Kugeln. Gesucht ist also dasjenige  $\Phi$  bei dem der Radius  $R$  in Gleichung 5.14 minimal wird.

Um dieses Optimierungsproblem zu lösen, werden zunächst Lagrange'sche Multiplikatoren  $\beta_i$  and  $\mu_i$  eingeführt, wodurch sich die erste Form für das Minimierungsproblem ergibt:

$$L(\boldsymbol{\xi}, \boldsymbol{\mu}, \boldsymbol{\beta}) = R^2 - \sum_i (R^2 + \xi_i - \|\Phi(\mathbf{x}_i) - \mathbf{a}\|^2)\beta_i - \sum_i \xi_i \mu_i + C \sum_i \xi_i. \tag{5.14}$$

In Gleichung (5.13) bzw. (5.14) repräsentieren die Schlupfvariablen  $\xi$  die Toleranz für falsch klassifizierte Datenpunkte und sind somit als Maß für die Constraintverletzungen zu sehen (Liang et al. 2009). Der Parameter  $C$  hingegen gibt einen Trade-off zwischen der Größe der Kugel und der Zahl akzeptierter Fehler an und muss vom Nutzer vorab eingestellt werden. Eine ausführlichere Diskussion zur Bedeutung und zur Wahl dieses Parameters findet sich im Anschluss in Abschnitt 5.3.4 auf Seite 68. In (Zheng et al. 2006) wird eine Erweiterung vorgeschlagen,  $C$  nicht als skalare Konstante sondern als Fuzzygröße zu behandeln.

Zur Lösung des Problems hier werden zunächst die partiellen Ableitungen von  $L$  bzgl.  $R$ ,  $\mathbf{a}$  und  $\xi_i$  auf 0 gesetzt, was zu folgenden Beziehungen führt (Ben-Hur et al. 2001):

$$\frac{\partial L}{\partial R} = 2R - \sum_i \beta_i \cdot 2R = 0 \quad \Rightarrow \quad 1 = \sum_i \beta_i \quad (5.15)$$

$$\frac{\partial L}{\partial \mathbf{a}} = \sum_i \beta_i (2 \cdot \Phi(\mathbf{x}_i) - 2\mathbf{a}) = 0 \quad \Rightarrow \quad \mathbf{a} = \sum_i \beta_i \Phi(\mathbf{x}_i) \quad (5.16)$$

$$\frac{\partial L}{\partial \xi_i} = -\beta_i - \mu_i + C \quad \Rightarrow \quad \beta_i = C - \mu_i \quad (5.17)$$

Die Karush-Kuhn-Tucker (KKT) Komplementärbedingung nach Fletcher (1987) führt zu:

$$\xi_i \mu_i = 0 \quad (5.18)$$

$$(R^2 + \xi_i - \|\Phi(\mathbf{x}_i) - \mathbf{a}\|^2) \beta_i = 0. \quad (5.19)$$

Mit Hilfe von (5.18) und (5.19) lassen sich nunmehr einige Zusammenhänge ableiten. Liegt das Bild eines beliebigen Datenpunktes  $\mathbf{x}_i$  außerhalb der Kugel folgt aus Gleichung (5.19), dass  $\xi_i > 0 \wedge \beta_i > 0$ . Aus Gleichung (5.17) lässt sich in diesem Fall schließen, dass dann  $\beta_i = C$  sein muss. Nach Gleichung (5.18) hat ein solches Datum einen Wert von  $\mu_i = 0$ .

Derartige Datenpunkte, die unter  $\Phi$  in einen Bereich außerhalb der Kugel abgebildet werden bezeichnet man als bounded Supportvektoren. Gebundene (bounded) Supportvektoren stellen prinzipiell falsch klassifizierte Vektoren aus der Trainingsmenge dar. Sie liegen, obwohl sie als Supportvektoren zur Kugeldefinition beitragen, außerhalb der Kugel.

Ein Datenpunkt  $\mathbf{x}_i$  wird genau dann in das Innere der Kugel abgebildet, wenn der zugehörige Wert  $\beta_i = 0$  ist. Im speziellen Fall  $0 < \beta_i < C$  wird der korrespondierende Datenpunkt genau auf die Oberfläche der Kugel abgebildet. Diese speziellen Datenpunkte werden als (unbounded) Supportvektoren bezeichnet. Mit Hilfe dieser Beziehungen können nun zunächst die Parameter  $R$ ,  $\mathbf{a}$  and  $\mu_i$  in der Lagrangeform (Gleichung (5.14)) eliminiert werden (vgl. hierzu beispielsweise Tax & Duin 2004; Ben-Hur et al. 2001; Liu et al. 2013; Chang et al. 2013).

$$\begin{aligned} L(\xi, \mu, \beta) &= R^2 + C \sum_i \xi_i - \sum_i (R^2 + \xi_i - \|\Phi(\mathbf{x}_i) - \mathbf{a}\|^2) \beta_i - \sum_i \xi_i \mu_i \\ &= R^2 \left( 1 - \sum_i \beta_i \right) + \sum_i \xi_i (C - \beta_i - \mu_i) + \sum_i \beta_i (\|\Phi(\mathbf{x}_i) - \mathbf{a}\|^2) \\ &= R^2 \left( 1 - \sum_i \beta_i \right) + \sum_i \xi_i (C - \beta_i - \mu_i) + \sum_i \beta_i (\|\Phi(\mathbf{x}_i) - \sum_i \beta_i \Phi(\mathbf{x}_i)\|^2) \end{aligned} \quad (5.20)$$

Hieraus ergibt sich schließlich die Wolfe Dualform (5.21) des ursprünglichen Optimierungsproblems zur Bestimmung der Kugel (vgl. Ben-Hur et al. 2001).

$$W(\beta) = \sum_i \Phi(x_i)^2 \beta_i - \sum_{i,j} \beta_i \beta_j \Phi(x_i) \Phi(x_j). \quad (5.21)$$

Diese ist nur noch eine Funktion des Vektors  $\beta$ . Gesucht ist  $\beta$ , welcher Gleichung (5.21) maximiert, sodass

$$\begin{aligned} 0 &\leq \beta_i \leq C \\ \sum_i \beta_i &= 1. \end{aligned} \quad (5.22)$$

An diesem Punkt kann auf das bereits in Abschnitt 5.2.1 auf Seite 55 beschriebene Theorem von Mercer (Kern-Trick) zurückgegriffen werden, um die Berechnung von Skalarprodukten in  $\mathcal{H}$  durch geeignete Kernfunktionen direkt in  $\mathbb{R}^d$  durchzuführen (Schölkopf et al. 1999a):  $\Phi(x_i) \cdot \Phi(x_j) = k(x_i, x_j)$ .

Für die SVDD hat es sich allgemein als vorteilhaft erwiesen, einen Gaußkern zu verwenden (Ben-Hur et al. 2001; GhasemiGol et al. 2009):

$$k(\mathbf{x}_i, \mathbf{x}_j) = e^{-\frac{1}{2\sigma^2} \|\mathbf{x}_i - \mathbf{x}_j\|^2}, \quad (5.23)$$

da dieser eine bessere Anpassung an den Umriss der Datenmenge liefert.  $\sigma$  bezeichnet hier die so genannte Breite oder Bandbreite des Gaußkern und korrespondiert mit dem in der Statistik geläufigeren Begriff Standardabweichung. Alle anderen Mercer-Kerne können prinzipiell jedoch auch verwendet werden.

Ersetzt man alle Skalarprodukte in Gleichung (5.21) durch  $k$  so ergibt sich schließlich Gleichung

$$W(\beta) = \sum_i k(x_i, x_i) \beta_i - \sum_{i,j} \beta_i \beta_j k(x_i, x_j), \quad (5.24)$$

welche zu maximieren ist. Diese Suche nach dem Maximum kann vollständig in  $\mathbb{R}^d$  berechnet werden, ohne  $\Phi$  kennen zu müssen.

Ergebnis dieser Optimierung ist eine Belegung für den Vektor  $\beta$  der Lagrange'schen Multiplikatoren als Gewichtungsvektor mit der folgenden Bedeutung (Tax & Duin 1999; Ben-Hur et al. 2001):

- $\|\mathbf{x}_i - a\|^2 < R^2 \Rightarrow \beta_i = \xi_i = 0$ . Diese Trainingsvektoren werden in die Sphäre hinein abgebildet. Sie können (wg.  $\beta_i = 0$ ) für die Beschreibung vernachlässigt werden.
- $\|\mathbf{x}_i - a\|^2 = R^2 \Rightarrow 0 < \beta_i < C \wedge \xi_i = 0$ . Alle Trainingsvektoren  $\mathbf{x}_i$  mit  $0 < \beta_i < C$  werden auf die Oberfläche der Sphäre abgebildet. Sie werden für die Beschreibung benötigt. Diese sind die Supportvektoren.
- $\|\mathbf{x}_i - a\|^2 > R^2 \Rightarrow \beta_i = C \wedge \xi_i > 0$ . Im Fall schlecht separierbarer Cluster in der Trainingsmenge kann es sinnvoll sein,  $C < 1$  zu wählen, um Ausreißer zuzulassen (Ben-Hur et al. 2001). Solche Trainingsvektoren, die dann außerhalb der Sphäre liegen werden als gebundene Supportvektoren bezeichnet.

Hierdurch ist nach Gleichung (5.16) auch der Mittelpunkt  $a$  der gesuchten Kugel gegeben als Expansion im Hilbertraum:  $a = \sum_i \beta_i \Phi(\mathbf{x}_i)$ .

Der Abstand des Bildes (unter  $\Phi$ ) eines beliebigen Datenpunktes  $z \in \mathbb{R}^d$  vom Mittelpunkt  $a$  der Kugel in  $\mathcal{H}$  lässt sich bestimmen durch

$$\begin{aligned}
 R^2(z) &= \|\Phi(z) - a\|^2 \\
 &= \|\Phi(z) - \sum_i \beta_i \Phi(x_i)\|^2 \\
 &= \Phi(z) \cdot \Phi(z) - 2 \cdot \Phi(z) \sum_i \beta_i \Phi(x_i) + \left(\sum_i \beta_i \Phi(x_i)\right)^2 \\
 &= \Phi(z) \cdot \Phi(z) - 2 \sum_i \beta_i \Phi(x_i) \cdot \Phi(z) + \sum_i \beta_i \Phi(x_i) \sum_i \beta_i \Phi(x_i) \\
 &= \Phi(z) \cdot \Phi(z) - 2 \sum_i \beta_i \Phi(x_i) \cdot \Phi(z) + \sum_i \beta_i \Phi(x_i) \sum_j \beta_j \Phi(x_j) \\
 &= \Phi(z) \cdot \Phi(z) - 2 \sum_i \beta_i \Phi(x_i) \cdot \Phi(z) + \sum_{i,j} \beta_i \beta_j \Phi(x_i) \Phi(x_j)
 \end{aligned} \tag{5.25}$$

Auch hier lassen sich wieder die Skalarprodukte in  $\mathcal{H}$  berechnen durch Ersetzen mit der für die Optimierung verwendeten Kernfunktion, sodass auch hier  $\Phi$  nicht bekannt sein muss:

$$R^2(z) = k(z, z) - 2 \sum_i \beta_i k(x_i, z) + \sum_{i,j} \beta_i \beta_j k(x_i, x_j). \tag{5.26}$$

Im Falle eines Gaußkerns (5.23) vereinfacht sich Gleichung (5.26) weiter zu

**Definition 5.4** (Abstand eines Bildes vom Kugelzentrum).

$$R^2(z) = 1 - 2 \sum_i \beta_i k(x_i, z) + \sum_{i,j} \beta_i \beta_j k(x_i, x_j). \tag{5.27}$$

Die Funktion  $R$  gibt also den Abstand des Bildes (unter der Abbildung mit  $\Phi$ ) eines beliebigen Datenpunktes vom Mittelpunkt der Hypersphäre in  $\mathcal{H}$  an. Diese Abstandsfunktion kann nun als Entscheidungsfunktion für die Klassifikation wie folgt verwendet werden: bekannt ist die Menge der Supportvektoren. Diese werden unter  $\Phi$  auf die Oberfläche der Kugel abgebildet; ihr Abstand ist somit gerade der Radius der Kugel  $R_S$ . Somit ist über diese Beziehung der Radius der Kugel bekannt. Ein beliebiger Datenvektor gehört zur Positivklasse genau dann, wenn er unter  $\Phi$  in die Kugel oder auf die Kugeloberfläche abgebildet wird; wenn also sein Bild einen Abstand nach Gleichung (5.26) kleiner oder gleich des Radius der Kugel hat:

$$\mathcal{X} = \{x \mid R(x) \leq R_S\}. \tag{5.28}$$

### 5.3.2. Geometrische Interpretation

Der zuvor beschriebene Sachverhalt soll an dieser Stelle zunächst noch einmal graphisch dargestellt werden. Abbildung 5.2 zeigt zunächst schematisch die Verhältnisse in  $\mathcal{H}$ . Auch wenn der Featureerraum bei Verwendung eines Gaußkerns unendlichdimensional ist, so werden doch alle Trainingsvektoren auf eine endlichdimensionale Mannigfaltigkeit in  $\mathcal{H}$  abgebildet. Diese Mannigfaltigkeit ist maximal  $n$ -dimensional bei  $n$  linear unabhängigen

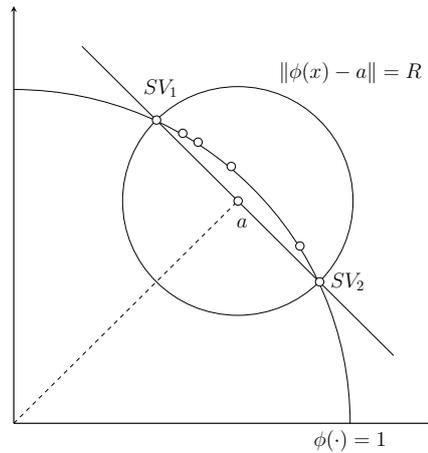


Abbildung 5.2.: Geometrische Zusammenhänge in  $\mathcal{H}$  und Lage der Hyperkugel, nach (Kim et al. 2008), verändert.

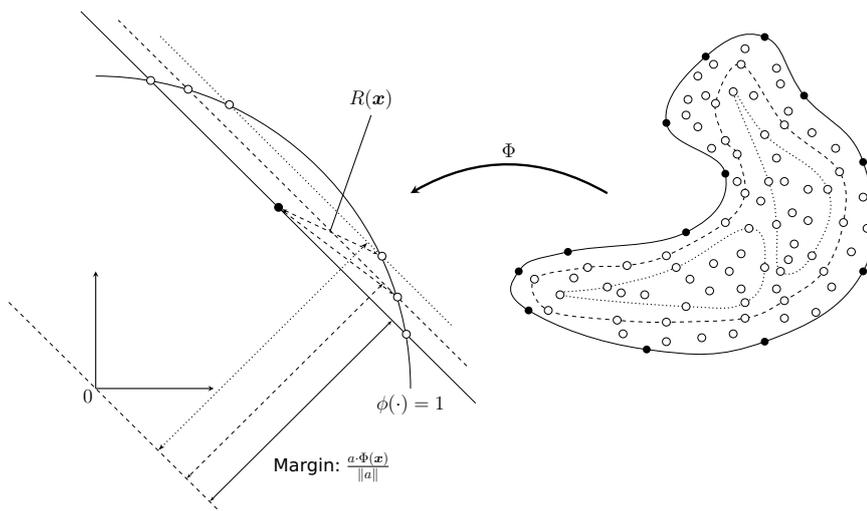


Abbildung 5.3.: Abbildung der Trainingsmenge unter  $\Phi$  und Zusammenhang zwischen Margin und Abstand zum Mittelpunkt in  $\mathcal{H}$ , nach (Kim et al. 2008), verändert.

Trainingsvektoren. Alle Bilder werden unter  $\Phi$  abgebildet auf den Einheitskreis in  $\mathcal{H}$  (Kim et al. 2008). Die den Datenraum beschreibende SVDD Kugel ist die kleinere Kugel in Abbildung 5.2 und schneidet die Einheitskugel in den Supportvektoren. Die Hyperebene durch Supportvektoren und Kugelmittelpunkt  $a$  zeigt die Verwandtschaft zur 1-Klassen-SVM (von Schölkopf et al. 1999b), wo diese zur Abgrenzung der Positivklasse verwendet wird. Abbildung 5.3 zeigt den Zusammenhang zwischen  $R$  und der Lage der Urbilder im ursprünglichen Datenraum. Es wird insbesondere noch einmal die Lage der Trainingspunkte auf der Einheitshyperkugel in  $\mathcal{H}$  dargestellt und verdeutlicht, dass der Abstand in  $\mathcal{H}$  nicht direkt mit einem Abstand in  $\mathbb{R}^d$  übersetzt werden darf. Die Linien im rechten Teil der Abbildung kennzeichnen Punkte deren Bilder in  $\mathcal{H}$  auf demselben Breitengrad auf der Einheitshyperkugel liegen. Sie haben den gleichen Abstand vom Zentrum.

### 5.3.3. Lösung des Optimierungsproblems zur Bestimmung der Hypersphäre

Die Bestimmung des Maximums von Gleichung (5.21) mit dem die Gewichte  $\beta$  bestimmt werden, welche schließlich die Unterteilung in Support- und nicht Supportvektoren ermöglichen sowie deren individuellen Beitrag zur Klassengrenze definieren, ist allgemein in  $\mathcal{O}(n^3)$  für  $n$  Trainingsvektoren lösbar. Für große Datenmengen ist diese Rechenzeit ggf. unbefriedigend. In der Vergangenheit wurden daher bereits verschiedene Näherungsverfahren entwickelt, um hier eine Beschleunigung zu erreichen. Die wichtigsten sollen im Folgenden kurz vorgestellt werden.

Die Lösung des zuvor beschriebenen Optimierungsproblems zur Bestimmung der Supportvektoren, welche die minimale Kugel definieren, erfordert wie bei der herkömmlichen SVM auch die Lösung eines Quadratischen Programms (QP) und somit die Lösung eines speziellen Optimierungsproblems aus der Komplexitätsklasse  $\mathcal{O}(n^3)$  für  $n$  Trainingsdaten (Platt 1999; Tavakkoli et al. 2008b; Liang et al. 2009). Die Speicherkomplexität liegt in  $\mathcal{O}(n^2)$ . Für größere Trainingsdatensätze stellt insbesondere die kubisch wachsende Berechnungszeit ein Problem da. Daher finden sich in der Literatur eine Reihe unterschiedlicher Verfahren zur näherungsweisen Bestimmung.

Eine der ersten Methoden geht bereits auf Vapnik (1982) zurück und ist seitdem als Chunking-Methode bekannt (Platt 1998). Dieses Verfahren macht sich zunutze, dass die quadratische Form die gleiche bleibt, wenn aus der Matrix Reihen und Spalten entfernt werden, die zu Lagrange Multiplikatoren mit einem Wert von 0 gehören. Ziel ist es letztlich, alle 0 Multiplikatoren zu entfernen und nur diejenigen mit einem Wert  $> 0$  zu identifizieren. In jedem Schritt wird ein verhältnismäßig kleines QP gelöst aus den bereits identifizierten nicht 0 Multiplikatoren und einigen weiteren Datenpunkten aus dem Trainingsatz, welche die KKT-Bedingungen am stärksten verletzen (vgl. auch (Boser et al. 1992)).

Einige weitere Vorgehensweisen, die im Rahmen dieser Arbeit eher Bedeutung haben sollen im folgenden kurz erläutert werden.

**Problemzerlegung durch Clustering** Kim et al. (2007) schlagen eine Partitionierung des Trainingsproblems basierend auf einer Zerlegung in Cluster vor. Hierfür wird die Menge der Trainingsdaten zunächst mit einem k-Means-Clusteringansatz auf mehrere Teilmengen aufgeteilt. Anschließend kann auf jede dieser Teilmengen das herkömmliche SVDD-Verfahren angewendet werden. Der Geschwindigkeitsvorteil resultiert dann aus der Tatsache, dass die Lösung mehrerer kleinerer Optimierungsprobleme aufgrund der kubischen Komplexität in Summe immer schneller ist als die Lösung eines großen. Nun könnte als Klassifikator bereits die Menge der einzelnen Klassifikatoren genutzt werden, da eine Klassenzugehörigkeit genau dann vorliegt, wenn eine Zugehörigkeit zu einer der Teilklassen festgestellt werden kann. Die Autoren schlagen jedoch noch vor, als weiteren Schritt eine Vereinigung der Teilklassen vorzunehmen indem ein zusätzlicher SVDD-Lernvorgang mit der Vereinigung der Mengen aller Supportvektoren aus den Teilklassifikatoren vorgenommen wird. Hierbei dürfte allerdings die Wahl Verfahrensparametrierung essentiell sein, um keine Löcher in der gelernten Region zu produzieren mit unnötigen falschen Negativklassifikationen. Aufgrund dieser Problematik scheint dieses Verfahren zur Beschleunigung des Lernvorgangs nur bedingt geeignet.

**Inkrementelles Lernen** Beim inkrementellen Lernen werden von dem entsprechenden Verfahren nicht alle Trainingsdaten auf einmal verarbeitet. Vielmehr werden dem Lernprozess die Trainingsdaten sukzessive zugeführt. Auf diese Weise erfolgt mit jedem neuen Trainingsdatum ein Update der bis dato gefundenen Hypersphäre.

Tavakkoli et al. (2008b) schlagen hierzu vor, die KKT-Bedingungen zu nutzen. Basierend auf (Osuna et al. 1997) kann ein großes Problem der Quadratischen Programmierung zerlegt werden in eine Reihe kleinerer Probleme, welche einfacher zu lösen sind. Bedingung für ein Konvergieren ist die Verletzung mindestens einer KKT-Bedingung beim Hinzufügen eines neuen Trainingsdatums. Das Arbeitstrainingsset besteht lediglich aus den (bisher gefundenen) Supportvektoren. Wird ein neuer Trainingsvektor hinzugefügt, der zur Positivklasse gehört (Bild innerhalb der Kugel), dann erfüllt dieser Datenpunkt per Definition die KKT-Bedingungen. Ein Hinzufügen dieses Punktes zur aktuellen Arbeitstrainingsmenge (bestehend nur aus Supportvektoren) würde nach (Osuna et al. 1997) somit das aktuell gefundene Minimum der SVDD-Trainings-Zielfunktion nicht weiter verringern. Wird also ein Datum hinzugefügt, welches innerhalb der Kugel liegt, die durch die aktuell bereits bestimmten Supportvektoren definiert wird, so wird es nicht weiter beachtet. Es wird bereits korrekt klassifiziert und bringt für die Beschreibung keine weitere Information.

Erfüllt das neue Datum die KKT-Bedingungen jedoch nicht, so wird das Datum zum Arbeitstrainingsset hinzugefügt und das Trainings-Optimierungsproblem für das neue Arbeitstrainingsset gelöst. Tavakkoli et al. (2008b) schlagen hierzu vor, analog zur sequentiellen Minimierung vorzugehen. Da nur das neue Datum die KKT-Bedingungen verletzt, können sukzessive die jeweils kleinsten Optimierungsprobleme – bestehend aus dem neuen und jeweils einem alten Datum – aus dem Trainingsset gelöst werden. Dies ist analytisch möglich.

Im Gegensatz zu diesem Vorgehen schlagen Kim et al. (2008) vor, Informationen über die Margin eines Trainingsdatums zu verwenden und definieren eine  $\epsilon$ -Umgebung innerhalb derer keine Trainingspunkte entfernt werden dürfen, da diese aufgrund ihrer Nähe zu den bisher gefundenen Supportvektoren selbst noch eine Chance haben im Zuge des fortschreitenden Lernprozesses zu Supportvektoren zu werden.

Der große Vorteil inkrementeller Vorgehensweisen liegt darin, dass nicht das gesamte Trainingsset bearbeitet werden muss. Es muss nicht einmal vollständig bekannt sein, um mit dem Training beginnen zu können. Auf diese Weise wird es möglich, beispielsweise das Training online anhand eines Datenstroms von sequentiell ankommenden Daten vorzunehmen. Aber auch wenn die Trainingsmenge zu Beginn vollständig bekannt ist, ergibt sich ein Vorteil: Das Training kann sukzessive mit den Daten aus der Trainingsmenge durchgeführt werden und (beinahe) jederzeit mit dem bis dato erreichten Ergebnis abgebrochen werden (analog einem evolutionären Vorgehen bei der Suche nach einem Optimum). Da jederzeit ein gültiges Ergebnis existiert (zumindest näherungsweise, da es auf einer Teilmenge der Trainingsmenge basiert) kann der Vorgang beispielsweise bei Erreichen einer Deadline aus gegebenen Echtzeitbedingungen mit einem gültigen Ergebnis abgebrochen werden.

### Sequential Minimization für SVDD

Im Rahmen dieser Arbeit wird als Verfahren SMO (vgl. Abschnitt 5.2.2 auf Seite 57) in einer Eigenimplementierung zum Einsatz kommen. Die Implementierung orientiert sich an (Platt 1998) und wird in der Programmiersprache Java umgesetzt. Als wesentliche Anpassung an dem in (Platt 1998) vorgeschlagenen Pseudocode für eine Standard-SVM muss im Wesentlichen die Implementierung der Ausgabe des Klassifikators an die SVDD-Entscheidungsfunktion angepasst werden.

### 5.3.4. Parametrierung

Praktisch alle Verfahren zur Erstellung eines Modells für gegebene Datenklassen erlauben eine gewisse Parametrierung wodurch sich die Güte des Modells (und ggf. auch andere Eigenschaften) beeinflussen lassen. Somit stellt sich die Suche nach einem Satz an Parameterwerten, der Modelle mit gewünschten Eigenschaften erzeugt, als eine Optimierungsfragestellung dar. In der Regel ist es das Ziel, ein Modell mit einer möglichst guten Vorhersagegüte zu erhalten. Für Klassifikatoren wird meist die Genauigkeit (engl.: accuracy) als Kriterium für die Güte verwendet, wobei diese häufig empirisch über die Falschklassifikationsrate gemessen wird (Steinwart & Christmann 2008). Meist kommt hierbei auch zusätzlich ein Kreuzvalidierungsansatz zum Einsatz. Hierbei wird der verfügbare Datensatz in  $n$  gleich große Teildatensätze unterteilt, wobei jeweils ein Teildatensatz zur Validierung und die übrigen als Trainingsdaten fungieren. Die Güte ergibt sich schließlich als Mittelwert. In der Regel wird mit einer 10-fach Kreuzvalidierung gearbeitet.

Für die systematische Suche nach einem geeigneten Satz an Parameterwerten können verschiedene Verfahren herangezogen werden. Üblich sind z. B. die Gridsuche (siehe z. B. Jiménez et al. 2008) bei der die Wertebereiche der einzelnen Parameter in äquidistante (oder geometrisch verteilte) Werte unterteilt werden, um alle Kombinationen systematisch zu testen, eine Zufallssuche, das Nelder-Mead Optimierungsverfahren (Nelder & Mead 1965) basierend auf einem Simplex über den Parametern oder auch der Einsatz heuristischer Ansätze für die Suche (Steinwart & Christmann 2008).

### Ausreißerakzeptanzrate

Der SVDD selbst ist nur der Parameter  $C$  als vom Benutzer vorzugebende Stellgröße zuzuordnen. Der Parameter lässt einen Tradeoff zwischen der Größe der Beschreibung und dem akzeptierten Fehler definieren (GhasemiGol et al. 2010); ein Umstand, der insbesondere bei Datensätzen mit sich teilweise überlappenden Clustern hilfreich ist, um eine Beschreibung zu erhalten, die diese beiden Cluster sauber voneinander trennt – z. B. bei der Anwendung im Clustering (Ben-Hur et al. 2001). Je kleiner  $C$  gewählt wird, desto beschränkter wird der Wertebereich für die Lagrange-Multiplikatoren  $\beta_i$  ( $\beta_i \leq C$ , vgl. Gleichung (5.17)) und desto mehr Trainingsinstanzen werden zu gebundenen Supportvektoren, also als Ausreißer behandelt. Hierdurch verkleinert sich der Durchmesser der beschreibenden Hypersphäre bei gleichzeitiger Vergrößerung des Fehlers (Tax & Duin 2004). Wenn keine Ausreißer im Datensatz zu erwarten sind, so kann  $C = 1$  gewählt werden (Tax & Duin 2004).

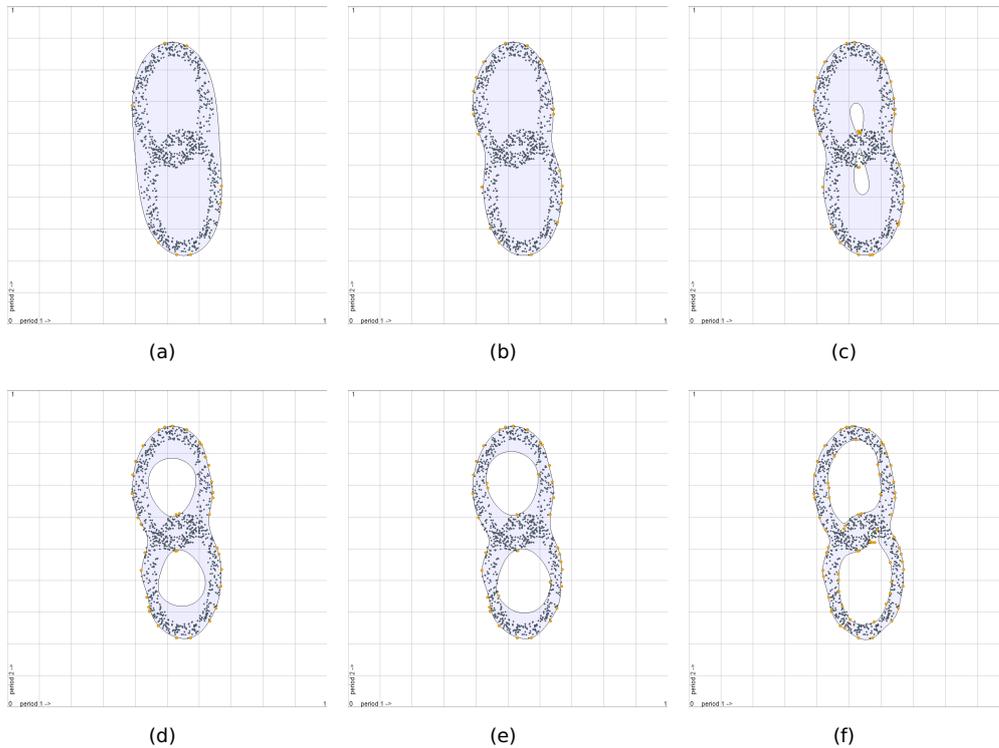


Abbildung 5.4.: Bedeutung der Bandbreite (beispielhaft von 0.32 in Abbildung 5.4(a) bis 0.05 in Abbildung 5.4(f)) des Kerns für die Angleichung der Entscheidungsgrenze an die Stichprobe (Bremer et al. 2011a).

Falls Ausreißer im Datensatz existieren – oder zumindest als existent angenommen werden müssen – so ist dieses Problem bei höherdimensionalen Problemstellungen (wie in dieser Arbeit betrachtet) weniger stark ausgeprägt, da mit steigender Dimension das Volumen der Negativklasse dazu tendiert stärker zu wachsen als das der Positivklasse (Tax & Duin 2004), wodurch mögliche Überlappungen kleiner werden. Das Klassifikationsproblem wird einfacher. Mit steigender Dimensionalität wächst auch die Größe der Begrenzungsfläche um die Positivklasse stark an, sodass mit steigender Dimensionalität auch mehr Trainingsdaten erforderlich sind (Tax & Duin 2004; Khan & Madden 2010). Probleme können auch entstehen, wenn es zu große Unterschiede in der Dichte der Trainingsdaten gibt, was in dünn besiedelten, höherdimensionalen Daten durchaus der Fall ist. Instanzen aus dünn besiedelten Regionen werden gegebenenfalls als Ausreißer behandelt, obwohl sie eigentlich korrekte Daten sind (Khan & Madden 2010). Dieser Umstand ist bei der Wahl des Parameters  $C$  zu beachten.

Dieser Parameter kann somit auch interpretiert werden als Rate von fälschlich als nicht zur Klasse gehörig klassifizierter Werte und somit als Toleranz des Klassifikators (Tavakkoli et al. 2008a).

### Bandbreite des Gaußkerns

Der wesentlich interessantere Parameter (zumindest für den in dieser Arbeit betrachteten Anwendungsfall) ist die Bandbreite  $\sigma$  in dem verwendeten Gauß-Kern. Dieser Parameter bestimmt die Größe der Hypersphäre (Lee & Daniels 2012) und damit wie eng sich die resultierende Entscheidungsgrenze an den tatsächlichen Datensatz anschmiegt. Abbildung 5.4 gibt ein anschauliches Beispiel. Ein weiteres Beispiel mit einem auf verschiedene Cluster aufgeteilten Beispieldatensatz ist in Abbildung 5.5 gegeben. Die hellblaue Fläche in den Abbildungen kennzeichnet jeweils die Region, die als durch die Trainingsmenge definiert klassifiziert wird. Bei dem zweiten Beispiel wird die Bedeutung der Bandbreite  $\sigma$  im Falle disjunkter Teilbereiche der Positivklasse deutlich. Bei ungünstig gewählter Bandbreite wie in den Abbildungen 5.5(a) bis 5.5(c) wird auch der Bereich zwischen den disjunkten Teilregionen zumindest teilweise als zur Positivklasse gehörig klassifiziert.

Ein weiterer Effekt lässt sich im Zusammenhang mit der Bandbreite beobachten: Eine Vergrößerung von  $\sigma$  führt zu einer geringeren Anzahl an Supportvektoren und zu einem größeren Fehler bei der Klassifikation (Tax & Duin 2004).

### 5.3.5. Anwendungen für die SVDD

Abschließend sollen nun noch einige Anwendungsfälle für das SVDD-Verfahren betrachtet werden. Die SVDD findet mittlerweile bei einer Reihe verschiedener Anwendungen Verwendung. Das Problem der 1-Klassen-Klassifikation wird häufig durch Abschätzen der Dichte der Daten gelöst oder durch anpassen eines Supportvektor-Modells (Khan & Madden 2010). Tax und Duin (Tax & Duin 1999; Tax & Duin 2004) haben hierfür speziell eine SVM Variante entwickelt, welche die einzelne, positive Klasse von Daten von allen anderen Mustern unterscheidet: die SVDD wie sie zuvor vorgestellt wurde. In einem verwandten Ansatz entwickelten Schölkopf et al. (1999b) eine Methode, die Bereiche mit Daten von denen ohne zu separieren durch Bestimmen einer maximal vom Ursprung (in  $\mathcal{H}$ ) entfernten Hyperebene, sodass die Bilder aller Trainingsdaten dennoch weiter vom Ursprung entfernt sind. Trotz Verwendung einer Trainingsmenge von Datenpunkten aus nur einer Klasse, ist das Ergebnis dieses Ansatzes aufgrund der Verwendung einer trennenden Hyperebene (wie bei der SVM) wieder ein binärer Klassifikator. Dieser Alternativansatz ist für diese Arbeit nicht von spezieller Bedeutung und nur der Vollständigkeit halber aufgeführt. Die Erweiterung von Manevitz & Yousef (2002) schlägt noch vor identifizierte Ausreißer als zweite Klasse zu verwenden. Allen gemein ist ihre Verwendung in Fällen, in denen fast ausschließlich positive Daten vorliegen. Dies ist z. B. der Fall bei der Fehlererkennung, wenn ein Fehler nur sehr selten auftritt und beinahe alle Beobachtungen sich auf den Nicht-Fehlerfall beziehen.

### Novelty Detection

Das Erkennen von neuen, bisher ungesehenen Mustern ist die Aufgabe der Novelty Detection. Im Falle der SVDD wird der Klassifikator trainiert mit einer Menge von gewöhnlichen Mustern, um einen Entscheidungsmechanismus zu gewinnen, der bei neuen Mustern entscheiden kann, ob diese zur Verteilung der normalen Muster passen oder als neu

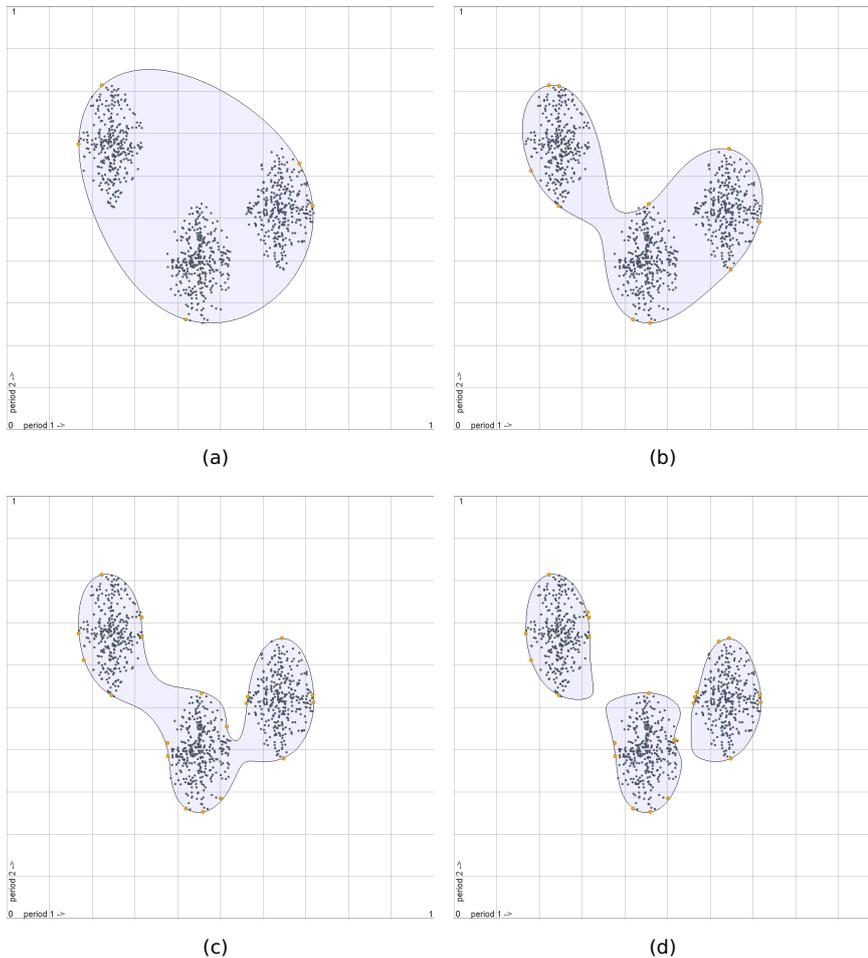


Abbildung 5.5.: Bedeutung der Bandbreite des Kerns (beispielhaft von 0.32 in Abbildung 5.5(a) bis 0.09 in Abbildung 5.5(d)) für den Zusammenhang einzelner Teilregionen der Stichprobe (Bremer et al. 2011a).

(und damit ungewöhnlich) zu klassifizieren sind. Anwendungsbeispiel ist das Erkennen ungewöhnlicher Vibrationen in Flugzeugmotoren (Hayton et al. 2000) oder die Überwachung von Umweltindikatoren (Rapp & Bremer 2012). Ein spezieller Anwendungsfall ist das Erkennen und die Trennung von Hintergrund und Vordergrund beispielsweise in Filmen (Tavakkoli et al. 2008a).

### Bildbearbeitung

Wesentlich stärker mit dem Thema dieser Arbeit verwandt ist die Anwendung der SVDD in der Bildbearbeitung. Hier werden 3-dimensionale Gegenstände mithilfe eines aus einer Punktwolke (z. B. von einem Laserscanner) trainierten SVDD-Modells modelliert, um eine kontinuierliche Figur zu erhalten. Beispiele für solche Modelle finden sich u. a. in (Schölkopf et al. 2004; Sen et al. 2008).

Ein weiterer Anwendungsfall in der Bildbearbeitung ist das Entrauschen von Bildern. Hier wird ein SVDD Modell mit einer Menge von unverrauschten Mustern trainiert. Mit einem verrauschten oder teilweise beschädigten Bild kann anschließend durch eine geodätische Projektion im Featureraum eine nahe gelegene, unverrauschte Instanz gefunden werden (Park et al. 2007; Park et al. 2006).

### **Clustering**

Das Supportvektor Clustering (Ben-Hur et al. 2001) ist eine weitere spezielle Anwendung der SVDD. Ziel ist hier ebenfalls das Lernen einer kompakten Beschreibung des Umrisses einer gegebenen Datenmenge, wobei speziell der Zerfall in separate Regionen bei hinreichend kleiner Bandbreite (vgl. Abschnitt 5.3.4 auf Seite 70) ausgenutzt wird zur Unterscheidung verschiedener Cluster in den Daten.

## **5.4. Zusammenfassung**

In diesem Kapitel wurden Supportvektor Maschinen als grundlegendes Konzept eingeführt. Im Rahmen vieler praktischer Anwendungen haben sich Supportvektor-Ansätze bereits bewährt und ihre gute Performanz und speziell ihre gute Generalisierungsfähigkeit unter Beweis gestellt. Zudem wurde speziell die 1-Klassen Erweiterung durch die Support Vector Data Description eingeführt. Letztere wird sich im nächsten Teil dieser Arbeit als eine ausschlaggebende Technik für die Constraintmodellierung in dem angestrebten Ansatz zur verteilten Wirkleistungsplanung erweisen. Ihre Fähigkeit, ein Modell für die kontinuierlich räumliche Figur ausgehend von einer Punktmenge zu erlernen wird sich als nützlich erweisen, um Fahrplanräume (mit Fahrplänen in ihrer vektoriellen Darstellung) zu modellieren. Zudem wird ein hierauf aufbauender Ansatz speziell die Möglichkeiten der Navigation bei der gezielten Verschiebung von Punkten im Featureraum ausnutzen, um Fahrpläne so zu verändern, dass gültige Fahrpläne im Rahmen der Lösungsfindung erzeugt werden können.

Dieses Kapitel schließt nunmehr die Vorstellung der Grundlagen und verwandten Arbeiten ab. Teil 2 der Arbeit wird im Anschluss eine Formalisierung des eingangs erläuterten Problems der Wirkleistungsplanung vornehmen und basierend auf den soeben vorgestellten Grundlagen einen Lösungsansatz erarbeiten, der eine Verbindung von Constraint-Behandlung in der Optimierung mittels auf Supportvektor-Ansätzen basierender Modelle vornehmen wird.

**Teil II.**

**Problemformalisierung &  
Lösungsansatz**



## 6. Formalisierung und Modellierung der Wirkleistungsplanung im Smart Grid

*Die Mathematiker sind eine Art Franzosen:  
redet man zu ihnen, so übersetzen sie es in ihre Sprache,  
und dann ist es alsobald ganz etwas anderes.*

---

*(Johann Wolfgang von Goethe)*

### 6.1. Einleitung

Das Problem der Lastplanung im Smart Grid, so wie es im Rahmen dieser Arbeit verstanden werden soll, wurde in den vorangegangenen Abschnitten zumeist eher umgangssprachlich und informell vorgestellt. Bevor in den nachfolgenden Kapiteln ein Lösungsansatz entwickelt und diskutiert wird, soll an dieser Stelle zunächst eine Formalisierung der Problemstellung vorgenommen werden. Hierzu gehört auch die formale Definition diverser benutzter Begriffe und Konstrukte.

In einem ersten Schritt werden daher aus der Problemstellung zunächst die beteiligten Akteure und deren Ziele, welche sich durch eine auf Märkten basierende Wirkleistungsplanung ergeben, abgeleitet und formal beschrieben. Zudem ist ein Verständnis für die Interpretation von Fahrplänen hilfreich, welche hier nicht – wie meist üblich – als Zeitreihen aufgefasst werden sollen, sondern als Elemente eines Vektorraums. Analog muss passend dazu auch der Begriff Constraint, welcher eine Einschränkung bezüglich der Realisierbarkeit einzelner Fahrpläne darstellt, geometrisch interpretiert und gefasst werden. Basierend auf diesen Definitionen kann schließlich das konkrete Optimierungsproblem formalisiert werden, für das im Nachgang ein neuer Lösungsweg zur dynamischen Modellierung und Constraintintegration beschrieben wird.

### 6.2. Beteiligte Akteure

Das Wirkleistungsplanungsproblem, wie es in dieser Arbeit betrachtet wird, umfasst verschiedene Akteure. Diese sowie ihre Eigenschaften und Rollen sollen zunächst formal gefasst werden. Insbesondere für das Zusammenwirken wichtig wird hierbei eine Formalisierung der Handlungsmöglichkeiten einzelner Akteure. Da es primär um die Planung von

Wirkleistungsverläufen im Rahmen eines Verbundes von Anlagen geht, steht als zentraler Akteur die einzelne Anlage oder Einheit im Vordergrund.

**Definition 6.1** (Einheit). *Eine Einheit  $U$  ist eine nicht leere Menge von Anlagen (Erzeuger, Verbraucher oder beides), gegebenenfalls einer einzelnen Anlage, die gemeinschaftlich von einer einzelnen Kontrollinstanz gesteuert werden und über einen gemeinschaftlichen Anschlusspunkt mit dem Stromnetz verbunden sind.*

Im Sinne dieser Definition als Blackbox ist von außen nicht zu unterscheiden, ob es sich bei einer Einheit um eine einzelne Anlage oder um mehrere handelt. Beide Fälle werden gleich behandelt. Der Fall, dass eine Einheit lediglich aus einer Einzelanlage besteht, steht in dieser Arbeit im Vordergrund, sodass o.B.d.A. im Folgenden der Terminus Anlage auch synonym für Einheit verwendet wird. Es ist in der Praxis aber durchaus denkbar eine gemeinschaftlich gesteuerte Menge von Anlagen zu haben, die nach außen als Einheit auftreten: beispielsweise verschiedene Fertigungsanlagen einer Fabrikationsstraße oder mehrere steuerbare Geräte in einem Haushalt für die sich aufgrund der Größe keine separate Kontrolleinheit lohnen würde. Solche Gruppen von Anlagen werden in dieser Arbeit jedoch nicht weiter betrachtet. Ein Beispiel für die Zusammenfassung von Kleinanlagen zu virtuellen Großgeräten findet sich in (Lünsdorf 2012a).

Es seien an dieser Stelle aber dennoch bevorzugt diejenigen, eher kleinen Energiesysteme betrachtet, die ohne ein Bundling mit anderen Anlagen (der eigenen oder anderer Art), bzw. mit ebenfalls kleinen Verbrauchern oder so genannten Prosumern (Einheiten, die sowohl als elektrische Erzeuger als auch als Verbraucher auftreten können; z. B. Speicher) allein keinen nennenswerten Betrag leisten könnten. Jedoch werden alle Anlagen einzeln und nicht als Gruppe modelliert. Ein zentraler Begriff in diesem Zusammenhang ist der Verbund von Einheiten der im Kontext dieser Arbeit folgendermaßen definiert wird:

**Definition 6.2** (Verbund). *Unter einem Verbund soll ein Zusammenschluss von Einheiten zur gemeinschaftlichen Erfüllung einer bestimmten Aufgabe verstanden werden. Ein Verbund  $\mathcal{V}$  ist ein Paar*

$$\mathcal{V} = (\mathcal{U}, \mathcal{A}), \quad (6.1)$$

wobei:

- $\mathcal{U}$ : Eine nicht leere Menge  $\{U_1, \dots, U_n\}$  von Einheiten ist, welche eine gemeinschaftliche Aufgabe haben (hier: die Erfüllung eines Wirkleistungsprodukts).*
- $\mathcal{A}$ : Eine nicht leere Menge von Agenten  $\{A_1, \dots, A_n\}$  ist, wobei Agent  $A_i$  für die Steuerung von Einheit  $U_i$  zuständig ist. Die Agenten vertreten die Einheiten nach außen.*

Ein solcher Verbund spezifiziert streng genommen wieder eine (gemeinsam gesteuerte) Einheit, allerdings ohne gemeinschaftlichen Anschlusspunkt an das Stromnetz. Als zusätzliche Qualität gehört zu einem Verbund jedoch, sich dynamisch zu rekonfigurieren im Hinblick auf ein Wirkleistungsprodukt. Ohne eine solche Rekonfiguration der Verbundzusammensetzung wäre ein Verbund konzeptionell vergleichbar mit einem (statischen) Virtuellen Kraftwerk. Ein Verbund wird die Systemgrenze für das in diesem Kapitel noch zu spezifizierende Optimierungsproblem im Rahmen der Wirkleistungsplanung darstellen.

Die Unterscheidung von physikalischen Einheiten und ihrer informationstechnischen Abstraktion, den Agenten und deren Modellvorstellung von den technischen Einheiten, ist für eine mögliche, spätere technische Umsetzung des Konzeptes in ein Produktivsystem von Bedeutung. Im Rahmen dieser Arbeit stehen jedoch zunächst keine real vorhandenen physikalischen Anlagen zur Verfügung. Daher wird im weiteren Verlauf der Arbeit auch keine konkrete Unterscheidung zwischen einer Einheit und dem Simulationsmodell der Einheit vorgenommen. In allen Evaluationen werden stets simulierte Einheiten verwendet.

### 6.3. Modellierung

Nach den Akteuren sollen nun die Einbettung in den Gesamtablauf des Anwendungsfalls und die verwendeten Modellierungen beschrieben werden. Ein Verständnis des zeitlichen Ablaufs ist wichtig für mögliche Fristen bezüglich der Einsatzplanbestimmung. Im Rahmen der Modellierung ist die Kenntnis der Fahrplandarstellung, insbesondere der Darstellung von Fahrplanmengen und der sie einschränkenden Bedingungen Voraussetzung für die Problemlösung. Für die spätere Bewertung von Lösungen schließlich werden noch Kennzahlen zur Beurteilung von Lösungsalternativen benötigt.

Das hier beschriebene Problem ist in einen sich regelmäßig wiederholenden Gesamtablauf eingebettet. Im Rahmen eines dezentral und autonom agierenden Kontrollmechanismus zum Abgleich von Wirkleistungseinspeisung und -bedarf stimmen sich Erzeuger und Verbraucher automatisiert untereinander ab. Unter dem Dach eines speziell entwickelten Marktmechanismus finden sich Anlagen z. B. zu Erzeugergemeinschaften (Verbünde) zusammen oder stellen im Rahmen von Einkaufsgemeinschaften Gesuche an den Markt. Aber auch gemischte Verbünde zum Zweck der Produktveredlung (Verbraucher kappen z. B. unerwünschte Spitzen in einem Erzeugerverbund) sind denkbar. Aber auch die Bereitstellung netzstützender Maßnahmen gehört zu einem solchen dezentralen Koordinationssystem (Nieße et al. 2012b).

Im Rahmen der vorausschauenden Wirkleistungsbereitstellung wird regelmäßig ein Verbundbildungsprozess angestoßen damit sich geeignete Verbünde bilden, die auf Wirkleistungsprodukte (für einen in der Zukunft liegenden Lieferzeitraum) am Markt bieten (Beer 2013). Nach dem ein Verbund einen (ggf. teilweisen) Zuschlag bekommen hat, muss durch den Verbund vor der Produktlieferung eine Optimierung des Einsatzplanes durchgeführt werden, da aus Komplexitätsgründen die ursprüngliche Verbundbildung nur eine Teilmenge der Möglichkeiten der Anlagen einbeziehen kann. Zudem haben sich Voraussetzungen und Annahmen unter der die Verbundbildung stattfand bis zum Lieferzeitpunkt mit großer Wahrscheinlichkeit schon verändert. Analog machen Ereignisse während der Produktlieferung Neuplanungen erforderlich (Nieße & Sonnenschein 2013). Diese Situation muss modelliert werden, um die Wirkleistungsplanung adäquat zu beschreiben.

Für die Einbettung in den Gesamtablauf soll zunächst folgende Annahme getroffen werden: ein Verbund von Anlagen dessen Einsatzplan bestimmt werden soll existiert bereits. Für das in dieser Arbeit betrachtete Problem ist es dabei unerheblich, ob es sich um einen statischen Verbund von Anlagen, also um ein Virtuelles Kraftwerk, handelt oder um einen dynamisch und selbstorganisiert gebildeten Verbund. Im Folgenden soll jedoch von letzterem Fall ausgegangen werden.

### 6.3.1. Fahrplanmodellierung und Wirkleistungsprodukt

Eine wichtige Voraussetzung für die Lösung des Wirkleistungsplanungsproblems im Verbund wird die Steuerbarkeit der beteiligten Anlagen sein. In der Praxis würde die eigentliche Steuerung durch einen integrierten Mikrocontroller erfolgen, der geeignete Steuersignale für die Anlage selbst erzeugt. Dieser Teil wird in dieser Arbeit nicht betrachtet. Es ist davon auszugehen, dass mit derzeitigen Mitteln eine Steuerung entwickelt werden kann, die Sollvorgaben bzgl. elektrischer Leistung in Steuersignale für die Anlage umsetzen kann. Viele Hersteller von KWK-Anlagen bieten derartige System bereits heute für ihre Anlagen an. Als Beispiel sei hier der *RemoteManager* von der Firma SOKRATHERM (<http://www.sokratherm.de/>) genannt. Entscheidendes Ergebnis der hier betrachteten Planung wird jedoch die Sollvorgabe sein, mit welcher eine Anlage zu fahren ist. Eine steuerbare Einheit benötigt üblicherweise eine Vorgabe, wie sie für einen gegebenen Zeitraum operieren soll. Die Steuerung der Anlage benötigt für die Erzeugung konkreter Steuersignale eine Vorgabe über den zeitlichen Verlauf der zu erzielenden Wirkleistung. Die Vorgabe wird über einen Fahrplan realisiert, der wie folgt definiert ist:

**Definition 6.3** (Fahrplan). *Ein Fahrplan für eine einzelne Einheit legt eine mittlere Wirkleistungsvorgabe je Zeitintervall fest:*

$$\mathbf{p} = (p_0, p_1, \dots, p_d) \in \mathbb{R}^d, \quad (6.2)$$

wobei  $p_i$  die im Zeitintervall  $i$  zu erzielende mittlere Wirkleistung bestimmt.

In der Praxis wären noch die Definition des Startzeitpunktes (Beginn von Zeitintervall 0) zu einem Fahrplan sowie der Länge eines Zeitintervalls notwendig. Liegt für jede Anlage eines Verbundes ein Fahrplan vor, so ergibt sich hieraus der Einsatzplan für einen Verbund.

**Definition 6.4** (Einsatzplan). *Ein Einsatzplan für einen Verbund  $\mathcal{V}$  stellt eine Menge von Fahrplänen dar:*

$$\mathcal{E} = \{\mathbf{p}_i\}_{\forall U \in \mathcal{V}} \quad (6.3)$$

Von Bedeutung für die spätere Modellierung wird auch der Begriff des skalierten Fahrplans sein, der wie folgt definiert wird:

**Definition 6.5** (Skalierter Fahrplan). *In einem skalierten Fahrplan liegen alle Leistungswerte zwischen 0 und 1:*

$$\mathbf{p}_{[0,1]} = (\pi_0, \pi_1, \dots, \pi_d) \in [0, 1]^d, \quad (6.4)$$

mit  $\pi_i = \frac{p_i - p_{min}}{p_{max} - p_{min}}$

wobei  $\pi_i$  die im Zeitintervall  $i$  zu erzielende mittlere prozentuale Wirkleistung angibt bezogen auf die maximale ( $p_{max}$ ) bzw. minimale ( $p_{min}$ ) elektrische Leistung der Einheit.

Das Skalieren der Fahrpläne erlaubt zum einen die Gleichbehandlung von Erzeugern, Verbrauchern und auch von Prosumern, zum anderen ist bekannt, dass ein solches Featu-

rescaling zu besseren Ergebnissen bei der Anwendung bestimmter Klassifikationsverfahren (wie sie später Verwendung finden sollen) führt (Juszczak et al. 2002).

Ein Fahrplan (auch innerhalb eines Einsatzplanes) definiert jedoch nur die Vorgabe eines Wirkleistungsverlaufs, welcher durch die Einheit erzielt werden soll. Für die tatsächliche Umsetzung müsste eine solche Vorgabe von einem Anlagencontroller jedoch noch in konkrete Steuerbefehle (z. B. zur Drehzahlkontrolle bei einem BHKW) umgesetzt werden. Gegebenenfalls ist eine 1:1 Umsetzung nicht immer möglich. Zudem können Unregelmäßigkeiten der Anlagenregelung für kleine Abweichungen sorgen. Ob ein Fahrplan von der Anlage wie vorgesehen realisiert (abgefahren) werden kann oder nicht, hängt davon ab, ob alle Annahmen bezüglich des Betriebszustandes zum Startzeitpunkt und ggf. Prognosen so eintreffen, wie sie bei der Erstellung des Fahrplans vorausgesetzt wurden; bzw. davon, ob Abweichungen hinreichend klein sind. Möglicherweise muss die Anlage aufgrund aktueller Situationen von der Vorgabe abweichen. Daher soll hier zusätzlich zum Fahrplan noch der Wirkleistungsverlauf definiert werden, der tatsächlich von der Einheit umgesetzt wird.

**Definition 6.6** (Lastgang / Wirkleistungsverlauf). *Der realisierte Wirkleistungsverlauf einer Einheit wird analog zum Fahrplan definiert als:*

$$\mathbf{p}^{(r)} = (p_0, p_1, \dots, p_d) \in \mathbb{R}^d, \quad (6.5)$$

wobei  $p_i$  die im Zeitintervall  $i$  tatsächlich erzielte mittlere Wirkleistung angibt.

Es ist zu erwarten, dass die heutigen Subventionen aus dem EEG auslaufen werden und Anlagen zukünftig marktorientiert produzieren müssen. Gleiches gilt für die Beschaffung von elektrischer Energie. Es ist davon auszugehen, dass durch geeignete, neue Marktstrukturen Elektrizität nicht über langfristige Lieferverträge, sondern über Produkte mit überschaubaren Produktzeiträumen gehandelt wird. Derartige Marktstrukturen werden derzeit beispielsweise bei (Wissing & Appelrath 2013; Shah 2012b; Nieße & Sonnenschein 2013) untersucht. In dieser Arbeit sollen keine speziellen Produkte oder Märkte betrachtet werden. Es wird aber angenommen, dass ein Verbund und ein ihm zugeordnetes Wirkleistungsprodukt bereits existieren, wobei durch das Produkt im Wesentlichen der vom Verbund erwartete Gesamtlastgang sowie ein zeitlicher Horizont (der Lieferzeitraum) definiert wird, anhand dessen der Verbund seine Planung vornehmen kann.

**Definition 6.7** (Wirkleistungsprodukt). *Ein Wirkleistungsprodukt  $P$  ist ein Tupel*

$$P = (\mathcal{V}, \mathbf{p}, t_0), \quad (6.6)$$

wobei:

$\mathcal{V}$ : Verbund für den das Produkt gilt.

$\mathbf{p}$ : Lastgang der den vom Produkt geforderten Wirkleistungsverlauf kennzeichnet.

$t_0$ : Definiert den Startzeitpunkt des Erbringungszeitraums, also den Startzeitpunkt des Lastgangs  $\mathbf{p}$ .

### 6.3.2. Anlagenmodellierung und Randbedingungen

Eine wichtige Anforderung wird es sein, die Möglichkeiten einer Anlage für einen gegebenen zukünftigen Zeitraum zu kennen. Anlagenmodelle haben üblicherweise den Zweck, das Verhalten einer Anlage auf physikalisch-technischer Ebene nachzubilden, um in Simulationsstudien nicht vorhandene reale Anlagen zu substituieren. Der Fokus liegt zumeist darauf, zu einer gegebenen Eingabe, die Reaktion der Anlage zu ermitteln. In dieser Arbeit ist die Fragestellung eine andere.

Jede Anlage die steuerbar ist, kann ihre eigentliche Aufgabe meist auf verschiedene Weise erfüllen. Als Beispiel mag ein Mini-BHKW dienen. Eigentliche Aufgabe ist es, für jeden Zeitpunkt stets genug Wärme zu liefern um die vorgegeben Raumtemperatur zu halten. Zeitgleich wird auch Elektrizität erzeugt. Wenn aber die Wärmeproduktion vom Wärmeverbrauch durch einen Pufferspeicher in gewissen Schranken zeitlich entkoppelt werden kann. So kann das Wärmeprofil auf verschiedene Weise erzeugt werden, wodurch sich auch verschiedene elektrische Erzeugungsprofile ergeben. Es existieren also verschiedene elektrische Fahrpläne mit denen eine Anlage betrieben werden kann, sodass unter Beachtung weiterer Randbedingungen dennoch die eigentliche Aufgabe (im Beispiel die Wärmeversorgung) erfüllt wird. Der Fokus liegt hier auf der Ermittlung der Menge der Möglichkeiten (verschiedene realisierbare Fahrpläne), die eine Anlage ausgehend von einem gegebenen Startzustand hat. Daher ist zunächst der Betriebszustand einer Anlage zu definieren.

**Definition 6.8** (Betriebszustand). *Ein physikalisch-technisches System kann beschrieben werden durch eine Menge von variablen Zustandsgrößen, ihren Veränderungen und ggf. konstanten Beschreibungsparametern. Der Momentanzustand des Systems wird dann beschrieben durch die momentane Belegung der Zustandsgrößen mit Zahlenwerten. Gegeben sei eine Einheit  $U$ . Sei  $\{\zeta_1, \dots, \zeta_n\}$  eine Menge von Zustandsgrößen. Eine Belegung*

$$\mathcal{B}_U[t] : \{\zeta_1, \zeta_2, \dots, \zeta_n\} \rightarrow \mathcal{W}_{\zeta_1} \times \mathcal{W}_{\zeta_2} \times \dots \times \mathcal{W}_{\zeta_n} \quad (6.7)$$

*ordne den Zustandsgrößen eine Belegung zum Zeitpunkt  $t$  aus der Menge  $\mathcal{W}_{\zeta_1}$  der zulässigen Belegungen für die Zustandsgröße  $\zeta_i$  zu. Dann beschreibt  $z_t = \mathcal{B}_U[t](\{\zeta_1, \dots, \zeta_n\})$  den Zustand der Einheit zum Zeitpunkt  $t$ . Eine Menge diskreter, zeitlich aufeinanderfolgender Betriebszustände sei bezeichnet mit  $Z_U$  mit Startzustand  $z_0$ .*

Im Rahmen der Vereinfachung bei der Modellierung einer Einheit werden in der Regel voneinander abhängige Größen eliminiert, da sie keine zusätzliche Information zur Beschreibung beisteuern. Ein Beispiel für einen Betriebszustand ist der aktuelle (prozentuale) Füllstand (state of charge, SOC) eines Batteriespeichers.

Verschiedene Betriebsmodi einer Anlage werden meist durch eine ganze Reihe von Randbedingungen eingeschränkt. Solche Constraints beim Betrieb einer Anlage bestimmen maßgeblich die Möglichkeiten bezüglich der Menge von verschiedenen Fahrplänen. Constraints werden meist modelliert als Satz von Regeln, die auf Entscheidungsvariablen anzuwenden sind (Eiben & Ruttkay 1997). Die Modellierung solcher Regelsysteme (auf die dann Constraint-Solver angesetzt werden können) erfordern Erfahrung und Experten-

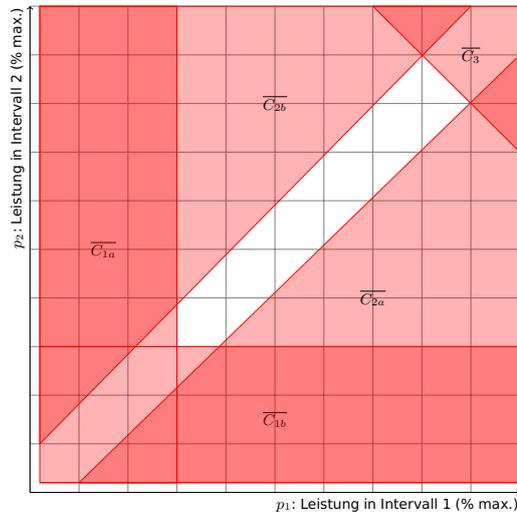


Abbildung 6.1.: Vereinfachendes, 2-dimensionales Beispiel für den Suchraum eines modulierenden BHKW mit drei exemplarischen Constraints.

wissen (Frisch et al. 2005). Zudem ist eine solche Modellierung schwer automatisierbar, insbesondere in sich ändernden Szenarien. Hier soll daher ein anderer Weg verfolgt werden, indem nur die Auswirkungen der Constraints auf den Suchraum modelliert werden.

An dieser Stelle soll zunächst eine geometrische Beschreibung der verschiedenen Constraints für den Betrieb eingeführt und erläutert werden. Ein Wirkleistungslastgang (und somit auch ein Fahrplan, der einen Wirkleistungslastgang vorgibt) wird aufgefasst als Punkt in  $\mathbb{R}^d$ . Ohne Beschränkung der Allgemeinheit soll im Folgenden davon ausgegangen werden, dass die Wirkleistung auf das Intervall  $[0, 1]$  skaliert ist.

Abbildung 6.1 zeigt das Prinzip an einem vereinfachten 2-dimensionalen Beispiel. Mit lediglich dem Box-Constraint, dass alle Leistungswerte immer zwischen 0 und 1 liegen müssen als einziger Einschränkung wäre der Handlungsspielraum einer Anlage der gesamte Hyperwürfel  $[0, 1]^d$  für  $d$  Zeitintervalle. In Abbildung 6.1 wäre dies die gesamte dargestellte Rechteckfläche. Die verschiedenen Anlagenconstraints schränken aber diesen Bereich sukzessive immer weiter ein. Dies gilt prinzipiell für beliebige Anlagen und Anlagentypen. Beispielhaft illustriert ist dieser Sachverhalt in der Darstellung 6.1 an den folgenden Constraints für ein modulierendes BHKW:

**$C_1$  Leistungsbegrenzungen:** Üblicherweise können modulierende BHKW ihre Leistung innerhalb eines Fensters zwischen einer minimalen Leistung  $p_{min}$  (einzuhaltende Mindestmotordrehzahl) und der Maximalleistung  $p_{max}$  der Anlage variieren. Das vollständige Abschalten der Anlage ( $p = 0$ ) ist eine weitere Option. Ein Constraint  $C_1 : p_{min} \leq p_i \leq p_{max} \vee p_i = 0, \forall 0 \leq i \leq d$  verbietet daher Lösungen  $p_1 \in ]0, p_{min}[$  ( $C_{1a}$ ) oder  $p_2 \in ]0, p_{min}[$  ( $C_{1b}$ ).

**$C_2$  Trägheit:** Eine Anlage kann ihren Betriebszustand nicht instantan ändern. Gegebenenfalls sprechen auch nicht physikalische Gründe gegen eine zu schnelle Änderung der Ausgangsleistung (z. B. Lärmemissionen). Das Constraint  $C_2$  verbietet somit Lösungen  $|p_i - p_{i+1}| > \epsilon$ , und somit eine Änderung zwischen zwei aufeinander folgenden

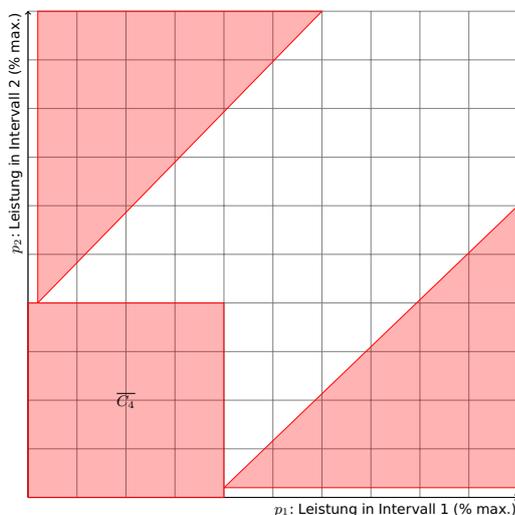


Abbildung 6.2.: Vereinfachtes, 2-dimensionales Beispiel für einen nicht konvexen Lösungsraum für ein BHKW (verbleibende weiße Fläche).

Zeitintervallen, die größer als eine zulässige Schranke  $\epsilon$  ist. Dieses Constraint stellt auch ein Beispiel dar, welches in einer nicht zusammenhängenden Ausschlussfläche ( $\overline{C_{2a}}$  bzw.  $\overline{C_{2b}}$ ) resultiert.

- $C_3$  **Speicherfüllstand:** Jedes BHKW muss die gleichzeitig produzierte Wärme jederzeit verbrauchen können oder diese in einem Pufferspeicher zwischenspeichern. Ein teilweise gefüllter Pufferspeicher ohne gleichzeitige Wärmeentnahme sowie seine Speicherkapazität würden die Summe  $p_1 + p_2$  nach oben beschränken und somit die Fläche  $\overline{C_3}$  ebenfalls aus dem Lösungsraum ausschließen.

Lediglich der nach Abzug aller durch Constraints verbotenen Regionen verbleibende Bereich (der weiße Bereich in den Abbildungen) ist der Lösungsraum aus dem bei der Durchführung eines Planungsalgorithmus Lösungskandidaten verwendet werden dürfen. Üblicherweise sind Constraints in realen Anwendungsfällen aus der Praxis nicht notwendigerweise so linear wie in den vorgestellten Beispielen.

Eine weitere Frage ergibt sich noch nach der möglichen Struktur der entstehenden Lösungsräume, da hiervon maßgeblich die Anwendbarkeit diverser Suchalgorithmen sowie deren Laufzeit abhängen. Wie bereits in Abbildung 6.1 zu erkennen, ist der Lösungsraum nicht notwendigerweise eine monolithische, zusammenhängende Region im Suchraum. In diesem Beispiel zerfällt der Lösungsraum in die (überhöht dargestellt) Nulllinie (ausgeschaltetes BHKW) und die Region für den restlichen Betrieb. Die Frage, ob von konvexen Strukturen ausgegangen werden kann, lässt sich mit einem weiteren Beispiel beantworten. In Abbildung 6.2 ist ein weiteres Constraint  $C_4$  zu dem vorigen Beispiel hinzugefügt worden. Dieses gibt vor, die Leistung in mindestens einem Zeitintervall nicht unter einen Minimalwert fallen zu lassen (beispielsweise um eine minimale mittlere Leistung nicht zu unterschreiten). Offensichtlich ergibt sich hieraus bereits ein nicht konvexes Polytop als Lösungsraum.

Maßgeblich für die Struktur des Lösungsraums sind die Constraints welche ihn beschränken. Sie sind wie folgt definiert:

**Definition 6.9** (Constraint). Constraints werden üblicherweise als Satz von Regeln modelliert, die auf Entscheidungsvariablen anzuwenden sind (Eiben & Ruttkay 1997). In der Definition nach Eiben & Ruttkay (1997) ist ein Constraint Optimization Problem (COP) ein Triple

$$(S, f, \phi) \quad (6.8)$$

mit

$S$ : der freie Suchraum  $S = \mathcal{D}_1 \times \mathcal{D}_2 \times \dots \times \mathcal{D}_n$  über die Definitionsmengen  $\mathcal{D}_i$  der jeweiligen Parameter,

$f$ : Optimalität definierende Zielfunktion,

$\phi$ : Boole'sche Funktion für  $m$  Constraints  $\phi = \phi_{c_1} \wedge \phi_{c_2} \wedge \dots \wedge \phi_{c_m}$ , welche die Gültigkeit einer Lösung repräsentiert.

Hier sollen Constraints als Regionen im Vektorraum der möglichen Lösungen definiert werden. Ein Constraint  $C \in \mathcal{C}$  beschreibt eine Region (einen Unterraum) im freien Suchraum  $S$ .

Sei  $\mathcal{R}(C) \subseteq S$  eine Region, die von einem Constraint  $C$  induziert wird,  $s \in S$  ein Lösungskandidat, dann gilt

$$\phi_C(s) = \text{true} \Leftrightarrow s \in \mathcal{R}(C). \quad (6.9)$$

Die Menge aller gültigen Lösungen  $\{s \in S \mid \phi(s) = \text{true}\}$  (Eiben & Ruttkay 1997) sei bezeichnet als  $\mathcal{L}$ .

Der freie Suchraum ist die Menge aller möglichen Fahrpläne, ungeachtet ob ein Fahrplan aufgrund technischer Restriktionen realisierbar ist oder nicht. Er ist definiert wie folgt.

**Definition 6.10** (Suchraum). Mit der mittleren Wirkleistung  $p_i$ , welche eine Anlage  $U$  im entsprechenden Zeitintervall  $i$  erbringen kann, ergibt sich als Definitionsbereich  $\mathcal{D}_{1 \leq i \leq d} = [p_{\min}, p_{\max}]$  und somit für den freien Suchraum  $S$  für Einheit  $U$  der Hyperwürfel

$$S = [p_{\min}, p_{\max}]^d. \quad (6.10)$$

Im skalierten Fall bildet der Einheitshyperwürfel  $[0, 1]^d$  den freien Suchraum.

Im Falle skalierten Fahrpläne ist der freie Suchraum offensichtlich für alle Anlagen gleich. Die Anlagen unterscheiden sich jedoch nach ihren Lösungsräumen. Der Lösungsraum beschreibt die Menge der tatsächlich realisierbaren Fahrpläne.

**Definition 6.11** (Lösungsraum). Der Lösungsraum einer Einheit ist definiert als Schnittmenge aller Constraint-induzierten Regionen.

$$\mathcal{F} = \bigcap_{C \in \mathcal{C}} \mathcal{R}(C). \quad (6.11)$$

Der Lösungsraum im skalierten Fall sei bezeichnet mit  $\mathcal{F}_{[0,1]}$ .

Ein wichtiger Aspekt ist die Integration der zeitlichen Abhängigkeit bei dieser Art der Constraint- bzw. Suchraummodellierung. Betrachtet werden immer vollständige Fahrpläne. Entscheidungen zu einem frühen Zeitpunkt (für die Wahl einer Leistungsstufe) können bei bestimmten Anlagentypen Auswirkungen auf die späteren Möglichkeiten haben: ein früh gefüllter Pufferspeicher schränkt ohne hinreichenden Verbrauch die späteren Möglichkeiten ein und umgekehrt. Fahrpläne derartiger Anlagen müssen also immer als Ganzes betrachtet werden. Änderungen der Leistung an einzelnen Stellen ziehen zumeist auch notwendige Nachfolgeänderungen an anderen Stellen aufgrund der Abhängigkeit durch die verursachten Zustandsänderungen (z. B. Speicherfüllstand) nach sich. Gleiches gilt für den Suchraum der Fahrpläne als Ganzes. Die zeitliche Abhängigkeit der Leistungswerte im Fahrplan untereinander wird bei der hier vorgeschlagenen Art der Modellierung repräsentiert durch die geometrische Lage eines Fahrplans im Suchraum.

Üblicherweise existieren Constraints, welche abhängig von einem bestimmten Startzustand sind. Als Beispiel mag der Füllstand eines Pufferspeichers dienen. Hieraus lässt sich folgern, dass ein Lösungsraum nicht nur anlagenspezifisch ist, sondern sich auch für eine Anlage von Betriebszustand zu Betriebszustand unterscheidet. Ein Lösungsraum kann daher im Allgemeinen nicht statisch vorab berechnet werden.

### 6.4. Kennzahlen zur Lösungsbewertung

In den hier betrachteten Szenarien spielen in der Praxis zumeist mehrere Interessen eine Rolle. Oberstes Ziel ist zwar das gemeinschaftliche Erreichen eines gemeinsamen Ziellastgangs zur Produkterfüllung, jedoch gibt es weitere Ziele für ein Optimierungspotenzial. So ist es beispielsweise vorteilhaft, wenn Anlagen, die sich Betriebsmittel des Stromnetzes als gemeinsame Ressource teilen, sich hinsichtlich der Nutzung abstimmen. Ein Beispiel hierzu ist die Minimierung des Gleichzeitigkeitsfaktors.

Aber auch individuelle Interessen einzelner Anlagen (Minimierung lokaler Kosten, Maximierung lokalen Komforts, usw.) müssen vertreten werden. Zusätzliche Optimierungsziele, welche nur von den Fahrplänen (bzw. den Leistungswerten) abhängen oder lediglich anlagenspezifisch sind, können einfach integriert werden. Wenn eine Bewertung jedoch fahrplanspezifisch vorgenommen werden muss, so werden Kennzahlen benötigt, welche für jeden Fahrplan eine individuelle Bewertung bezüglich bestimmter Optimierungsziele erlauben.

Unabhängig davon, welche Kennzahlen in einem konkreten Anwendungsfall als Annotation eines Lösungskandidaten verwendet werden, dürfte in der Regel eine lokale Bestimmung dieser Werte notwendig werden. Motiviert werden soll auch diese Aussage am Beispiel Smart Grid.

Wenn eine dezentrale Anlage, z. B. ein BHKW, verschiedene Lastgänge als Lösungsalternativen anbieten kann, so sind diese individuellen Lastgänge immer auch mit individuellen Kosten verbunden. Dies reicht von Unterschieden durch individuell verschiedenen Kraftstoffverbrauch bis hin zu Benutzerpräferenzen (wenn beispielsweise gewisse Lastgänge aufgrund von Lärm unerwünscht, aber nicht völlig verboten sind). Solche Soft Constraints

lassen sich so ebenfalls gut über Kennzahlen modellieren, um anschließend in der Zielfunktion der Optimierung angemessen berücksichtigt zu werden.

Derartige Kennzahlen hängen oft nicht nur vom zu erbringenden Lastgang allein ab, sondern auch von weiteren Rahmenbedingungen wie etwa dem jeweils aktuellen Zustand der Anlage. Wird etwa ein Lastgang aus dem Kaltstart heraus gefahren anstatt von einer Anlage, die bereits auf Betriebstemperatur ist, so hat dies auch Auswirkungen auf einige Kennzahlen (z. B. den Kraftstoffeinsatz).

Die Werte solcher Kennzahlen müssen spezifisch für einzelne Fahrpläne bestimmt werden, da es später möglich sein soll, jeden Fahrplan hinsichtlich verschiedener Ziele bewerten zu können. Basierend auf lokalen Parametern muss die Berechnung vor jedem Planungsintervall neu vorgenommen werden. Für viele Kennzahlen und Modelle wäre es sicher denkbar, eine mathematische Funktion zur direkten Bestimmung der Werte aufzustellen (beispielsweise für die reinen Brennstoffkosten in Abhängigkeit von der erzeugten elektrischen Energie), jedoch würden viele Betreiber solche Interna nur ungern außer Haus geben. Zum gegenwärtigen Zeitpunkt soll lediglich vorausgesetzt werden, dass es einen beschreibenden Zusammenhang gibt, nicht dass dieser auch in seiner mathematischen Formulierung bekannt ist.

Ein zusätzliches Problemfeld ergibt sich möglicherweise durch die kontinuierliche und fluktuierende Reorganisation der Verbände. Die kurzfristig betrachtete Betriebsweise einer Einheit für die nähere Zukunft hat höchstwahrscheinlich auch Einfluss auf die Möglichkeiten (Freiheitsgrade der Einheit) danach und somit auch auf erreichbare Minimierungen negativer Auswirkungen über längerfristige Zeiträume. Anders gesagt, sollten Kennzahlen für solche Einheiten immer einheitenspezifisch berechnet werden und beispielsweise auch Verluste berücksichtigen, die erst später in den nachfolgenden Verbänden zum tragen kommen. Andernfalls würde z. B. das besonders starke Herabkühlen eines Gefrierhauses für den aktuellen Verbund nur teilweise berücksichtigt, da statische Verluste erst für den nächsten Verbund zum Tragen kommen. Dies ist im Übrigen speziell eine Eigenschaft der Nachhaltigkeitskennzahlen (Bremer & Sonnenschein 2013d). Eine ebenfalls notwendige Betrachtung der Kompatibilität zwischen solchen lokal bestimmten Kennzahlen wird später in Abschnitt 16.3 auf Seite 228 vorgenommen.

Zunächst sei der Begriff der Kenngröße in Anlehnung beispielsweise an (Kütz 2008; Jamous & Müller 2013) definiert als:

**Definition 6.12** (Kenngröße). *Eine Kenngröße ist das Produkt aus Kennzahl und Einheit. Die Kennzahl ist eine Maßzahl welche einen messbaren Sachverhalt zu einem gegebenen Zeitpunkt quantifiziert.*

Kennzahlen konzentrieren Sachverhalte quantitativ (Kütz 2008), können aber ohne Kenntnis der Einheit nicht interpretiert werden und sind daher allein nicht aussagekräftig. Dennoch hat sich der Begriff Kennzahl im Sprachgebrauch als Synonym für Kenngröße eingebürgert. Für den Rest der Arbeit wird auch der Begriff Kennzahl verwendet, wobei implizit angenommen wird, dass bei der Anwendung die Einheit als Interpretationsvoraussetzung bekannt ist. Um die Gefahr von Fehlinterpretationen zu vermindern werden Kennzahlen in der Praxis häufig zu Kennzahlensystemen zusammengefasst (Kütz 2008).

Im Folgenden wird zunächst eine Kennzahl zur Charakterisierung eines Fahrplanlösungskandidaten allgemein wie folgt definiert:

**Definition 6.13** (Kennzahlen für die Lastplanung). Seien  $x \in \mathbb{R}^d$  ein Fahrplan und  $x_{[0,1]} \in \mathcal{F}_{[0,1]}^{(U)}$  seine skalierte Variante aus dem Suchraum einer Anlage  $U$ . Sei ferner  $Q$  eine Kenngröße zur Beschreibung von Lastgängen. Dann ist

$$\begin{aligned} \mathcal{I}^{(Q)}(x, z_0) : \mathbb{R}^d \times Z_U &\rightarrow \mathbb{R}, \\ x &\mapsto \mathcal{I}^{(Q)}(x, z_0) \end{aligned} \tag{6.12}$$

eine Indikatorabbildung welche dem Kandidaten  $x$  eindeutig einen Wert für die Kenngröße  $Q$  zuordnet für den Fall, dass  $x$  aus Anlagenzustand  $z_0$  heraus gefahren wird. Die Funktion

$$\begin{aligned} \mathcal{I}_{[0,1]}^{(Q)}(x, z_0) : \mathbb{R}^d \times Z_U &\rightarrow [0, 1], \\ x &\mapsto \frac{\mathcal{I}^{(Q)}(x, z_0)}{\max_{x \in \mathbb{R}^d, z \in Z_U} \{\mathcal{I}^{(Q)}(x, z)\}}, \end{aligned} \tag{6.13}$$

bildet analog auf die skalierte Kennzahl  $Q_{[0,1]}$  ab.

Die Funktionen  $\mathcal{I}$  sind nicht notwendigerweise als explizite Form bekannt. Es sei lediglich gefordert, dass ein funktionaler Zusammenhang zwischen Lösungskandidat und Kennzahl besteht und dass für alle Elemente der Trainingsmenge die zugehörige Kennzahl (gegebenenfalls auch durch Simulation o. ä. ) bestimmt werden kann.

Abhängig vom konkret gegebenen Optimierungsproblem bei der Lastplanung ist eine Vielzahl verschiedener Kenngrößen denkbar. Beispiele hierfür sind die Kosten (z. B. Brennstoffkosten) die beim Abfahren eines bestimmten Fahrplans entstehen oder Umweltkenngrößen welche den durch unterschiedliche Betriebsmodi variierenden Schadstoffausstoß reflektieren. Aber auch Kenngrößen zur Beurteilung der Zuverlässigkeit mit der ein durch die Optimierung erstellter Einsatzplan von den Einheiten umgesetzt werden kann, ist eine wichtige Größe zur Beurteilung einer Lösung. Schließlich können auch Kenngrößen betrachtet werden zur Einflussnahme auf den gewünschten Anlagenzustand. Ein Beispiel hierzu wäre die Einflussnahme auf die Größe der verbleibenden Flexibilität der Anlage nach Abarbeiten eines Fahrplans, um eine möglichst große Bandbreite an Alternativen für nachfolgende Planungsintervalle zu reservieren. Eine konkrete Auswahl an geeigneten Kenngrößen für die Evaluation wird in Abschnitt [10.3 auf Seite 135](#) getroffen.

## 6.5. Problemdefinition

Die hier betrachteten Szenarien beinhalten explizit die Annahme, dass insbesondere kleinere dezentrale Erzeugungseinheiten sich in der Nach-EEG-Zeit zusammenschließen müssen, um sich am Markt zu behaupten. Erzeugter Strom wird selbst vermarktet (oder selbst verbraucht) werden müssen. Abzulesen ist dieser Trend bereits in der Novellierung des EEG von 2012 (Deutsche Bundesregierung (Bundeskabinett) [2012](#)), in der mit

Teil 3a eine Direktvermarktungsmöglichkeit geschaffen wurde. Hierdurch soll eine mehr marktorientierte Integration von erneuerbaren Energieträgern induziert werden.

Bereits heute stellen Markteintrittsbarrieren eine Hürde an die Mindestleistung dar, die einen Zusammenschluss kleinerer Anlagen sinnvoll erscheinen lassen (Beer 2013). Anforderungen an die Zuverlässigkeit (bei der Erbringung der zugesagten Leistung), können insbesondere Erzeuger aus fluktuierenden Quellen wie Wind oder Sonne allein schwerlich erfüllen. Es ist zu erwarten, dass sich Anlagen aus verschiedenen Besitzverhältnissen mehr oder weniger ad hoc (je nach Situation und Bedarf) zu immer wieder neuen und individuell spezialisierten Verbänden zusammenfinden müssen, um eine Chance zu haben, das eigene Potenzial weitestgehend zu vermarkten und damit auch auszunutzen.

Wie auch in dem Eckpunktepapier (Bundesnetzagentur 2012) bekräftigt, werden dezentrale Einheiten zukünftig ihre Energie selbst vermarkten. Dies wird auch für Dienste wie Regelleistungsbereitstellung oder ähnliches gelten. Zur umfänglichen Ausschöpfung des Potenzials aller Anlagen wird derzeit an verteilten dynamischen Kontrollstrukturen gearbeitet (Nieße et al. 2012b). Wichtigstes Konzept in diesem Szenario ist die dynamische, produktorientierte Neuorganisation der Anlagen. In regelmäßigen Abständen (bei den heute nicht unüblichen Day-Ahead-Marktstrukturen zunächst täglich) sollen sich Anlagen selbstorganisiert so zusammenfinden, dass sich die aktuell existenten Flexibilitäten in einem Verbund sinnvoll ergänzen, um ein gefordertes Produkt abzubilden. Nach dieser Verbundbildung folgt für diejenigen Verbände, die den Zuschlag für ein Produkt bekommen haben ab einem späteren Zeitpunkt die Produktlieferphase in der der vom Produkt definierte Lastverlauf kollektiv erbracht werden muss. Hierfür benötigt jede Anlage in dem Verbund einen Fahrplan, der den geforderten Gesamtlastgang auf die einzelnen Anlagen aufteilt. Eine erste Aufteilung, wie sie bereits während der Verbundbindung vorgenommen wird, kann vermutlich direkt vor Beginn der Produktlieferung noch durch eine initiale Optimierung (jetzt insbesondere unter Beachtung mehrerer Optimierungsziele) verbessert werden. Durchzuführen ist sie in jedem Fall, da sich die während der Verbundbildung gemachten Annahmen geändert haben können. Während der Verbundbildung muss eine kontinuierliche Umplanung jederzeit durchgeführt werden können, falls sich Abweichungen vom vorgesehenen Betrieb zur Produktlieferung ergeben, wobei nunmehr die größtmögliche Beibehaltung der Netzähnlichkeit zum vorherigen Plan als Optimierungsziel integriert werden muss (Nieße & Sonnenschein 2013).

Dieser komplexe Prozess der dynamischen Wirkleistungsplanung erfordert das nahtlose Ineinandergreifen einer Vielzahl von Konzepten und algorithmischen Ansätzen, die in verschiedenen, externen Arbeiten behandelt werden (Beer 2013; Nieße & Sonnenschein 2013; Hinrichs et al. 2013a).

Die regelmäßige Neukonfiguration der Verbände in wechselnden Zusammensetzungen erfordert aber auch, dass die Akteure in der Lage sind sich in einen verbundinternen Steuerungsmechanismus zu integrieren mit dessen Hilfe die Erzeugung und der Verbrauch einzelner Anlagen so koordiniert werden, dass der Verbund als ganzes ein über den Markt abgestimmtes Wirkleistungsprodukt abfahren kann.

Aus Sicht des Gesamtsystems läuft also ein sich wiederholender Prozess aus selbstorganisierter, produktspezifischer Verbundbildung, Markt-basierter Platzierung des Produktes und einer anschließenden Phase der Produktlieferung ab. In dieser Prozesskette tritt an

mehreren Stellen die Notwendigkeit einer Wirkleistungsplanung auf: Direkt vor dem Beginn der Produktlieferung und im Rahmen möglicher Umplanungen für den Fall des Eintretens unvorhergesehener Ereignisse. Für den Prozess der Verbundbildung sind aber ebenfalls bereits Kenntnisse über den Suchraum der Fahrpläne erforderlich, um zu entscheiden für welche Produkte eine Beteiligung sinnvoll erscheint. Dieses Wirkleistungsplanungsproblem soll im Folgenden formalisiert werden, um hierfür einen neuen Mechanismus zur dynamischen Integration der individuellen Anlagenflexibilitäten zu entwerfen.

Gegeben sind ein Wirkleistungsprodukt  $P$  und ein Verbund  $\mathcal{V}$  mit  $n$  Anlagen, der das Produkt erbringen soll. Gesucht ist ein Einsatzplan  $\mathcal{E}$  für den Verbund, der das Produkt erfüllt. Sei  $\mathcal{U}$  die Menge der Einheiten,  $\mathbf{p}_i$  der zur Einheit  $U_i$  gehörige Fahrplan im Einsatzplan,  $\zeta^{(P)}$  der Zielfahrplan aus dem Produkt und  $\mathcal{Q}$  die Menge der Kenngrößen. Sei ferner  $\mathcal{F}^{(U_i)}$  der Lösungsraum von Einheit  $U_i$  und  $\mathcal{I}^{(Q,U_i)}$  die für  $U_i$  spezifische Abbildung der Kenngröße  $Q$  auf die aus Fahrplan und Startzustand  $z_{0,U_i}$  der Einheit resultierende Kennzahl. Dann ist das Wirkleistungsplanungsproblem wie folgt definiert:

$$\delta \left( \sum_{i=1}^n \mathbf{p}_i, \zeta^{(P)} \right) \rightarrow \min \quad (6.14)$$

$$\wedge \sum_{U \in \mathcal{U}} \mathcal{I}^{(Q,U_i)}(\mathbf{p}_i, z_{0,U_i}) \rightarrow \text{opt}_{Q_i} \in \{\min, \max\} \forall Q \in \mathcal{Q} \quad (6.15)$$

sodass

$$\mathbf{p}_i \in \mathcal{F}^{(U_i)} \forall U_i \in \mathcal{U}. \quad (6.16)$$

Hierbei bezeichne  $\delta$  ein Abstands- oder Fehlermaß zur Bewertung der Abweichung von der Zielvorgabe. Im praktischen Einsatz wären verschiedene Maße denkbar. Für den Rest dieser Arbeit wird jedoch o.B.d.A. der Euklidische Abstand  $L_2, \|\cdot\|_2$  verwendet.

Offensichtlich handelt es sich beim Wirkleistungsplanungsproblem um ein kombinatorisches Optimierungsproblem aus der Klasse der Rucksack- (engl. Knapsack-) Probleme; genauer gesagt um ein Multiple-Choice Knapsack Problem, da für jede Anlage eine eigene Menge von Wahlmöglichkeiten existiert aus denen jeweils genau eine auszuwählen ist. In der Integer-Form (nur ganzzahlige Werte für  $p_i$ ) sind solche Probleme üblicherweise schwach NP-schwer, können aber beispielsweise durch Methoden der dynamischen Programmierung in polynomieller Zeit gelöst werden (Hinrichs et al. 2013b; Martello et al. 1990), sofern es möglich wird sich speziell die Struktur des Problems zunutze zu machen. Aufgrund der stetigen dynamischen Rekonfiguration der Verbünde wird es jedoch nicht möglich sein ein statischen Optimierungsmodell aufzustellen, dessen Struktur zu untersuchen und ein hierfür statisch angepasstes Optimierungsverfahren für jeden Verbund wiederzuverwenden. Für viele heutige Anlagen trifft es auch zu, dass Leistungsstufen nur in diskreten Schritten gewählt werden können. Ein einfaches Beispiel stellen Wärmepumpen dar, welche üblicherweise nur an oder aus sein können. Für solche Geräte konnte bereits das Finden eines nach gegebenen Kriterien optimalen Fahrplans bereits für einzelne Anlagen als NP-vollständig nachgewiesen werden (Bremer 2006). Für Anlagen mit kontinuierlichem Modulationsbereich gilt die NP-Schwere nicht mehr allgemein. Im Rahmen dieser Arbeit werden keine diskreten Fahrplanräume betrachtet. Die besondere Schwierigkeit des Wirkleistungsplanungsproblems liegt vielmehr im Finden einer geeigneten Modellierung der Constraints und deren geeigneter Behandlung während der Optimierung.

Der in dieser Arbeit behandelte Anwendungsfall der Wirkleistungsplanung weist als wichtigste Besonderheit eine Verteiltheit des Wissens um die Lösungsräume auf. Aufgrund der wiederholten, dynamischen Rekonfiguration zur Laufzeit ist es nicht möglich, ein Modell des Gesamtsystems für das Optimierungsproblem vorab zu erstellen und wiederzuverwenden, so wie es beispielsweise bei statischen Virtuellen Kraftwerken heute üblich ist. Ein Modell des Gesamtsystems müsste die individuellen Konfigurationen der beteiligten Anlagen sowie ein mathematisches Modell aus dem die individuelle Flexibilität abgeleitet werden kann bereits enthalten. Die Zusammensetzung der Verbünde ist aber vorab nicht bekannt und auch nicht konstant. Zudem können jederzeit neue Anlagen in den Prozess der Verbundbildung eintreten, deren Modellierung zur Laufzeit noch integriert werden müsste.

Es muss also eine Möglichkeit gefunden werden, die Flexibilitäten (einschließlich der individuellen Kosten für Fahrplananpassungen) der einzelnen Anlagen so zu modellieren, dass eine anlagenunabhängige, einheitliche Beschreibung entsteht aus der im Rahmen der Wirkleistungsplanung realisierbare Fahrpläne als Lösungskandidaten in systematischer Weise hergeleitet werden können. Wenn dies gelingt, kann für die eigentliche Optimierung auf Standardverfahren ohne Constraintbehandlung zurückgegriffen werden.

## 6.6. Zusammenfassung

Basierend auf den Definitionen dieses Kapitels konnte nunmehr die Problemstellung formal gefasst werden. Zu lösen ist ein kombinatorisches (Mehrziel-)Optimierungsproblem, wobei die einzelnen zu kombinierenden Komponenten der Lösung (die Fahrpläne der Anlagen) in verschiedenen, individuell durch Constraints eingeschränkten Lösungsräumen zu suchen sind. Die praktische Bedeutung, die durch die konkrete Anwendungsdomäne Smart Grid gegeben ist, wurde bereits im Einleitungsteil unterstrichen.

In diesem Abschnitt wurden insbesondere die modellierungstechnischen Besonderheiten deutlich, welche bei der Umsetzung zu beachten sind. Wie in Kapitel 3 ab Seite 23 bereits erläutert, wird in herkömmlichen Ansätzen gerade von einer adäquaten Behandlung technischer Anlagenconstraints häufig abstrahiert oder es muss auf monolithische Optimierungsmodelle mit integrierter Anlagenmodellierung zurückgegriffen werden, um durch die Einsatzplanung auch zulässige Lösungen zu generieren. Dies trifft gerade auch bei der Betrachtung längerer Zeiträume zu. Viele Ansätze vereinfachen die Problemstellung auch zu einem 1-dimensionalen Problem, bei dem lediglich das jeweils nächste Zeitintervall betrachtet wird. In diesem Fall entstehen natürlich keine Abhängigkeiten zwischen den Entscheidungen in verschiedenen Zeitschritten bezüglich der Umsetzbarkeit von Fahrplänen, jedoch wird so auch keine fahrplanorientierte Vorabplanung der verteilten Produktion möglich.

Sobald sich Anlagen in einer dynamischen Umgebung produktspezifisch zusammenfinden sollen, ist eine detaillierte Kenntnis der verschiedenen Möglichkeiten über einen längeren Zeitraum (ggf. auch über mehrere Produktzeiträume hinweg) für eine Day-Ahead-Planung erforderlich. Dann müssen aber auch zeitliche Abhängigkeiten bei der Erzeugung (oder dem Verbrauch) berücksichtigt werden. Es bedarf der Kenntnis des individuellen Lösungsraums für einen gegebenen Zeitrahmen.

## *6. Formalisierung und Modellierung der Wirkleistungsplanung im Smart Grid*

---

Genau diesem Umstand wird sich der in dieser Arbeit entwickelte Ansatz widmen, der in den nachfolgenden Kapiteln erläutert wird. Im folgenden Kapitel wird zunächst das grundlegende Konzept zur Lösung des Wirkleistungsplanungsproblems Problems vorgestellt, bevor im Anschluss die vollständige Prozesskette abschnittsweise ausgearbeitet wird.

# 7. Lösungskonzept

*Right now it's only a notion,  
but I think I can get the money to make it into a concept,  
and later turn it into an idea.*

---

*(Woody Allen)*

## 7.1. Einleitung

Bevor in den nachfolgenden Kapiteln der allgemeine Lösungsansatz im Detail beschrieben wird, soll an dieser Stelle der Lösungsweg auf konzeptioneller Ebene im Überblick hergeleitet und dargestellt werden. Insbesondere werden die einzelnen Komponenten des Ansatzes und ihr späteres Zusammenspiel vorgestellt. Für die Erläuterungen des Verfahrens erfolgt eine Beschränkung auf die reine Wirkleistungsplanung, wie sie im vorigen Kapitel beschrieben wurde. Dies stellt keine generelle Einschränkung des Verfahrens dar, die Beschränkung erleichtert lediglich die Illustration. Eine Erweiterung, die auch die Betrachtung komplexer Leistungswerte einbezieht erfolgt abschließend in Kapitel [17](#) ab Seite [231](#).

Für eine verteilte Lösung des Wirkleistungsplanungsproblems ist es notwendig, den individuellen Lösungsraum einer jeden Anlage geeignet zu modellieren. Auch im Falle eines statischen Verbundes mit zentraler Kontrollinstanz liegen essentielle Informationen über Anlagenzustände und damit über daraus resultierende individuelle Handlungsoptionen für nachfolgende Planungszeiträume verteilt bei den einzelnen Anlagen vor, wodurch eine zentrale Modellierung erschwert wird.

Ansätze, welche eine Wirkleistungsplanung auf Basis eines statischen, monolithischen Optimierungsmodells (vgl. z. B. Kallrath [2013](#), aber auch Kapitel [3](#) ab Seite [23](#)) für die Planung vornehmen, werden im Falle sich regelmäßig rekonfigurierender, dynamischer Verbünde zudem nicht länger handhabbar sein. Die Modellierung der Anlagen sowie insbesondere die Modellierung ihrer in vielfältiger Gestalt auftretenden Constraints erschwert das automatische Ableiten eines Optimierungsmodells für den Verbund. Gesucht ist daher ein Ansatz, bei dem jede Anlage eine Modellierung ihres Suchraums für einen gegebenen Planungszeitraum selbst vornehmen kann. Diese Modellierung muss auf eine Weise geschehen, die eine einheitliche Definition der verschiedenen Suchräume und somit deren Integration zu einer gemeinsamen Planung zulässt.

In diesem Kapitel soll zunächst das allgemeine Konzept zur Lösung der im vorigen Kapitel skizzierten Problemklasse vorgestellt werden. Kennzeichnend für die Problemklasse sind insbesondere:

## 7. Lösungskonzept

---

- Das verteilte Wissen um die Lösungsräume, die für eine gemeinsame Lösungsfindung geeignet zu integrieren sind.
- Die Vielzahl verschiedener Constraints unterschiedlicher Akteure, die eine einheitliche Beschreibung der Lösungsräume erschweren.

Als Lösung wurde daher ein zweistufiger Ansatz entworfen, der im Kern aus einer Komponente zur einheitlichen Modellierung der Suchräume sowie einem System zur systematischen Generierung gültiger Fahrplanalternativen während der Planung besteht. Der Raum der realisierbaren Fahrpläne muss effizient modelliert werden, sodass zum einen von beliebigen Constraints (auch von ggf. bisher unbekanntem Anlagen) abstrahiert werden kann und zum anderen auch individuelle (allerdings für alle Teilnehmer gleich ausgeprägte) Kosten- oder Effizienzkennzahlen zur Bewertung einzelner Fahrpläne integriert werden können. Eine solche Beschreibung kann effizient beispielsweise über Surrogatmodelle erfolgen, welche das technische Modell einer Anlage so abstrahieren, dass nur noch das beobachtbare Verhalten (ohne Ursachen) näherungsweise abgebildet wird. Durch solche Modelle kann für jede dezentrale Einheit der Suchraum gültiger Fahrpläne durch eine verhältnismäßig kleine Menge von Beispielalternativen vollständig beschrieben werden ohne dass eine explizite Modellierung der Constraints für das eigentliche Planungsverfahren notwendig wird. Diese Darstellung allein ermöglicht jedoch noch keine effiziente Suche in den Fahrplanräumen.

Eine besondere Schwierigkeit wird der extrem kleine Anteil zulässiger Lösungen sein. Dies gilt in besonderem Maße bei höherdimensionalen Problemstellungen wie sie in den hier betrachteten Smart Grid Anwendungen vorliegen. Hierdurch wird es äußerst schwierig, existierende Heuristiken beispielsweise durch Ergänzen einer Penaltyfunktion für ungültige Lösungen out-of-the-box einzubinden. Die Wahrscheinlichkeit, im Zuge der Suchraumexploration eine gültige Lösung zu erhalten, ist aufgrund des Verhältnisses von gültigen und ungültigen Lösungen viel zu gering. Eine vorherige Einschränkung der Suche auf die Region gültiger Lösungen gestaltet sich aber ebenfalls schwierig, da zum einen keine funktionale Beschreibung der Randbedingungen vorliegt und zum anderen ggf. in nicht zusammenhängenden Regionen gesucht werden muss. Das Suchraummodell wird es ermöglichen, im Nachhinein zu überprüfen, ob ein gefundener Lösungskandidat zur Lösungsmenge gehört oder nicht. Da dies für eine systematische Suche nicht reicht, wird zusätzlich ein so genannter Dekoder entwickelt, welcher eine Abbildung des Raumes aller Fahrpläne in den Raum der gültigen Fahrpläne erlaubt und somit dazu dienen kann, das Optimierungsproblem vollständig in einem unbeschränkten Suchraum zu lösen.

Dieses Konzept wird im weiteren Verlauf dieses Kapitels detaillierter ausgearbeitet.

### 7.2. Lösungskonzept

Für die Lösung des Optimierungsproblems zur Wirkleistungsplanung, wie im vorigen Kapitel beschrieben, müssen verschiedene, verteilte Lösungsräume einbezogen werden. Generell sind zwei grundsätzliche Optimierungsansätze denkbar:

1. Im Falle eines zentralen Ansatzes (auch in einem Multiagentensystem mit einem speziell für die Steuerung des Verbundes zuständigen Agenten) müssen die Lösungs-

räume von den einzelnen Einheiten bestimmt und kommuniziert werden, damit die zentrale Kontrolleinheit darauf zugreifen kann. Kennzahlen zur Bewertung der einzelnen Fahrpläne müssen in dieser Fahrplanraumdarstellung enthalten sein, da diese ebenfalls zentral benötigt werden.

2. Im Falle eines verteilten Ansatzes muss der Lösungsraum zulässiger Fahrpläne zum Zeitpunkt der Optimierung ebenfalls bekannt und geeignet beschrieben sein. Um dem Optimierungsverfahren eine systematische und zielgerichtete Navigation in diesem Lösungsraum zu ermöglichen, muss dieser auch hier zu Beginn der Optimierung vollständig beschrieben vorliegen. Je nach verwendetem Verfahren müssen gegebenenfalls auch Lösungsräume anderer Einheiten bekannt sein und somit kommuniziert werden. Als Beispiel mag das Verfahren OptAPO (Mailler & Lesser 2004) für DCOP-Problem dienen, welches nach dem Konzept der kooperativen Mediation arbeitet, bei dem Agenten dynamisch als Mediatoren gewählt werden, um für ihre Nachbarschaft Teilprobleme zentral zu lösen, bevor diese Teillösungen sukzessive zur Gesamtlösung aggregiert werden.

In jedem Fall wird eine Beschreibung der Lösungsräume der verschiedenen Einheiten benötigt. Prinzipiell ist es möglich, die Flexibilitäten einer Einheit ausgehend von einem gegebenen Anlagenzustand und einer Beschreibung ihrer Freiheitsgrade analytisch herzuleiten. Häufig sind jedoch für eine analytische, exakte Herleitung keine hinreichenden mathematischen Modelle gegeben. Meist wird eine Bestimmung der Flexibilitäten ohnehin zumindest teilweise auf Prognosen (beispielsweise über den zukünftigen thermische Bedarf oder das Wetter) beruhen und damit eine Abschätzung bleiben. Als weitere Schwierigkeit kommt bei diesem analytischen Ansatz hinzu, dass ein solches Verfahren detailliert auf jede Anlage angepasst werden müsste. Das gilt nicht nur für verschiedene Anlagentypen, sondern auch für Anlagen eines Typs an verschiedenen Standorten, da die individuelle Einbettung in den jeweiligen Betriebskontext in die Bestimmung der Flexibilitäten einbezogen werden muss. Das automatisierte Ableiten einer Beschreibung der Flexibilitäten stellt sich also als schwierig dar.

Bei geeigneter Darstellung der Daten ist gerade dies aber eine Stärke von Machine-Learning-Verfahren. Regressionsverfahren können aus vergangenem, beobachtetem Verhalten ein Modell des zukünftigen Verhaltens automatisiert erstellen, um beispielsweise Vorhersagen zu machen. Klassifikationsverfahren können aus einer Menge von Daten einer bestimmten Klasse (im Beispiel hier: realisierbare Fahrpläne ausgehend von einem gegebenen Startzustand) ein Modell ableiten, mit dem weitere Daten hinsichtlich ihrer Zugehörigkeit zu dieser Klasse getestet werden können.

Letzteres soll in dem hier vorgestellten Konzept genutzt werden, um ein Modell des Suchraums zu erstellen, das ...

- ... eine Modellierung der Suchräume auf Anlagenebene vornimmt, da der dynamische Charakter der betrachteten Verbünde keine statische Modellierung auf Verbundebe-  
ne zulässt.
- ... eine Abstraktion von jedweder konkreter mathematischer Anlagen-Modellierung  
sowie von konkreten, individuellen Constraintformulierungen vornimmt, um die  
dynamische, algorithmische Integration zu einem Gesamtmodell bewerkstelligen zu  
können.

### 7.2.1. Suchraummodell

Für den Suchraum wird ein Modell gesucht, das folgende Eigenschaften besitzt:

- Das Modell soll in der Lage sein, zu entscheiden ob ein gegebener Fahrplan (wie er im Rahmen eines Einsatzplanungsverfahrens als Lösungskandidat generiert wird) von der Anlage technisch realisiert werden kann oder nicht.
- Für bestimmte Constraint-Handling-Techniken ist es notwendig eine zumindest qualitative Aussage darüber zu machen welcher von zwei gegebenen Fahrplänen eher realisierbar ist. Eine Aussage hierüber muss möglich sein.
- Beliebige Constraints müssen in das Modell integriert werden können.
- Eine gegebene Menge von Suchraummodellen einzelner Anlagen muss sich dynamisch und vor allem automatisiert zu einem Gesamtmodell für die Planung zusammenfassen lassen.
- Die verschiedenen Fahrpläne müssen individuell mit Bewertungskennzahlen für die Mehrzieloptimierung versehen werden können.
- Das Modell muss von konkreten Annahmen über die Fahrpläne (Zeitraum, Länge und Anzahl eines Zeitintervalls, usw.) abstrahieren.

Wird die Modellierung heruntergebrochen auf die Modellierung der Menge realisierbarer Fahrpläne, wodurch die Notwendigkeit verschiedenste Constraints mathematisch beschreiben zu müssen entfällt, so bietet sich die Verwendung eines 1-Klassen-Klassifikators (Moya et al. 1993) an. Während bei der Mehrklassen-Klassifikation versucht wird, ein Datum in eine von mehreren (meistens 2) vordefinierten Klassen einzusortieren, gibt es beim 1-Klassen-Fall nur eine Klasse, die gerade durch die bekannten Trainingsdaten definiert ist (Khan & Madden 2010). Üblicherweise kommen solche Verfahren zum Einsatz, wenn im Wesentlichen nur positive (zur Klasse gehörig) Instanzen von Trainingsdaten bekannt sind. Die Schwierigkeit liegt darin, die Grenze (zwischen Zugehörigkeit und Nicht-Zugehörigkeit) nur von einer Seite aus definitiv bestimmen zu können. Daher ist das Klassifikationsproblem im 1-Klassen-Fall auch schwieriger und bedarf in der Regel einer größeren Anzahl an Trainingsinstanzen (Khan & Madden 2010).

Constraints sind in einem solchen Modell nicht mehr mathematisch beschrieben, sondern sind lediglich noch impliziert durch ihre Auswirkungen vertreten. Beschrieben ist ein Unterraum des Vektorraums aller Fahrpläne, welcher durch die Constraints begrenzt ist, jedoch beschrieben durch eine Menge von Beispielen aus diesem Unterraum und nicht länger durch die Constraints selbst.

Eine Umsetzungsmöglichkeit für einen solchen Klassifikator wurde bereits in Abschnitt 5.3 auf Seite 60 vorgestellt. Die Supportvektor-Variante hat den Vorteil, dass über das Konzept des Abstandes vom Kugelmittelpunkt auch die Möglichkeit einer Ordnungsrelation (hier sogar qualitativ) für die Nähe zur Realisierbarkeit gegeben ist. Ein solches auf einem Klassifikator basierendes Modell ist jedoch nur für die Überprüfung von bereits bekannten Fahrplänen hinsichtlich ihrer Realisierbarkeit geeignet. Für die Umsetzung einer systematischen Generierung von realisierbaren Fahrplänen jedoch bedarf es noch einer Erweiterung.

### 7.2.2. Systematische Lösungsgenerierung

Mit dem Modell liegt zwar eine Beschreibung der individuellen Möglichkeiten der einzelnen Anlagen vor, benötigt wird jedoch noch eine Technik zur systematischen Suche nach einer optimalen Lösung. Optimierungsverfahren müssen in der Lage sein, verschiedene Lösungsalternativen systematisch zu bewerten. Auch bei Verwendung von Heuristiken (ohne vollständige Enumeration aller Lösungen) muss eine Möglichkeit existieren, eine Folge von gültigen Lösungen für die individuelle Bewertung zu erzeugen. Wie dies geschieht ist durch das jeweilige Verfahren vorgegeben. Bei allen evolutionären Optimierungsverfahren wird ausgehend von einer aktuell bekannten Lösung eine Menge von benachbarten weiteren Lösungen generiert und bewertet. Die Definition von Nähe und Nachbarschaft ist üblicherweise ebenfalls verfahrensabhängig und macht insbesondere meist keine Unterscheidung von gültiger und ungültiger Lösung. Will ein Optimierungsverfahren nur gültige Lösungen betrachten, so braucht es entweder einen Reparaturmechanismus, der in der Lage ist, eine ungültige Lösung zu erkennen und in eine gültige Lösung umzuwandeln, oder es wird ein Suchraum benötigt, der von vornherein nur gültige Lösungen enthält.

Hierzu wird im Folgenden ein Ansatz entwickelt, den unbeschränkten Suchraum mithilfe des Klassifikatormodells so auf den beschränkten Teilraum der gültigen Lösungen abzubilden, dass eine Definition von beispielsweise Nachbarschaftsbeziehungen im unbeschränkten Raum möglich wird während gleichzeitig die Auswertung der Zielfunktion unter der Abbildung im eingeschränkten Lösungsraum vorgenommen werden kann. Zugeordnete Kostenkennzahlen sollen bei diesem Vorgehen erhalten bleiben und mit transformiert werden. Auf diese Weise wird es möglich, Restriktionen und individuelle Kosten so in einen beliebigen Optimierungsansatz zu integrieren, dass diese nicht mehr explizit modelliert werden müssen. Das Optimierungsproblem als Ganzes wird in ein unbeschränktes transformiert, um dort mit erprobten Optimierungsverfahren gelöst zu werden, auch wenn diese nicht für die Verwendung in eingeschränkten Suchräumen konzipiert sind. Auf diese Weise wird es möglich sein, die Suchräume verschiedener Anlagentypen unabhängig von der Anlagenmodellierung und unabhängig von den individuellen Constraints die bei den verschiedenen Anlagentypen zum tragen kommen einheitlich zu behandeln und zu durchsuchen. Man spricht bei dieser Art von Ansatz bei dem ein Optimierungsproblem in einem einfacher zu handhabenden Suchraum gelöst wird auch vom Space-Mapping (Bandler et al. 1994; Brock 2004).

Konzeptionell eignen sich Dekoder (vgl. Kapitel 9 ab Seite 117) für diese Aufgabe. Daher wird in dieser Arbeit ein Dekoder entworfen, der automatisiert aus dem zuvor beschriebenen Klassifikator abgeleitet werden soll.

### 7.2.3. Prozesskette

Zusammenfassend ergibt sich aus den Anforderungen die in Abbildung 7.1 dargestellte Prozesskette für die dynamische Integration der Flexibilitäten der einzelnen Anlagen in ein Planungs- bzw. Optimierungsverfahren.

**Einheitensimulation** Um die Flexibilität einer Einheit abschätzen zu können, wird initial ein mathematisches Modell der Einheit benötigt. Dieses kann je nach Anlage individuell implementiert sein, da lediglich ein minimales Interface vorgeben werden

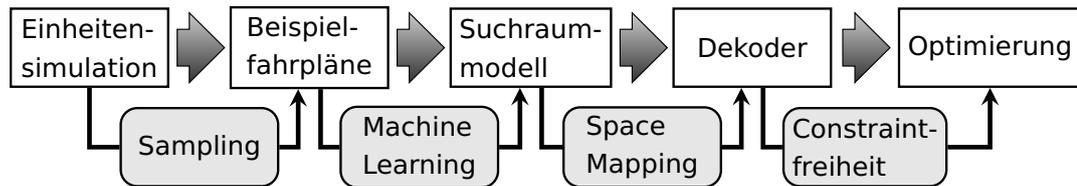


Abbildung 7.1.: Prozesskette für die dynamische Integration von Flexibilitäten in die (verteilte) Optimierung mittels Dekoder.

muss, mit dem gegebene Fahrpläne auf ihre Realisierbarkeit hin geprüft werden können.

**Beispielfahrpläne** Mit Hilfe des Einheitensimulationsmodells kann mit geeigneten Samplingverfahren eine Menge an realisierbaren Fahrplänen generiert werden (vgl. Abschnitt 8.2 auf Seite 102).

**Suchraummodell** Liegt erst einmal eine Trainingsmenge mit realisierbaren Fahrplänen vor, so kann hieraus mit geeigneten Klassifikatoren ein Modell gelernt werden, welches von konkreten mathematischen Einheitenmodellen, insbesondere aber von jedweder mathematischer Constraintbeschreibung abstrahiert (vgl. Abschnitt 8.3 auf Seite 110).

**Dekoder** Liegt ein Suchraummodell vor, welches ein geometrisches Modell des Unterraums mit den realisierbaren Fahrplänen enthält (der Klassifikator aus dem vorherigen Schritt), so kann hieraus eine Vorschrift abgeleitet werden, den Raum aller Fahrpläne auf diesen Unterraum abzubilden. Eine solche Abbildungsvorschrift schafft sodann ein Mittel, systematisch Fahrpläne zu generieren, die von der jeweiligen Anlage auch realisiert werden können. Eine Kenntnis des mathematischen Anlagenmodells ist hierfür nicht mehr erforderlich. (vgl. Kapitel 9 ab Seite 117).

**Optimierung** Mit Hilfe des Dekoders wird es schließlich möglich, die Optimierung bei der Aufteilung der Fahrpläne während der Wirkleistungsplanung in einem Raum vorzunehmen, der nicht durch Constraints beschränkt ist (vgl. Abschnitt 12.3 auf Seite 180).

Die einzelnen Komponenten werden in den nachfolgenden Kapiteln näher beschrieben.

### 7.3. Voraussetzungen an die Anlagen

In Laufe dieser Arbeit wird zumeist das BHKW als durchgängiges Beispiel verwendet. Dennoch ist der hier vorgestellte Ansatz unabhängig von konkreten Anlagentypen. Es sollen generell vorab keinerlei Einschränkungen bezüglich der Anlagen(-Typen) gemacht werden, die mit dem nachfolgend genauer vorgestellten Verfahren im Rahmen einer dynamischen Verbundbildung automatisiert zu einem gemeinschaftlichen Optimierungsmodell für die Wirkleistungsplanung integriert werden können. Es ist in der Praxis notwendig, ein derart offenes System zu haben, da vorab nicht bekannt ist welche Arten von Anlagen später einmal teilnehmen könnten. Es ist durchaus zu erwarten, dass auch nach Inbetriebnahme eines derartigen Systems regelmäßig neue Anlagentypen zu integrieren sind, die derzeit

noch gar nicht in Erwägung gezogen werden – beispielsweise weil die Rentabilität derzeit noch nicht gegeben ist. Bisher nicht betrachtete (oder gegebenenfalls noch nicht einmal bekannte) Anlagen unterliegen aber gegebenenfalls auch neuen technischen Randbedingungen, die ebenfalls ad hoc integriert werden müssen. Ebenfalls nicht im Fokus dieser Arbeit liegt die Problematik des Auffindens geeigneter, anderer Anlagen als mögliche Partner. Ansätze für eine verteilte Registry, die dies leisten könnten, finden sich beispielsweise in (Dethlefs & Renz 2013).

Dennoch müssen einige wenige Einschränkungen gemacht werden, um eine Anlage integrieren zu können. Folgende Voraussetzungen muss eine Anlage erfüllen, um mit dem hier vorgeschlagenen Konzept in den Prozess der Wirkleistungsplanung (auch in dynamischen Verbänden) eingebunden werden zu können:

1. Die Flexibilitäten der Anlage müssen durch Mengen von Fahrplänen ausdrückbar sein, die für diskrete Zeitintervalle eine mittlere Wirkleistung vorgeben. Wie später noch gezeigt wird ist auch die Vorgabe komplexer Leistungswerte möglich.
2. Die Anlage kann durch einen Softwareagenten gesteuert werden. Dieser muss nicht zwingend auf dem Anlagencontroller selbst ausgeführt werden, wenn der Agent Fernsteuerungszugriff auf die Anlage hat und zudem notwendige Anlagenparameter fernauslesen kann.
  - a) Der Softwareagent benötigt eine Modellvorstellung von der Anlage, die er kontrolliert. Hierzu muss er über ein Simulationsmodell der Anlage verfügen, dass mit aktuellen Messwerten von der Anlage parametrisiert werden kann. Alternativ kann das Simulationsmodell auch getrennt vom Agenten beispielsweise auf dem Anlagencontroller betrieben werden.
  - b) Das Simulationsmodell muss als Minimalanforderung die Überprüfung eines gegebenen Fahrplans auf Realisierbarkeit durch die Anlage ermöglichen.
  - c) Der Softwareagent benötigt eine Kommunikationsanbindung, um mit anderen Agenten in einem Verbund kommunizieren zu können. An die Transportschicht der Kommunikation werden keine speziellen Anforderungen gestellt. Anforderungen an das kommunizierte Datenformat ergeben sich aus der konkreten Anwendung (durch das jeweils eingesetzte Planungsverfahren definiert). Muss das Suchraummodell selbst kommuniziert werden, ist die Verfügbarkeit von (mehrdimensionalen) Arrays als Datenformat sinnvoll.
3. Die Anlage muss durch Vorgabe eines Fahrplans gesteuert werden können. Das bedeutet, dass die Wirkleistungsvorgaben aus dem Fahrplan von einem Anlagencontroller in geeignete Steuerbefehle für den Betrieb der Anlage umgesetzt werden müssen.

Die einzige Bedingung an zeitliche Anforderungen ist die, dass alle betrachteten Zeiträume immer durch diskrete Zeitintervalle dargestellt werden können. Dabei soll keine Beschränkung auf die heute in der Energiewirtschaft üblichen 15-Minuten-Intervalle gegeben sein, auch wenn im weiteren Verlauf dieser Arbeit diese Zeitgranularität als durchgehendes Beispiel verwendet wird. Das Verfahren selbst wird mit zeitlich beliebig aufgelösten Fahrplänen umgehen können, auch wenn für zu kleine Zeiträume (etwa im Sekundenbereich) die Sinnhaftigkeit beispielsweise thermischer Lastprognosen o. ä. angezweifelt

werden darf. Es ist ebenfalls keine Anforderung, dass alle Zeitintervalle die gleiche Länge haben. Diese darf auch zur Laufzeit variieren. Lediglich für die Dauer eines Wirkleistungsplanungs Vorgangs muss die Struktur eines Fahrplans (Anzahl Intervalle und die jeweilige Länge) fest bleiben und für alle beteiligten Anlagen gleich sein.

### 7.4. Zusammenfassung

Die Fragestellung nach der idealen Arbeitsteilung bei der Wirkleistungsbereitstellung innerhalb eines (volatilen oder statischen) Verbundes von Einheiten und somit die Frage nach der optimalen Aufteilung eines Gesamtlastgangs auf Fahrpläne für die einzelnen Einheiten wirft verschiedene Detailfragen auf. Die Aufteilung der Fahrpläne muss so erfolgen, dass die technischen Rahmenbedingungen jeder einzelnen Einheit bzgl. der Realisierbarkeit gewahrt werden. Um dies zu erreichen, ist in beinahe allen bekannten Ansätzen eine zentrale Modellierung in die Optimierung integriert. Ein solches Vorgehen hat aber einige Nachteile:

1. Änderungen im Profil einer Einheit ziehen Änderungen im zugehörigen Modell nach sich. Im Fall einer monolithischen Modellierung ist hiervon auch das Optimierungsmodell betroffen.
2. Die Zusammensetzung eines Verbundes von dezentralen Einheiten kann nicht dynamisch verändert werden ohne auch Änderungen am zentralen Modell vorzunehmen.
3. Die direkte Integration einer großen Anzahl von Einheitenmodellen (als mathematische Simulationsmodelle) in die Optimierung ist problematisch in Bezug auf die Performanz.
4. Eine zentrale Modellierung erlaubt schwerlich eine individuelle Abschätzung von beispielsweise Umweltwirkungen basierend auf aktuellen (vor Ort) Gegebenheiten, lokalen Kosten o. ä.

Einen eleganten Ansatz bietet hier die dezentrale Surrogat-Modellierung der Einheiten. Bei einem solchen Ansatz werden die Mengen von realisierbaren Lastgängen, die eine Einheit als Alternativen jeweils anbieten kann, durch ein Meta-Modell des Zustandsraums der Lastgänge und seiner Constraints beschrieben. Dieses Vorgehen bietet diverse Vorteile:

1. Jede (neue) Einheit, die eine solche standardisierte Modellierung verwendet, kann einschließlich mitmodellierter Kennzahlen (on-the-fly) in ein hierauf arbeitendes Optimierungsprotokoll integriert werden.
2. Während der Optimierung müssen keine individuellen Einheitenconstraints mehr berechnet werden; diese sind in dem Meta-Modell bereits implizit enthalten.
3. Die dezentrale Modellierung erleichtert die Verwendung verteilter Optimierungsverfahren.
4. Eine dezentrale Einheitenmodellierung verteilt die Rechenlast und führt zu einer deutlichen Steigerung der Performanz während der Optimierung.
5. Insbesondere können aber auch Kennzahlen zu den einzelnen Fahrplänen lokal ermittelt und in das Modell integriert werden. Somit stehen beispielsweise Umwelt-

kennzahlen (Reporting), Konfidenzwerte für die Umsetzbarkeit bestimmter Fahrpläne (robuste Planung) oder individuelle Kosteninformationen zur Verfügung.

In diesem Kapitel wurde auf konzeptioneller Ebene eine Prozesskette entwickelt. Sie zeigt bereits das angestrebte Vorgehen bei der Lösung des Wirkleistungsplanungsproblems als abstrakten Prozess. Der Vorteil dieser Zerlegung liegt darin, dass für die einzelnen Prozessschritte verschiedene Umsetzungen implementiert und einfach gegeneinander ausgetauscht werden könnten. Zwar werden im Rahmen dieser Arbeit nicht für alle Schritte mehrere Implementierungen betrachtet, für die Flexibilität und Anpassungsfähigkeit einer späteren Umsetzung bringt eine solche Modularisierung jedoch zusätzliche Vorteile. Nachdem in diesem Kapitel der generelle Ablauf erarbeitet und mit seinen Komponenten vorgestellt wurde, werden in den nachfolgenden Kapiteln die einzelnen Komponenten formal betrachtet und hinsichtlich ihrer Implementierungen diskutiert.



## 8. SVDD als Blackbox-Modell

*Modelle sollten sich bemühen,  
dem Porträt ähnlich zu sehen.*

---

*(Salvador Dalí)*

### 8.1. Einleitung

Als erste wichtige Komponente des Lösungsansatzes sollen in diesem Kapitel das eigentliche Suchraummodell und insbesondere seine Konstruktion hergeleitet werden. Aufgabe des Modells wird es sein, die Menge der realisierbaren Fahrpläne einer Einheit zu erfassen und zu beschreiben. Als Beschreibung für die Menge der realisierbaren Fahrpläne soll der in Abschnitt 5.3 auf Seite 60 beschriebene, auf Supportvektoren beruhende 1-Klassen-Klassifikator verwendet werden. Der Vorteil der geschlossenen Entscheidungsgrenze (wie sie von solchen Verfahren erzeugt wird) erlaubt die Definition einer einschließenden Hülle, welche alle realisierbaren Fahrpläne umschließt und somit ein Modell des Lösungsraums darstellt.

In der derzeitigen wissenschaftlichen Praxis dienen die mathematischen Anlagenmodelle im Rahmen von operativen Planungsfragestellungen meist dazu, den Handlungsspielraum analytisch zur Laufzeit (erst während der laufenden Optimierung) abzuleiten. Ein solches Vorgehen ist sowohl anlagen- als auch modellspezifisch. Hieraus folgt, dass die Modelle bei derartigem Vorgehen meist integraler und individuell umgesetzter Bestandteil der Implementierung eines Planungsverfahrens und somit des Optimierungsmodells sind. Änderungen aufgrund einer sich dynamisch ändernden Verbundzusammensetzung ziehen somit auch zwangsläufig Änderungen eines monolithischen Anlagen-Optimierungsmodells nach sich. Eine automatisierte Anpassung zur Laufzeit ist hier insbesondere für komplexere Modelle schwierig. Nebenbei sei angemerkt, dass auch solche analytischen Modelle eine Vereinfachung der Realität darstellen und häufig ebenfalls in Teilen auf Prognosen beruhen; sie stellen somit auch immer nur Näherungen der Realität dar. In dieser Arbeit soll ein anderer Weg gewählt werden.

Das angestrebte Modellierungskonzept wurde auf abstrakter Ebene bereits im vorigen Kapitel erläutert. Gelernt werden soll eine geometrische Repräsentation des Unterraums der realisierbaren Fahrpläne mithilfe des 1-Klassen-Verfahrens SVDD (vgl. Abschnitt 5.3 auf Seite 60). Für das Lernen eines Modells des Lösungsraums wird als Eingabe eine Trainingsmenge von Beispielfahrplänen benötigt. Die Erstellung einer solchen Trainingsmenge bringt einige Besonderheiten mit sich und wird daher nachfolgend als erstes behandelt. Anschließend erfolgt eine Beschreibung der Umsetzung des Suchraummodells als Supportvektor-Klassifikator.

## 8.2. Erstellen einer Trainingsmenge

Für das in diesem Kapitel betrachtete Suchraummodell wird als Ausgangsvoraussetzung das Vorliegen einer Trainingsmenge von realisierbaren Fahrplänen benötigt. Für eine gegebene Anlage (bzw. im Simulationsfall für deren Simulationsmodell) muss diese Menge bestimmt werden, um den individuellen Suchraum aus den Eigenschaften ableiten zu können. Die Trainingsmenge dient als Vorlage für die räumliche Lage der umsetzbaren Fahrpläne im Raum aller Fahrpläne. Diese räumliche Beziehung soll aus der Trainingsmenge als Muster für das Modell abgeleitet werden.

Zur Verfügung steht je Anlagentyp ein Einheiten-Simulationsmodell das in der Lage ist, als charakteristische Funktion die Zugehörigkeit eines Fahrplans zur Menge der realisierbaren Fahrpläne anzuzeigen. Wir interessieren uns nunmehr für die Definitionsmenge dieser Funktion ohne den Kern dieser Abbildung. Hiervon soll ein Sample erstellt werden, ohne die zu sampelnde Grundgesamtheit vorab explizit zu kennen. Diese kann somit auch nicht enumeriert werden. Es ist lediglich möglich, mit dem Simulationsmodell zu prüfen, ob ein gegebener Fahrplan realisierbar ist oder nicht.

Um die Trainingsmenge (ein Sample) erstellen zu können, wird ein geeignetes Samplingverfahren benötigt. Sampling bezeichnet ein Vorgehen, aus einer für die Weiterverarbeitung zu großen Grundgesamtheit eine repräsentative Stichprobe zu bestimmen. Unter der Voraussetzung, dass die Grundgesamtheit aus der eine Trainingsmenge ausgewählt werden soll (zumindest prinzipiell) bekannt ist, gibt es bekannte und anerkannt Verfahren auf die zurückgegriffen werden kann. Einen Überblick geben beispielsweise Cochran (1977) oder Moore & McCabe (1993). Mit solchen Verfahren können auch Anforderungen an die Repräsentativität der Stichprobe sichergestellt werden. Mit Verfahren wie beispielsweise dem Reservoir-Verfahren (Witten et al. 2011) kann z. B. sichergestellt werden, dass jedes Element die gleiche Wahrscheinlichkeit (zumindest bezogen auf die bisherigen Beobachtungen) hat in die Stichprobe zu kommen, wenn die Elemente der Grundgesamtheit lediglich nach und nach beobachtet werden können und die gesamte Menge nicht vorab bekannt ist.

Da die Menge der realisierbaren Fahrpläne vorab nicht bekannt ist, scheiden derartige Verfahren aber aus. Active Learning (Settles 2012) stellt eine im internationalen Forschungskontext aktuell untersuchte Methodik dar, die ein Modell für die Grenze zwischen zwei Klassen (hier zwischen realisierbaren und nicht realisierbaren Fahrplänen) iterativ annähert und stetig verbessert. Zwar lassen sich hiermit prinzipiell recht genaue Modelle erreichen, jedoch erfolgt die Verbesserung des Klassifikatormodells erst während es benutzt wird und erfordert somit zur Laufzeit noch ständigen Zugriff auf das Anlagensimulationsmodell (da es sich um ein semi-überwachtes Verfahren handelt), wodurch eine merkliche Reduktion der Performanz bei der Einsatzplanung zu erwarten wäre. Alternativ müsste der Vorab-Lernprozess iterativ wiederholt werden (Gieseke & Kramer 2013) bis eine hinreichende Genauigkeit erreicht ist. Ziel soll es hier aber sein, anhand eines vorab erstellten Samples ein Modell einmalig zu bestimmen. Unabhängig hiervon, ist im Falle einer eventuell notwendig werdenden späteren Umplanung der Anlagen gegebenenfalls eine vollständige Neuberechnung des Modells erforderlich (Nieße et al. 2014).

Benötigt wird also eine Trainingsmenge von durch die jeweilige Anlage realisierbaren Beispielfahrplänen. Ein mögliches Basisverfahren für den hier vorliegenden Fall ist das

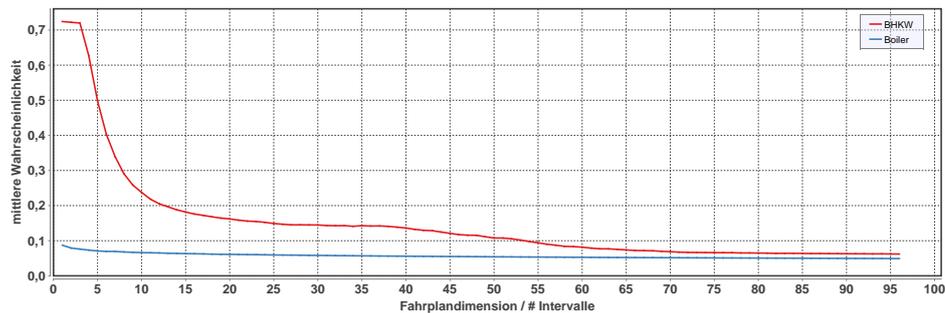


Abbildung 8.1.: Abhängigkeit der mittleren Wahrscheinlichkeit je Dimension für einen zufälligen Fahrplan realisierbar zu sein von der Länge der Fahrpläne; exemplarisch für ein BHKW und ein Untertischwarmwassergerät.

Rejection Sampling (Hammersley & Handscomb 1964) wie es beispielsweise in (Diaconis et al. 2012) zum Sampling einer Torusoberfläche verwendet wird. Hierbei werden Punkte gleichverteilt aus einem die zu sampelnde Zielregion einschließenden Hyperrechteck zufällig generiert und auf Zugehörigkeit zur Zielregion geprüft. Liegt der Kandidat außerhalb der Zielregion, so wird er verworfen. Die Effizienz dieses Verfahrens definiert sich offensichtlich durch das Verhältnis aus Größe der Zielfunktion zur Größe des umgebenden Rechtecks. Beim Fahrplansampling würde Rejection Sampling bedeuten, einen Fahrplan zufällig gleichverteilt aus dem Einheitshyperwürfel zu generieren und mit Hilfe des Anlagenmodells auf Zugehörigkeit zum Lösungsraum hin zu überprüfen. Aufgrund der hohen Dimensionalität der Problemstellung (96 bei Tagesfahrplänen mit Viertelstundenauflösung) ist es allerdings nicht praktikabel, vollständige Fahrpläne als Ganzes zu raten und durch das Anlagenmodell zu überprüfen. Folgendes Beispiel mag das verdeutlichen: Wenn in jedem Zeitintervall eines 96-dimensionalen Fahrplans ein Viertel der Möglichkeiten durch gegebene Constraints als nicht realisierbar ausgeschlossen ist, so ergibt sich (zumindest für Anlagen mit unabhängig wählbaren Leistungswerten) für den gesamten Fahrplan eine Wahrscheinlichkeit von  $2^{-96} = 0.25^{96} \approx 1.59 \times 10^{-58}$  realisierbar zu sein. Könnten 1 Milliarde zufällig generierter Fahrpläne pro Sekunde überprüft werden, so würde man dennoch nur etwa alle 1450 Jahre einen Fahrplan für die Trainingsmenge finden. Gegebenenfalls könnte diese Quote noch durch den Einsatz von a priori Wissen verbessert werden; so kann man etwa bei einem Untertischwarmwassergerät davon ausgehen (simulativ ermittelt), dass die mittlere (skalierte) Leistung je Intervall lediglich aus den Intervall  $[0, 0.25]$  stammen kann und nicht aus dem gesamten Intervall  $[0, 1]$ .

Abbildung 8.1 zeigt beispielhaft für zwei Gerätemodelle, eine Abschätzung (Monte Carlo Methode) wie die mittlere Wahrscheinlichkeit  $\bar{P}$  je Dimension und damit den Anteil der in jedem Zeitabschnitt im Mittel realisierbaren Fahrpläne von der Länge der Fahrpläne und damit der Dimension des Suchraums abhängt. Die unterschiedliche Abhängigkeit begründet sich zum einen durch den Gerätetyp: Je nach Speicherkapazität kann z. B. ein BHKW mit Pufferspeicher anfangs große Flexibilität (und damit viele Möglichkeiten) bieten, die in Abhängigkeit von der Restkapazität für längere Zeiträume dann stark abnimmt. Zum anderen liegt dies am so genannten Fluch der Dimensionalität. So hat ein in ein Einheitsquadrat einbeschriebener Kreis ein größeres Verhältnis zum Quadrat als eine im Einheitswürfel einbeschriebene Kugel.

## 8. SVDD als Blackbox-Modell

---

Tabelle 8.1.: Exemplarische Größe verschiedener Lösungsräume am Beispiel von BHKW und Unterschwarzwassergerät (jeweils aus dem gleichen Startzustand heraus simuliert).

Dim. $d$	BHKW	Boiler
8	5.0774e-5	4.5689e-10
16	8.4608e-13	6.0600e-20
24	2.4358e-20	4.5788e-30
32	8.6241e-28	2.1021e-40
48	9.8909e-46	2.3879e-61
96	1.3199e-116	2.6262e-126

Eine Abschätzung der Größe des Lösungsraums (Anteil realisierbarer Fahrpläne im Fahrplanraum) und damit die Wahrscheinlichkeit einen realisierbaren Fahrplan als Ganzes richtig zu raten ergibt sich zu  $\bar{P}^d$  für Fahrpläne über  $d$  Zeitintervalle. Tabelle 8.1 zeigt exemplarische Lösungsraumgrößen für die beiden Geräte.

Das Vorgehen des Rejection Sampling, einen zufälligen Lastgang zu generieren (raten), mit Hilfe des Anlagenmodells auf Realisierbarkeit zu prüfen und im Falle der Validität in das Sample aufzunehmen, kann hier also im Allgemeinen aufgrund der viel zu geringen Größe des Lösungsraums nicht angewandt werden. Um dennoch mit vertretbarem Zeitaufwand (perspektivisch auch auf eher leistungsbeschränkten, eingebetteten Mikrocontrollern) eine Trainingsmenge erstellen zu können, wird die Forderung an die statistische Repräsentativität aufgegeben, da es ohnehin nicht Ziel ist, mit Mitteln der Statistik aus Eigenschaften des Samples auf Eigenschaften der Grundgesamtheit zu schließen. Dann wird das folgende Verfahren zur Trainingsmengenerstellung möglich: statt einen vollständigen Fahrplan zur raten und im Nachgang auf Realisierbarkeit zu prüfen, soll dies nun Abschnittsweise geschehen.

Die Grundidee ist dabei folgende:

1. Rate eine zufällige Belegung (Leistungswert) nur für das erste Zeitintervall;
2. Überprüfe die Realisierbarkeit mit Hilfe des Anlagenmodells;
3. Wenn der bereits gefundene Teil realisierbar ist, dann fahre mit dem nachfolgenden Zeitintervall fort, sonst rate erneut eine zufällige Belegung für das aktuelle Zeitintervall.

Das Flussdiagramm Abbildung 8.2 zeigt eine mögliche Umsetzung, wobei die Implementierung des Samplingverfahrens und das Anlagenmodell (beispielsweise in einer möglichen Agentenumsetzung) zusammenarbeiten. Die Prozedur beginnt mit der Initialisierung eines Anlagenmodells (Simulationsmodell). Dieses Anlagenmodell muss folgende Eigenschaften besitzen:

- Es muss mit einem bestimmten Anlagenzustand parametrisiert werden können, so dass weitere Betrachtungen aus einem hypothetischen (zukünftigen) Startzeitpunkt heraus durchgeführt werden können.
- Das Modell muss entscheiden können, ob ein gegebener Fahrplan aus diesem Anlagenstartzustand heraus im Rahmen der zulässigen Freiheitsgrade gefahren werden könnte, oder ob er gegen technische Constraints der Anlage verstoßen würde.
- Hierzu müssen je nach modellierter Anlage ggf. auch Prognosen über externe Einflüsse (Wetter, Verlauf des geschätzten Wärmebedarfs, usw.) einbezogen werden können.

Voraussetzung ist die Existenz eines solchen Modells. Im Bereich der modellgestützten

prädiktiven Regelung (engl. Model Predictive Control) hat die Verwendung ähnlicher Modelle bereits eine hohe Akzeptanz in der Industrie erfahren (Grüne & Pannek 2011; Bemporad & Morari 1999), sodass an dieser Stelle eine gewisse Bereitschaft zum Einsatz solcher Modelle als gegeben angenommen sei.

Sind die zu erwartenden Parameterwerte zum Startzustand nicht direkt bekannt, wird das Modell mit Messwerten der Anlage initialisiert. Basierend auf einem gegebenen Fahrplan mit dem die Anlage aktuell gerade gefahren wird, kann hieraus der zu erwartende Anlagenzustand zum Startzeitpunkt extrapoliert werden. Nach der Initialisierung kennt das Anlagenmodell den Startzustand aus dem heraus Fahrpläne auf Realisierbarkeit hin überprüft werden sollen. In dem folgenden Ablauf (Abbildung 8.2) wird das Anlagenmodell von einer Samplingkomponente zur Erstellung des Samples benutzt. Die Erstellung eines neuen, zufälligen aber realisierbaren Fahrplans beginnt zunächst mit dem Zurücksetzen des Anlagenmodells auf den definierten Startzustand. Idealerweise kann dies durch Bereitstellung einer Resetfunktion durch das Anlagenmodell selbst geschehen, um die Parametrierung nicht wiederholt vornehmen zu müssen. Als nächstes wird eine zufällige Leistung (gleichverteilt aus  $[0, 1]$ ) für ein einzelnes Zeitintervall gewählt. Dies wird als Fahrplan der Länge 1 an das Anlagenmodell zur Validierung gegeben. Im Zuge dieser Validierung überprüft das Anlagenmodell die Realisierbarkeit eines Fahrplans mit dieser Leistungsstufe und berechnet im Erfolgsfall den durch Abfahren entstehenden Folgezustand aus dem heraus weiterevaluiert werden soll. Für den Fall, dass die Leistungsstufe aus dem aktuellen Anlagenzustand heraus nicht umgesetzt werden kann, verbleibt das Anlagenmodell im aktuellen Zustand und meldet den Fehlversuch an die Samplingkomponente zurück. Die Samplingkomponente generiert erneut zufällige Leistungsstufen bis entweder das Anlagenmodell eine Erfolgsmeldung gibt oder eine maximale Anzahl zulässiger Rateversuche je Zeitintervall überschritten wird. Letzteres führt zum Abbruch des aktuellen Fahrplans und nach einem Reset des Anlagenmodells zu einem Neubeginn des Vorgangs für den aktuellen Fahrplan. Im Erfolgsfall kann eine zufällige Leistung für das nächste Intervall gewählt und durch das Anlagenmodell evaluiert werden. Dieser Vorgang wiederholt sich für jedes Intervall im Planungshorizont bis ein Fahrplan hinreichender Länge entstanden ist, der tatsächlich von der Anlage umsetzbar ist. Nach einem Zurücksetzen des Anlagenmodells auf den Startzustand kann dann dieser Vorgang für den nächsten Fahrplan wiederholt werden bis genug Fahrpläne für das Sample geraten wurden. Im Idealfall ist also das Anlagenmodell nicht zustandslos, sondern kann direkt als Automat fungieren. Dies ist aber keine harte Anforderung, da es jederzeit möglich ist, einen Wrapper für das Anlagenmodell zu implementieren welcher diese Aufgabe übernimmt.

Betrachtet man das so erhaltene Sample geometrisch, so lässt sich folgendes feststellen: Das Sample umfasst  $n$  Punkte, die alle innerhalb einer durch die Constraints definierten geometrischen Figur liegen und in jeder Dimension gemäß derjenigen Verteilung verteilt sind, mit der auch die Zufallszahlen während des Samplings gezogen wurden. Dies gilt offensichtlich sowohl für das einfache Zufallssampling als auch für das sukzessive Sampling mit je einem Rateversuch.

Dies gilt aber nur, für den Fall, dass die Anzahl der Rateversuche für jede Dimension 1 ist. Unmittelbar einsichtig wird diese Eigenschaft, wenn man beide Verfahren aus Implementierungssicht betrachtet. Beim sukzessiven Sampling wird lediglich viel früher abgebrochen, wodurch deutlich weniger nicht realisierbare Fahrpläne betrachtet werden. Für den Fall,

## 8. SVDD als Blackbox-Modell

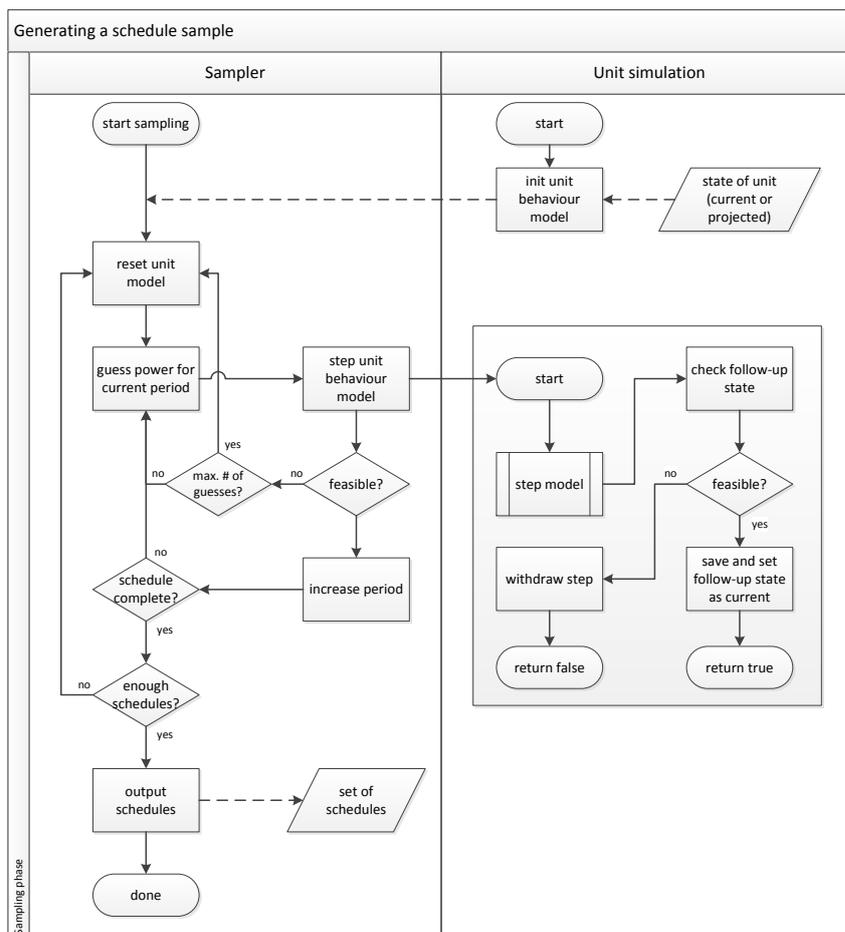


Abbildung 8.2.: Ablaufdiagramm für das sukzessive Sampling von Anlagensimulationsmodellen (Bremer & Sonnenschein 2013e).

dass mehrere Rateversuche je Dimension zugelassen sind, ist die Gleichverteilung des Samples innerhalb der Zielregion nicht mehr gegeben. Dies ist unmittelbar einsichtig. Da durch mehrere Rateversuche (in den Beispielen in diesem Abschnitt: 10 Versuche) die Wahrscheinlichkeit richtig zu liegen einer Bernoulli-Verteilung folgt. Dadurch sinkt die Wahrscheinlichkeit, kleine Bereiche zu finden unterproportional ab. Hierzu ein Beispiel: Im Falle eines einzigen Rateversuchs liegt das Verhältnis der Wahrscheinlichkeiten bei 5 : 1, falls 50% bzw. 10% des Leistungsspektrums zulässig sind. Bei beispielsweise 2 Rateversuchen liegt dieses Verhältnis – aufgrund der Bernoulli-Verteilung – bereits bei weniger als 4 : 1. Je höher die Zahl der Rateversuche, desto mehr nähert sich dieses Verhältnis an 1 : 1 an. Da bei hinreichender Zahl an Rateversuchen auch in schmalen Region beinahe genauso viele Punkte gefunden werden wie in breiten, ist hier die Dichte entsprechend höher.

Man kann dies in Abbildung 8.3(b) sehen. Ist eine Menge von Punkten innerhalb der Zielregion erst einmal bekannt, lässt sich deren Verteilung innerhalb der Region leicht verbessern. Dies kann beispielsweise mittels einer nachgeschalteten Diffusionssimulation oder z. B. mittels der Simulation einer Brown'schen Bewegung der einzelnen Punkte des

Samples innerhalb der durch die Zielregion vorgegebenen zulässigen Grenzen geschehen. Abbildung 8.3 zeigt ein solches Beispiel für eine artifizielle, bananenförmige Zielregion (artifizial begrenzt durch zwei gegeneinander verschobene Kreise). Das linke Sample wurde mit Hilfe des Rejection Sampling erstellt, das mittlere mittels sukzessivem Samplings. Im Nachgang, wurden die einzelnen Punkte aus dem mittleren Sample verschoben durch Simulation einer Brown'schen Bewegung (vgl. z. B. Binder & Braverman 2009), wobei ein eventuelles Herauslaufen aus der Zielregion mittel des Modells verhindert wurde. Bereits die simulierte Bewegung aus 20 Schritten reicht aus, das Sample aus der rechten Abbildung 8.3(c) aus dem mittleren Sample zu erzeugen. Da im weiteren Verlauf dieser Arbeit jedoch nur die räumliche Struktur interessiert, welche das Sample enthält, wird auf eine solche Nachbehandlung der Samples verzichtet. Die Struktur tritt auch beim sukzessiven Sampeln bereits deutlich hervor, sodass bereits dieses Sample als Trainingsmenge für das Lernen einer Hüllfläche ausreicht. Eine Verbesserung der Abbildung des Randbereichs könnte noch durch Erhöhung des Umfangs der Trainingsmenge erreicht werden. Man darf jedoch nicht vergessen, dass immer ein Bias bleiben wird, egal wie groß die Trainingsmenge ist.

Der große Vorteil des sukzessiven Samplingverfahrens liegt in seiner enormen Geschwindigkeitssteigerung. Beim sukzessiven Sampling sind für jedes Zeitintervall, für das eine mittlere Leistung geraten werden soll, mehrere Rateversuche zugelassen. Sei wieder  $\mathcal{P}$  die Wahrscheinlichkeit, dass eine für ein Zeitintervall zufällig gewählte Leistung realisierbar ist ohne Constraints zu verletzen. Dann gibt die Binomialverteilung  $B(i|\mathcal{P}, n) = \sum_{i=1}^n \binom{n}{i} \mathcal{P}^i (1 - \mathcal{P})^{n-i}$  für den Bernoulliprozess beim sukzessiven Raten die Wahrscheinlichkeit  $\mathcal{P}_{(n)}$  an, mit  $n$  Versuchen bei einer Erfolgsrate von  $\mathcal{P}$  genau  $i$  mal für ein Zeitintervall richtig zu liegen. Da jede Anzahl von Erfolgen zu einem Ergebnis führt, werden die Einzelwahrscheinlichkeiten aufsummiert. Mit den Wahrscheinlichkeiten für jedes Zeitintervall ergibt sich somit als Gesamtwahrscheinlichkeit für  $d$  Zeitintervalle

$$\mathcal{P}_{(n)}^d = \left( \sum_{i=1}^n B(i|\mathcal{P}, n) \right)^d. \quad (8.1)$$

Beispielsweise ergibt sich bei einer Erfolgsrate je Zeitintervall von  $\mathcal{P} = 0.05$  (vgl. Abbildung 8.1) bei 100 Rateversuchen eine Gesamterfolgsrate für 96-dimensionale Fahrpläne von ca. 0.5655.

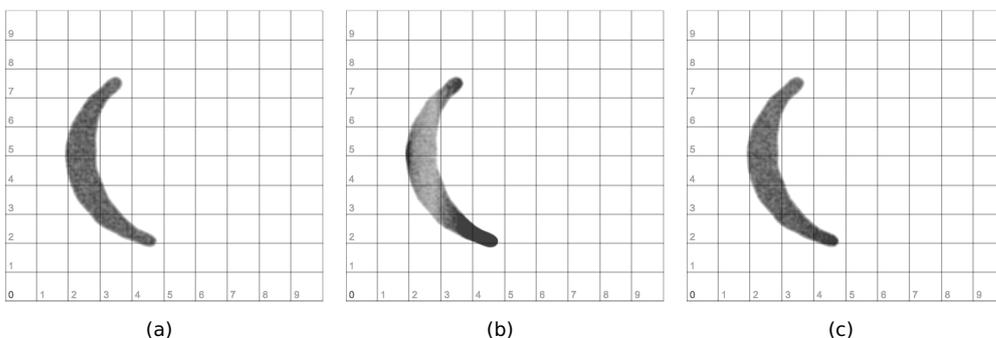


Abbildung 8.3.: Ausgleich der selektiven Stichprobe durch Brown'sche Bewegungen.

## 8. SVDD als Blackbox-Modell

---

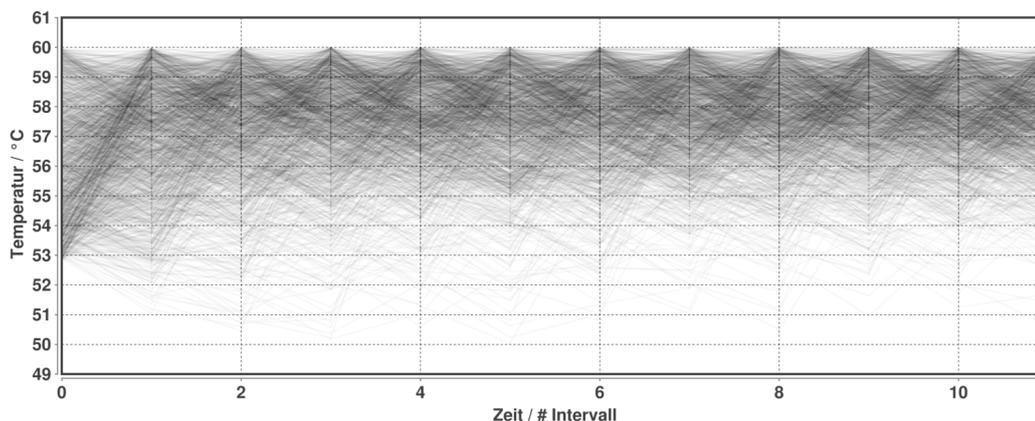


Abbildung 8.4.: Temperaturverläufe in einem Warmwasseruntertischgerät, gefahren nach Fahrplänen aus einem falsch erzeugten Sample. Dunklere Farben deuten mehrere Fahrpläne im entsprechenden Bereich an.

Ein weiterer wichtiger Punkt für die Erstellung guter Samples ist die Verwendung der richtigen Verteilung beim Generieren der Zufallswerte für mögliche Wirkleistungen. Ein Beispiel für die Auswirkungen ungeeigneter Verteilungen beim sukzessiven Sampling zeigt Abbildung 8.2. Dargestellt sind die Temperaturverläufe in einem Warmwassergerät wie sie entstehen, wenn es mit Fahrplänen aus einem Sample betrieben wird, das nach der sukzessiven Methode mit gleichverteilten Zufallszahlen erzeugt wurde. Durch die halbtransparente Farbgebung der einzelnen Temperaturverläufe deuten dunkle Bereiche das Vorhandensein von mehr Verläufen in dem entsprechenden Bereich an. Man sieht leicht, dass es ein Ungleichgewicht zugunsten von Fahrplänen gibt, die zu einer eher hohen Temperatur im Warmwassergerät führen. Da ein ungestörtes Warmwassergerät (ohne jedwede Warmwasserzapfung) simuliert wurde, wäre eigentlich eine eher gleichmäßige Verteilung der Temperaturen zu erwarten. Der Grund für dieses Ungleichgewicht ist in Abbildung 8.5 dargestellt.

Dargestellt sind die Dichten verschiedener Verteilungen von auftretenden Leistungswerten für Untertischwarmwassergeräte ausgehend von unterschiedlichen initialen Wassertemperaturen (äquivalent zu initialen Energieinhalten des Speichers) als Histogramme. Man sieht leicht, dass Leistungswerte im oberen noch zulässigen Bereich nicht mehr gleichverteilt auftreten, sondern mit zunehmender Leistung seltener werden. Bei Verwendung gleichverteilter Zufallszahlen, werden diese Werte also überproportional berücksichtigt. Die Lösung des Problems liegt in der Verwendung der korrekten Verteilungen bei der Erzeugung der Zufallswerte. Die Verteilung der auftretenden möglichen Leistungswerte ist einheitenspezifisch und offensichtlich auch abhängig vom Startzustand (vgl. Abbildung 8.5). Bei einigen Anlagentypen sollte jedoch eine Verteilung, welche über verschiedene Startzustände mittelt hinreichend genau sein, sodass diese einmalig vorab bestimmt werden kann. Eine Möglichkeit für die Abschätzung der unbekanntes Dichtefunktion ist die Verwendung eines Kerndichteschätzers, der im Nachgang leicht zum Generieren der Samples verwendet werden kann. Im Allgemeinen bezeichnet Dichteschätzung den Prozess der Konstruktion eines Schätzers für eine unbekanntes Wahrscheinlichkeitsdichte basierend auf

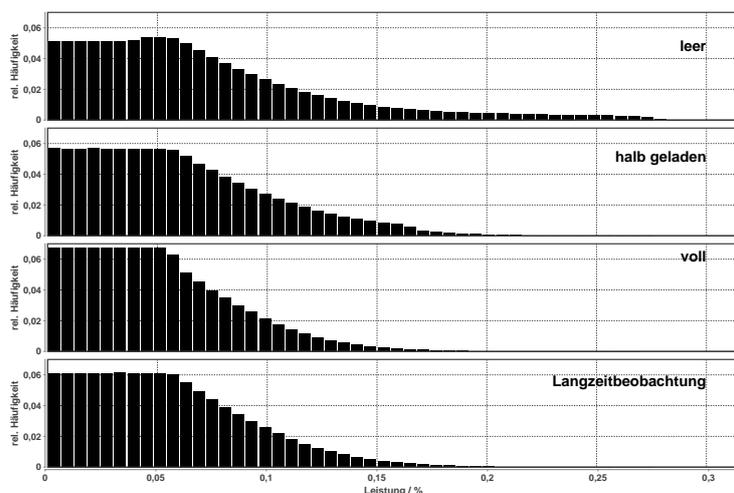


Abbildung 8.5:  
Histogramme empirischer  
Verteilungen gemessen an  
verschiedenen simulierten  
Untertischwarmwasser-  
geräten mit unterschied-  
lichem Energiefüllstand  
(repräsentiert durch die  
Temperatur).

einer Reihe von beobachteten Werten. Verschiedene Techniken sind hierfür bekannt. Für diskrete Schätzer wird häufig auf Histogramme zurückgegriffen. Eine seit langem bekannte und populäre Methode für kontinuierliche Dichtefunktionen ist der Kerndichteschätzer, auch bekannt als Parzen-Rosenblatt-Fenstermethode (Parzen 1962; Rosenblatt 1956).

Sei  $\mathcal{X} = x_1, \dots, x_n \in \mathbb{R}$  ein iid Sample von beobachteten Werten und  $K$  eine symmetrische (nicht notwendigerweise positive) Funktion mit dem Integral 0, dann ist

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (8.2)$$

ein Schätzer für die Dichte von  $\mathcal{X}$ . Als Kern  $K$  können stetige Lebesgue-Dichten verschiedener Wahrscheinlichkeitsmaße verwendet werden. Häufig kommen z. B. Gauß-, Cauchy- oder Epanechnikov-Kerne zum Einsatz (Bowman & Azzalini 1997). Der vorab zu wählende Parameter  $h$  gibt die Bandbreite (Glättungsfaktor) an, welcher maßgeblich die Qualität des Schätzers bestimmt. Eine bekannte Faustregel für die Wahl dieses Parameters stammt von Silverman (Silverman 1986). Als Vorbereitung für den Prozess des sukzessiven Samplings kann somit also zunächst ein Kerndichteschätzer mit einer durch das Anlagensimulationsmodell generierten Abfolge einzelner Leistungswerte trainiert werden. Während des eigentlichen Samplings kann dieser Schätzer dann zur Generierung geeigneter verteilter Zufallswerte genutzt werden. Im Falle eines Gaußkerns durch  $x \sim N(x_i, h)$  mit gleichverteilt zufällig gewähltem  $x_i$ . Es existieren auch Erweiterungen für den multivariaten Fall mit dem die Dichte ganzer Fahrpläne abgeschätzt werden könnte (Silverman 1986; Scott 1992; Rau 2003), allerdings würde dies zu den gleichen Komplexitätsproblemen wie beim Rejection Sampling führen.

Mit dem hier beschriebenen Verfahren des sukzessiven Samplings steht nunmehr eine Methode zur Verfügung, um aus einem gegebenen Anlagenmodell eine Trainingsmenge mit durch die Anlage realisierbaren Fahrplänen zu generieren. Die Anforderungen, die hierfür an das Anlagenmodell gestellt werden beschränken sich auf die Fähigkeit für einen bestimmten zukünftigen Zeitpunkt als Startzeitpunkt parametrisiert werden zu können sowie auf eine Funktion zur Überprüfung eines gegebenen Fahrplans auf Durchführbarkeit.

Mittels dieser Trainingsmenge soll im Folgenden nun ein SVDD-Klassifikator trainiert werden.

### 8.3. Modell

Das zuvor beschriebene Sampling eines gegebenen Simulationsmodells für eine dezentrale Anlage liefert eine Menge von Beispiellastgängen für realisierbare Fahrpläne der Anlage. Diese Beispielmenge realisierbarer Fahrpläne stellt als Punktwolke quasi eine Schablone des Unterraums (Lösungsraums) dar, den sie im Fahrplanraum einnimmt. Sie charakterisiert den zu modellierenden Lösungsraum geometrisch und enthält somit implizit eine Beschreibung aller Constraints welche den Fahrplanraum als Suchraum einschränken; genauer betrachtet beschreibt sie die Auswirkungen der Constraints. Aus dieser Punktmenge soll nun eine mathematische Beschreibung und somit ein Modell desjenigen Raums abgeleitet werden, den sie einnimmt.

Hierfür bietet sich das in Abschnitt 5.3 beschriebene Verfahren der Supportvektor Domain Description an. Wird das SVDD-Verfahren auf die Fahrplantrainingsmenge angewandt, so stehen im Ergebnis eine Teilmenge der Fahrpläne (die Supportfahrpläne) und ein Gewichtungsvektor für diese Fahrpläne zur Beschreibung des Lösungsraums zur Verfügung. Mit Hilfe der Abstandsfunktion (vgl. Gleichung 5.27 auf Seite 64) kann zunächst der Abstand eines Supportfahrplans (bzw. seines Bildes) zum Mittelpunkt der Hypersphäre bestimmt werden. Dieser dient dann als Radius und somit als Vergleichswert, um beliebige (auch nicht zur Trainingsmenge gehörende) Fahrpläne auf Zugehörigkeit zum Lösungsraum und somit auf Realisierbarkeit hin zu überprüfen. Wird für die Bestimmung der Supportfahrpläne ein Näherungsverfahren (eine ggf. zufallsbehaftete Heuristik) verwendet, so bietet es sich alternativ an, als Vergleichswert vorab den Erwartungswert der Abstände aller Supportfahrpläne zu bestimmen.

Es bringt Vorteile, für die Trainingsmenge skalierte Fahrpläne zu verwenden. Zum einen schmiegt sich die Hüllfläche besser an die Trainingspunkte an (Juszczak et al. 2002). Zum anderen können Anlagen unterschiedlicher Größe (unterschiedlicher Leistungsklassen) und Anlagen mit unterschiedlichem bzw. wechselndem Vorzeichen der Leistung (Prosumer) bei dem Verfahren bezüglich Parametrierung (z. B. die Bandbreite des Kerns) ähnlich behandelt werden. Ein Feintuning wird aufgrund verschiedener Lösungsraumstrukturen immer noch notwendig sein, jedoch ist der Wertebereich über den Parameterbelegungen gesucht werden müssen deutlich kleiner, sodass eine Automatisierung dieses Vorgangs deutlich einfacher wird. Durch die Skalierung müssen nur wenige zusätzliche Anlagenparameter für die spätere Rekonstruktion der ursprünglichen Fahrpläne in das Modell aufgenommen werden.

Für die Erstellung des Suchraummodells wird das Verfahren der SVDD verwendet, so wie es in Abschnitt 5.3 auf Seite 60 erläutert wurde. Als Kern wird ein Gauß-Kern verwendet und als Trainingsverfahren zur Bestimmung der Supportvektoren das SMO-Verfahren aus Abschnitt 5.3.3 auf Seite 68.

Abbildung 8.6 fast noch einmal das Vorgehen bei der Modellerstellung zusammen. Ausgehend von einer Trainingsmenge mit gültigen Fahrplänen wie sie beispielsweise

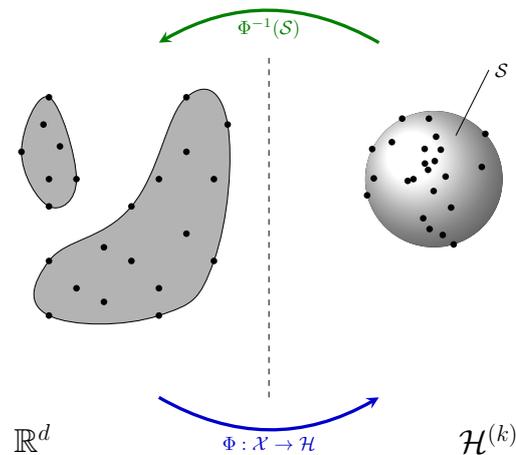


Abbildung 8.6.: Vorgehen bei der Erstellung des Modells mittels SVDD.

das Ergebnis des eingangs erläuterten Samplingverfahrens sind, wird eine Abbildung  $\Phi$  bestimmt, welche alle Fahrpläne in einen (nicht näher bekannten) hochdimensionalen Raum abbildet. Als Ergebnis eines Optimierungsprozesses wird diese Abbildung dergestalt bestimmt, dass die Menge der Bilder der Fahrpläne unter der Abbildung  $\Phi$  innerhalb einer Hypersphäre  $\mathcal{S}$  mit minimalem Durchmesser liegen. Im Ergebnis ist für die Beschreibung der Hypersphäre diejenige Teilmenge der Fahrpläne aus dem Trainingssatz ausreichend, deren Bilder genau auf der Oberfläche der Hypersphäre liegen. Das Urbild der Hypersphäre wiederum stellt nun ein Modell des Unterraums dar, der genau die gültigen Fahrpläne enthält. Für dieses Urbild ist jedoch keine explizite mathematische Beschreibung gegeben. Es muss indirekt aus einer Entscheidungsfunktion abgeleitet werden, welche direktes Ergebnis der SVDD ist.

### 8.3.1. Suchraummodell

Das auf SVDD basierende Suchraummodell besteht somit aus den folgenden Komponenten (Bremer et al. 2010; Bremer et al. 2011a):

1. Eine Menge  $SV = \{s_i \in \mathcal{X} \in [0, 1]^d \mid \beta_i \neq 0\}$  von Supportvektoren. Jeder Supportvektor  $s \in SV$  ist ein skaliertes Fahrplan aus der Trainingsmenge der Fahrpläne, die durch das Sampling aus dem Einheitensimulationsmodell erzeugt wurden.
2. Ein Gewichtungsvektor  $\mathbf{w} = (\beta_1, \dots, \beta_n) \quad \forall \beta_i \neq 0$ , der nur noch die Gewichtungen für die Supportvektoren enthält. Die Gewichte aller anderen Fahrpläne aus der Trainingsmenge sind 0, wessenthalb sowohl die zugehörigen Fahrpläne als auch deren Gewichte nicht benötigt werden.
3. Anlagenparameter, die zur Skalierung der Fahrpläne benötigt werden. Dies sind in der Regel die minimale und maximale Wirkleistung der Anlage.
4. Die Abstandsfunktion  $R^2(\mathbf{x}) = 1 - 2 \sum_i w_i k(\mathbf{s}_i, \mathbf{x}) + \sum_{i,j} w_i w_j k(\mathbf{s}_i, \mathbf{s}_j)$ , mit  $1 \leq i, j \leq |SV|$ , welche die Klassifikatorausgabe des Modells berechnet. Der Suchraum, der durch dieses Modell kodiert wird ist somit definiert als  $\Sigma = \{\mathbf{x} \mid R(\mathbf{x}) \leq R_S\}$ , mit

$R_S = \frac{1}{n} \sum_{s \in SV} R(s)$ . Kann bei der Berechnung der SVDD auf exakte Verfahren zurückgegriffen werden könnte statt des mittleren Abstands  $R_S$  einfach der Abstand eines beliebigen Supportvektors verwendet werden.

Das Supportvektormodell kapselt bereits alle Informationen bezüglich der räumlichen Verteilung der realisierbaren Lastgänge. Der eigentliche Raum der realisierbaren Lastgänge ist als Subraum eingebettet in den Raum aller Fahrpläne  $[P_{min}, P_{max}]$ .

Der auf Basis der realisierbaren Beispielfahrpläne trainierte Supportvektor-Klassifikator ist in der Lage, zu einem gegebenen Fahrplan zu entscheiden, ob dieser von der zugehörigen Anlage umsetzbar ist, oder nicht. Die Funktionsweise der Klassifikation ist in Abbildung 8.7 dargestellt. Ein gegebener Eingabefahrplan wird über den verwendeten Kern jeweils mit den verschiedenen Supportvektoren verglichen. Bei diesen Supportvektoren handelt es sich hier ebenfalls um Fahrpläne, die während des Trainings aus der Trainingsmenge als die am meisten relevanten Beispiele ausgewählt wurden. Diese sind nach dem Training für die Dauer der Gültigkeit des Suchraummodells statisch im Modell hinterlegt. Die Ergebnisse dieser Vergleiche gehen unterschiedlich gewichtet in die Funktion  $R$  zur Abstandsbestimmung ein. Einmalig vor der ersten Verwendung des trainierten Suchraummodells wird dieser Prozess auch für einen der Supportvektoren durchgeführt. In Abbildung 8.7 durch den grauen Pfad angedeutet. Auch hier findet analog der Vergleich mit allen anderen Supportvektoren statt. Das Ergebnis dieser speziellen Abstandsbestimmung (Radius  $R_S$  der Hyperkugel) kann nun ebenfalls für regelmäßige Wiederverwendung hinterlegt werden. Falls zur Bestimmung der Hyperkugel ein Näherungsverfahren (vgl. Abschnitt 5.2.2 auf Seite 57) zur beschleunigten Berechnung verwendet wurde, bietet es sich an, den Abstand für alle Supportvektoren zu bestimmen und als Radius den Mittelwert festzulegen. Als letzter Schritt der Klassifikation kann dann der zu einem beliebigen Eingabefahrplan bestimmte Abstand mit dem Radius  $R_S$  verglichen werden. Ein kleinerer (oder gleich großer) Abstand ( $R - R_S < 0$ ) kennzeichnet einen realisierbaren Fahrplan, ein größerer einen nicht realisierbaren.

Somit liegt ein Suchraummodell vor, welches die geometrische Region, die alle realisierbaren Fahrpläne enthält, anhand einer verhältnismäßig kleinen Auswahl aus der Menge der Trainingsfahrpläne beschreibt. Wie gut diese Beschreibung tatsächlich ist und welche Parameter des Verfahrens diese Güte wie beeinflussen ist im Kapitel 11 ab Seite 147 ausführlich diskutiert. Hier soll das Verfahren zunächst anhand von Illustrationen veranschaulicht werden.

### 8.3.2. Illustration

Eine Illustration der Klassifikation ist leider nur für niedrig-dimensionale Beispiele sinnvoll darstellbar.

Abbildung 8.8 zeigt am Beispiel eines BHKW (die beiden oberen Abbildungen) sowie eines Warmwasserboilers (unten) wie ein mit Hilfe einer SVDD erstelltes Modell für Fahrpläne mit Leistungswerten für zwei Zeitintervalle aussehen kann. Die roten Flächen in den beiden linken Abbildungen kennzeichnen das Ergebnis des Klassifikators. Alle Punkte für die der Klassifikator die Nichtzugehörigkeit zur Trainingsmenge ergibt, wurden rot eingefärbt. Der Klassifikator wurde mit Hilfe der durch blaue Punkte (in den jeweils linken

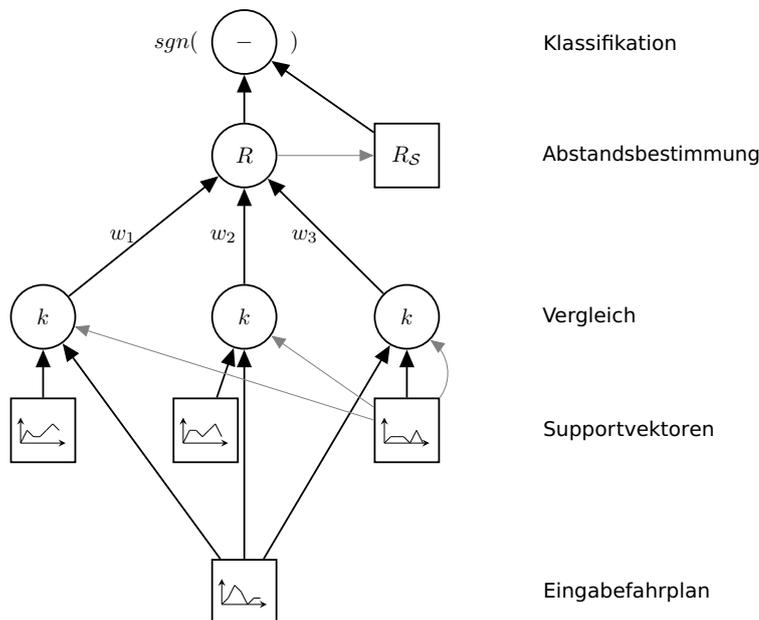


Abbildung 8.7.: Schematische Darstellung des Klassifikationsprozesses. Die Kernfunktion  $k$  ist a priori bekannt. Das übertragene Modell enthält die Menge der Supportvektoren und den Gewichtungsvektor  $w$ . Der Vergleichsabstand (Radius der Sphäre  $R_S$ ) wird anhand eines beliebigen Supportvektors einmal vorab berechnet (grauer Pfad). Darstellung nach einer Idee aus (Schölkopf 1997), auf SVDD angepasst.

Bildern) dargestellten Trainingsmenge erstellt. Die orangefarbene Linie kennzeichnet die Entscheidungsgrenze  $\{x \mid R(x) = R_S\}$  zwischen Zugehörigkeit zur Trainingsmenge und Nichtzugehörigkeit. Die beiden Bilder auf der rechten Seite verdeutlichen genauer die unterschiedlichen Entfernungen der jeweiligen Bilder der Fahrpläne (unter der Abbildung  $\Phi$ ) vom Rand der Hypersphäre. Die blaue Fläche stellt den Bereich außerhalb der Hypersphäre dar: je weiter außen, desto dunkler das Blau. Der olivfarbene Teil kennzeichnet das Urbild des inneren Teils der Hypersphäre (die orangefarbene Linie entspricht wieder der Oberfläche). Dunklere Olivtöne signalisieren, dass das Bild dieser Teile weiter im Inneren der Hypersphäre liegt.

Man erkennt auch deutlich, dass das Verfahren am Rand keine gerade Begrenzungslinie liefert, sondern aufgrund der natürlichen Unvollständigkeit der Trainingsmenge eine Näherung darstellt sowie aufgrund des verwendeten Gaußkerns Ecken eher rund abbildet. Wie in dem unteren Beispiel ersichtlich, ist es auch möglich, dass innerhalb des Bereichs in dem eigentlich ausschließlich zulässige Fahrpläne liegen, Löcher entstehen, falls die Trainingsmenge an der entsprechenden Stelle zu dünn besiedelt ist. In gewissen Grenzen lassen sich derartige Unzulänglichkeiten durch geeignete Parameterbelegung aber noch abfangen. Ein Beispiel, wie solche Lücken ansonsten gehoben werden können zeigt Abbildung 8.9. Hier wurde die Entscheidungsfunktion dadurch modifiziert, dass der Radius der Hypersphäre marginal größer angenommen wird, als er wirklich ist. Dadurch werden Punkte nahe am Rand noch als zulässige Fahrpläne klassifiziert, obwohl sie eigentlich nicht zulässig sind. Hierdurch können die Lücken im inneren beseitigt werden zum Preis

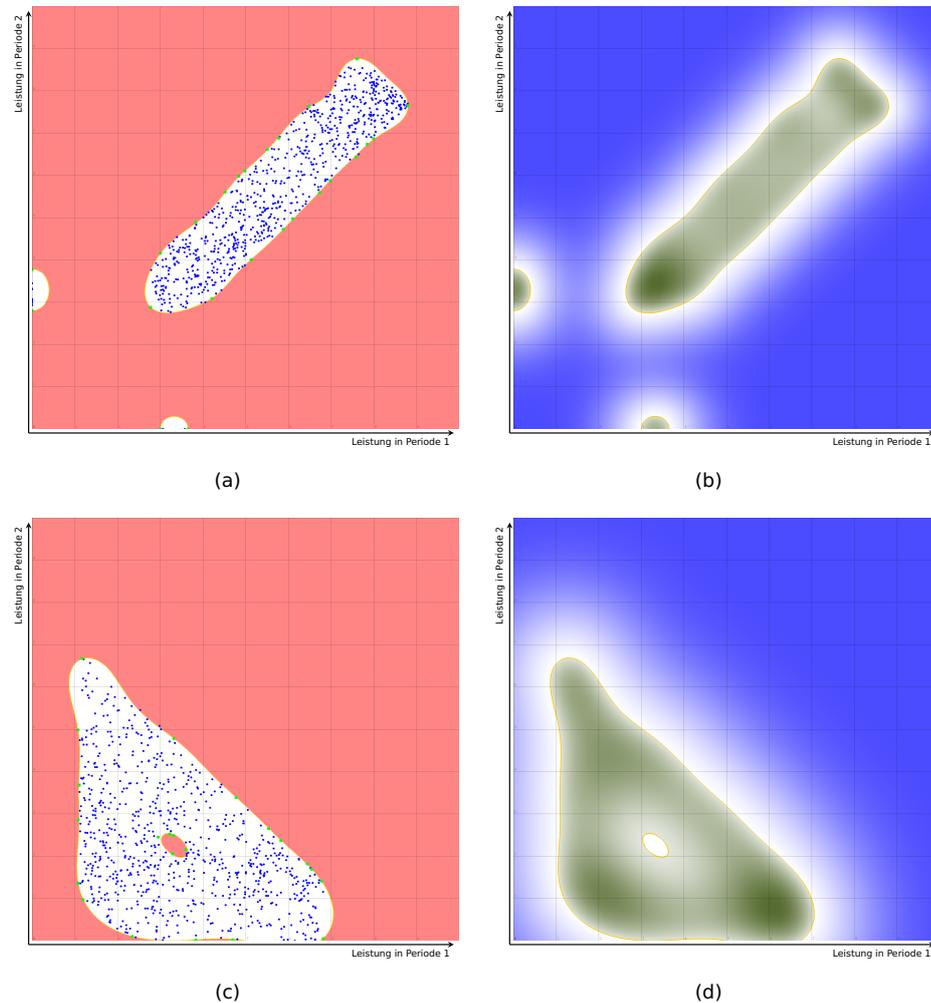


Abbildung 8.8.: Illustratives Beispiel zweier Suchraummodelle. In den oberen Abbildungen 8.8(a) und 8.8(b) ist der 2-dimensionale Suchraum eines BHKW dargestellt, unten ( 8.8(c) und 8.8(d)) wurde ein Warmwasserbereiter visualisiert.

von wenigen Fehlklassifikationen am Rand. In dem Beispiel in Abbildung 8.9 beträgt der Fehler am Rand schlimmstenfalls (bei 0.2%o Radiusenerweiterung) ca. 0.75 Watt bei einer Geräteenennleistung von 800 Watt.

### 8.3.3. Einbettung und Anwendbarkeit

Mit dem Suchraummodell in seiner jetzigen Form sind zwei Anwendungsfälle möglich:

1. Überprüfung eines gegebenen Fahrplans auf Gültigkeit. Die eigentliche Anlagenmodellierung wird hierfür nicht mehr benötigt.
2. Vergleich zweier gegebener Fahrpläne. Anhand der Abstandsfunktion ist es möglich

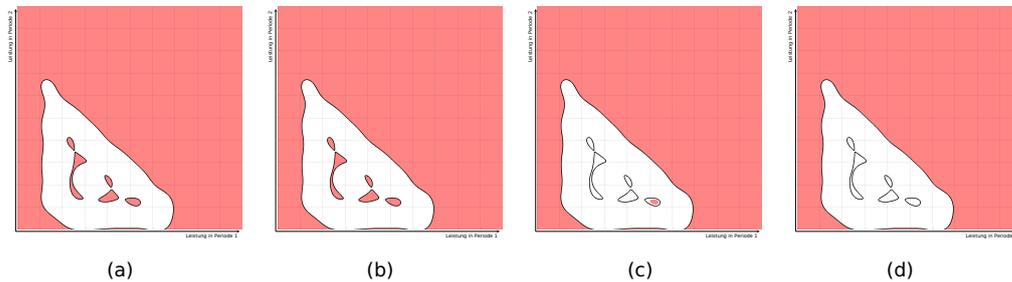


Abbildung 8.9.: Beispiel zur möglichen Reparatur von Löchern im Suchraummodell. Der Toleranzwert für die Klassifikation beträgt von links nach rechts 0, 0.05‰, 0.1‰ bzw. 0.2‰ des Kugelradius.

zu entscheiden welcher von zwei gegebenen, ungültigen Fahrplänen näher an der Region der gültigen Fahrpläne gelegen ist.

Somit könnten im Rahmen eines Planungsalgorithmus Fahrpläne als Lösungskandidaten anderweitig generiert werden, während das Modell zur Bewertung der Gültigkeit einer Lösung eingesetzt wird. Die Abstandsfunktion könnte beispielsweise auch als externer Penaltywert verwendet werden, um ein evolutionäres Planungsverfahren in Richtung gültiger Lösungen zu leiten. Problematisch ist in beiden Fällen, dass an der Lösungsfindung im Rahmen einer ganzen Gruppe von Anlagen immer auch mehrere Suchraummodelle (eins je Anlage) beteiligt sind.

Penalty-Ansätze für Mehrzielprobleme wurden in der Literatur bereits vorgeschlagen, beschränken sich in ihrer Evaluation allerdings meist auf eine überschaubare Anzahl von Penalties (z. B. (Angantyr et al. 2003)) oder aggregieren zunächst alle Nebenbedingungen gewichtet zu einer einzigen, um diese mit einem einzigen Penalty behandeln zu können (z. B. (Meng et al. 2011)). Im Falle des Wirkleistungsplanungsproblems beziehen sich die einzelnen Penalties auf gegebenenfalls verteilt gehandhabte separate Suchräume verschiedener Anlagen, sodass eine Zusammenfassung nicht möglich ist. Eine solche Zusammenfassung von Penalties würde auch immer eine individuelle Anpassung an die jeweilige Problem Instanz (aufgrund der individuellen Verbundzusammensetzung) bedeuten, welche ja eine Automatisierung der Wirkleistungsplanung wieder erschweren würde.

Zudem tritt bei hochdimensionalen Fahrplänen das Problem auf, dass weiter außen gelegene Punkte anhand der Abstandsfunktion kaum zu unterscheiden sind, sodass ein Penalty-Ansatz hier keine Richtung feststellen könnte. Andere Constraint-Handling Techniken (vgl. Abschnitt 4.5 auf Seite 42) würden sicher ebenso einer wiederholten, jeweils spezifischen Anpassung an die Problem Instanz bedürfen.

Der umgekehrte Ansatz eines internen Penalty in der Lösungsbewertung würde zwar dafür sorgen, dass immer mit einer gültigen Lösung (eine gültige Startlösung ist allerdings erst einmal zu finden) gearbeitet wird, andererseits verhindert dieser Ansatz auch das Durchsuchen nicht zusammenhängender Teilbereiche der Lösungsräume.

Das Modell, so wie es in diesem Kapitel bisher beschrieben wurde, ermöglicht zunächst die Identifizierung realisierbarer Fahrpläne ohne hierfür das eigentliche, zugrunde liegende mathematische Anlagenmodell kennen zu müssen. Wie auch Abbildung 8.10 illustriert

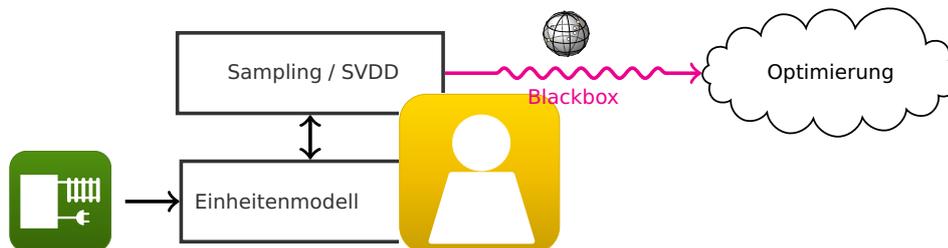


Abbildung 8.10.: Vorläufiges rein Klassifikator-basiertes Integrationsmodell zur Kommunikation individueller Möglichkeiten ohne eine konkrete Modellierung preisgeben zu müssen.

könnte das Modell in Szenarien mit zentraler Kontrollstelle bereits zur Kommunikation der individuellen Suchräume an eine Kontrollinstanz dienen (vgl. auch Bremer et al. 2010).

### 8.4. Zusammenfassung

Mit dem SVDD-basierten Modell eines Suchraums (und dem Sampling als Voraussetzung zur Erstellung) steht nunmehr ein Mittel zur Verfügung, um beliebige Einheitensimulationsmodelle in ein einheitliches Darstellungsformat bezüglich ihres Suchraums und seiner Einschränkungen zu bringen. Einzige Voraussetzung für die Konstruktion ist eine Menge von Beispielfahrplänen als Trainingsmenge. Für Einheitensimulationsmodelle im Smart Grid Umfeld darf wohl angenommen werden, dass die Erstellung solcher Trainingsdaten immer möglich ist, da die Simulation lediglich in der Lage sein muss, einen gegebenen Leistungswert auf Zulässigkeit zu prüfen. Die weitere Funktionalität des Sampling kann leicht ausgelagert und separat implementiert werden. Als Funktionalität bietet das Modell soweit die Möglichkeit einen beliebigen, gegebenen Fahrplan daraufhin zu überprüfen, ob er für die Einheit für die das Suchraummodell erstellt wurde, umsetzbar ist oder nicht. Durch Vergleich der Abstände vom Mittelpunkt der Hypersphäre ist darüber hinaus noch möglich zu zwei gegebenen Fahrplänen zu entscheiden welcher von beiden näher an der Umsetzbarkeit ist.

Wie gezeigt werden konnte, ist das Suchraummodell allein noch nicht hinreichend für eine Integration in Planungsverfahren. Aufgrund der Struktur der Problemstellung müssen immer eine Vielzahl von Suchraummodellen für die Lösungsfindung gleichzeitig betrachtet werden, wobei jedes einzelne nur eine Aussage über jeweils eine Anlage erlaubt. Durch Einbinden dieser Vielzahl von Suchraummodellen über einen Penalty-Ansatz werden offenbar entsprechend viele lokal Optima induziert, sodass ein Suchverfahren welches allein auf den Penalties fußt inakzeptabel häufig in zumindest teilweise ungünstigen Lösungen steckenbleibt (Bremer & Sonnenschein 2013b).

In jedem Fall ist es bisher nur mithilfe des Modells eine Aussage über einen bereits gegebenen Fahrplan zu machen. Die gezielte Konstruktion gültiger Fahrpläne ist so noch nicht möglich. Dies wird die Aufgabe des Dekoder sein, dessen Konstruktion im nächsten Kapitel hergeleitet wird.

## 9. Supportvektor-Dekoderansatz

*A discerning eye needs only a hint, and understatement leaves the imagination free to build its own elaborations.*

*(Russell Page)*

### 9.1. Einleitung

Als letzte Komponente der Prozesskette wird in diesem Kapitel ein Dekoder entwickelt, der in der Lage ist, systematisch korrekte Lösungen zu produzieren. In dem in dieser Arbeit betrachteten Wirkleistungsplanungsproblem ist dies gleichbedeutend mit der systematischen Erzeugung von durch die verschiedenen Anlagen realisierbaren Fahrplänen.

Ziel wird es sein, die Nachbarschaftsbeziehung eines beliebigen Optimierungsansatzes so definieren zu können, dass keinerlei Constraints beachtet werden müssen. Hierzu wird die Nachbarschaftsbeziehung zwischen Lösungen in einem freien (unbeschränkten) Suchraum definiert. Ein so genannter Dekoder wird eine Beziehung zwischen dieser Lösung und einer gültigen Lösung aus den jeweiligen Suchräumen der Anlagen herstellen. Hierdurch wird das Wirkleistungsplanungsproblem überführt in eine constraintfreie Formulierung. Das Optimierungsverfahren braucht keinerlei Kenntnis mehr über die ursprüngliche mathematische Formulierung eines Anlagenmodells oder des hieraus resultierenden Suchraums. Gleiches gilt für die Beschreibung der Constraints.

Nach einer Skizzierung der generellen Konstruktionsidee werden nachfolgend die einzelnen Schritte formalisiert. Anschließend wird die Funktionsweise anhand von Illustrationen verdeutlicht. Die Evaluation erfolgt zusammen mit dem Modell in Kapitel 11 ab Seite 147; die Einbettung in den Anwendungsfall wird anschließend in Kapitel 12 ab Seite 177 und Kapitel 13 ab Seite 199 betrachtet.

### 9.2. Problemstellung und Lösungsidee

Wie bereits im vorigen Kapitel erläutert, liegt die Schwierigkeit bei der direkten Verwendung des Modells in einem Standard-Penaltyansatz zur Constrainthandhabung in der Tatsache, dass beispielsweise nicht ein einzelner Penaltywert für ungültige Lösungen verwendet werden kann, sondern verschiedene Strafwerte zu einem einzelnen Wert zusammengefasst werden müssen. Die individuellen Penaltywerte beziehen sich aber auf individuelle Teile der Lösung (jeweils den Beitrag einer Anlage zur Lösung des gemeinsamen Planungsproblems) und geben für verschiedene Teile der Lösung verschiedene Richtungen der Gradienten (in Richtung Gültigkeit der Lösung) vor. Dies kann durch einen

einzelnen Penaltywert nicht adäquat reflektiert werden. Methoden, welche Constraints als zusätzliche Optimierungsziele behandeln, hätten hiermit ähnliche Probleme. Als Alternative wäre eine Barriere-Methode möglich bei der das Verlassen des gültigen Bereiches bestraft wird. Die hierzu zwingend notwendig gültige, initiale Lösung könnte relative leicht gefunden werden indem für jede Anlage eine beliebige Startlösung solange in Richtung des durch die jeweilige Radiusfunktion vorgegebenen Gradienten iteriert wird, bis der zulässige Bereich erreicht ist. Ein solcher Ansatz käme allerdings nicht mit nicht zusammenhängenden Teillösungsräumen zurende. Wesentlich eleganter (und auch universeller im Einsatz) wäre es, wenn zulässige Lösungen einfach systematisch konstruiert werden können.

Für diese Aufgabe wird in diesem Kapitel ein Dekoder entwickelt. Wie bereits in Abschnitt 4.5.4 auf Seite 48 erläutert hat ein Dekoder die Aufgabe, eine Beziehung zwischen einer gültigen Lösung und der Dekoderlösung herzustellen. Ein Dekoder stellt – umgangssprachlich formuliert – eine Konstruktionsvorschrift für gültige Lösungen dar. Durch die Verwendung eines Dekoders muss ein Optimierungsverfahren damit auch nur gültige Lösungen verarbeiten können (Koziel & Michalewicz 1999). Üblicherweise wird der gewünschte Zusammenhang durch die Abbildung des ursprünglichen Optimierungsproblems auf eine andere, einfacher zu handhabenden Funktion erreicht (Coello Coello 2002). Beispielsweise nutzen Koziel & Michalewicz (1998) eine homomorphe Abbildung zwischen dem Einheitshyperwürfel und dem Lösungsraum, die ausgehend von einem Referenzpunkt im Inneren des Lösungsraums Strahlen (Absuchen des Raums in Polarkoordinaten) abbildet durch Skalierung des Suchstrahls anhand der maximal entfernten Randpunktes des Lösungsbereichs in der jeweiligen Richtung (Koziel & Michalewicz 1999). Dieses Verfahren bedarf allerdings zusätzlicher Berechnungen bei jeder Lösungsvariation für die Suche nach dem jeweils entferntesten Punkt des Lösungsraums sowie einer mathematischen Beschreibung desselben. Dennoch lassen die Ergebnisse die Verfahrensidee als einen vielversprechenden Ansatz erscheinen (Coello Coello 2002). Ältere Ansätze wie z. B. (Kim 1998) verwenden Riemann'sche Abbildungen, sind aber nur für niedrigdimensional Probleme geeignet und bedürfen der Zielfunktion in algebraischer Form (Coello Coello 2002), d.h. die Abbildung muss problemspezifisch algebraisch abgeleitet werden. Ein solches Vorgehen ist für das dynamische Wirkleistungsplanungsproblem nicht durchführbar.

Daher soll hier ein anderer Weg gegangen werden. Ausgehend von dem Supportvektormodell soll der geometrische Zusammenhang im RKHS  $\mathcal{H}$  zur Konstruktion eines Dekoders genutzt werden. Abbildung 9.1 auf der nächsten Seite zeigt die zugrundeliegende Idee.

Durch das Anwenden der SVDD entsteht die Beschreibung eines Unterraums der Fahrpläne, welcher nur noch die realisierbaren Fahrpläne enthält (die graue Fläche in der linken Hälfte der Abbildung). Diesem Unterraum entspricht eine hochdimensionale Kugel: das Bild des Unterraums (graue Kugel in der rechten Hälfte der Abbildung). Die Abbildung, welche das kugelförmige Bild des Lösungsraums erzeugt ist unbekannt, jedoch existieren Verfahren, diese durch eine empirische Abbildung anzunähern. Mit einer solchen Abbildungsvorschrift kann aber jeder Punkt (Fahrplan in Leistungswert-Vektordarstellung) in den Raum abgebildet werden, in dem auch die Kugel als Bild des Lösungsraums existiert. Hierbei gilt: gültige Fahrpläne (schwarz in Abbildung 9.1) werden in die Kugel oder auf deren Oberfläche abgebildet, ungültige Fahrpläne landen außerhalb der Kugel. In letzterem Fall sind aber folgende geometrische Zusammenhänge bekannt: das Bild des ungültigen

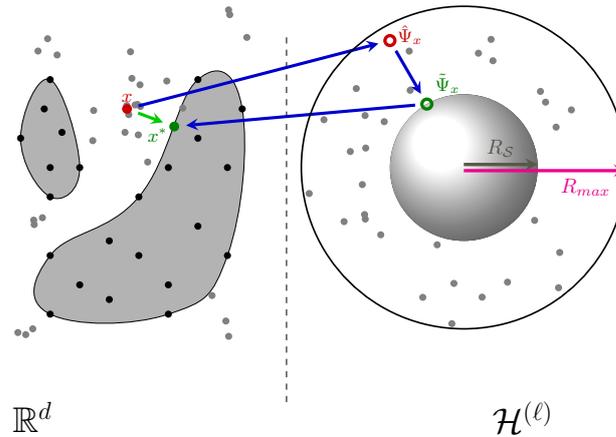


Abbildung 9.1.: Prinzipskizze des Supportvektor-Dekoderansatzes.

Fahrplans, der Mittelpunkt der Kugel sowie der Radius der Kugel. Somit kann ein Differenzvektor konstruiert werden, der vom Mittelpunkt der Kugel zum Bild des ungültigen Fahrplans zeigt. Anhand des Radius  $R_S$  kann dann seine Länge skaliert werden, sodass er auf einen Punkt auf der Kugeloberfläche zeigt. Hierdurch wird ein Punkt auf der Kugeloberfläche gefunden, der gerade noch das Bild eines gültigen Fahrplans darstellt. Findet man in einem letzten Schritt das Urbild dieses neu gefundenen Punktes, so erhält man einen gültigen Fahrplan am Rand des Lösungsraums. Würde man den größtmöglichen Abstand  $R_{max}$  eines beliebigen Bildes vom Kugelmittelpunkt kennen, so könnte die Skalierung des Differenzvektors auch so vorgenommen werden, dass ungültige Fahrpläne auch in das Kugellinnere hinein abgebildet werden (und zwar umso stärker, je näher sie bereits dran waren), damit nicht alle außenliegenden nur auf die Oberfläche abgebildet werden. Auch bereits innen liegende Punkte müssten hierbei weiter nach innen gerückt werden. Auf diese Weise kann es möglich werden, den Raum aller Fahrpläne in  $\mathbb{R}^d$  abzubilden auf das Urbild der Kugel in  $\mathcal{H}$ .

Diese einführende Darstellung hat die tatsächlichen Verhältnisse vereinfacht wiedergegeben. In der Praxis müssen Einschränkungen, die zwangsläufig zu Näherungslösungen führen werden hingenommen werden. Im Detail wird das Verfahren in den nachfolgenden Abschnitten erläutert.

### 9.3. Dekoderabbildung auf Basis der empirischen Kernabbildung

Für die Umsetzung des Dekoders soll die folgende Abbildung verwendet werden:

$$\begin{aligned} \gamma : [0, 1]^d &\rightarrow \mathcal{F}_{[0,1]} \subseteq [0, 1]^d \\ \mathbf{x} &\mapsto \gamma(\mathbf{x}), \end{aligned} \tag{9.1}$$

welche den Einheitshyperwürfel  $[0, 1]^d$  auf die Region der realisierbaren Fahrpläne (Feasible Region  $\mathcal{F}_{[0,1]}$ ) abbildet.

Konzeptionell werden für die Konstruktion wie im vorigen Abschnitt erläutert die bekannten geometrischen Zusammenhänge um die hochdimensionale Sphäre mit den Bildern der Fahrpläne ausgenutzt. Hierzu müssen drei Schritte bewerkstelligt werden:

1. Ein gegebener (ungültiger) Fahrplan  $x$  muss abgebildet werden auf  $\Phi(x)$ . Da  $\Phi$  nicht bekannt ist (vgl. Abschnitt 5.3 auf Seite 60), muss eine alternative Abbildung  $\hat{\Phi}_\ell$  als Näherung verwendet werden (Schölkopf et al. 1999a).
2. Das Bild von  $x$  muss so verändert werden, dass es das Bild eines gültigen Fahrplans darstellt. Hierzu muss es mit einer geeigneten Abbildung  $\Gamma_a$  in Richtung des Mittelpunktes  $a$  der Sphäre verschoben werden (vgl. Park et al. 2007).
3. Schließlich muss von dem verschobenen Bild das Urbild gefunden werden, welches den nunmehr gültigen Fahrplan repräsentiert. Dies ist nicht direkt, sondern nur mittels einer Näherung  $\hat{\Phi}_\ell^{-1}$  möglich (vgl. z. B. (Kwok & Tsang 2004)).

Die gesuchte Abbildung wird somit schrittweise als Komposition dreier einzelner Funktionen konstruiert:

$$\gamma = \hat{\Phi}_\ell^{-1} \circ \Gamma_a \circ \hat{\Phi}_\ell. \quad (9.2)$$

Anstatt also eine direkte Abbildung nach  $\mathcal{F}_{[0,1]}$  zu suchen, wird ein Umweg über den durch den Kern  $k$  induzierten RKHS eingeschlagen. Das Kommutativdiagramm 9.3 verdeutlicht noch einmal den eingeschlagenen Weg und den Zusammenhang zwischen den beteiligten (Unter-)Räumen. Ein beliebiger Vektor  $x \in [0, 1]^d$  (im Anwendungsfall ein skaliertes Fahrplan) wird zunächst abgebildet in den Raum  $\mathcal{H}^{(\ell)}$ , der durch die  $\ell$  Supportvektoren aufgespannt wird. Das Bild  $\hat{\Psi}_x$  von  $x$  wird anschließend in  $\mathcal{H}^{(\ell)}$  verschoben in Richtung Kugelmittle. Als letzter Schritt wird das Urbild  $x^*$  des modifizierten Bildes  $\tilde{\Psi}_x$  bestimmt um einen Vektor aus der zulässigen Region  $\mathcal{F}_{[0,1]}$  zu erhalten. Die gesuchte Funktion  $\gamma$  ergibt sich schließlich durch Hintereinanderausführung dieser drei Schritte.

$$\begin{array}{ccc} x \in [0, 1]^d & \xrightarrow{\hat{\Phi}_\ell} & \hat{\Psi}_x \in \mathcal{H}^{(\ell)} \\ \downarrow \gamma & & \downarrow \Gamma_a \\ x^* \in \mathcal{F}_{[0,1]} \subseteq [0, 1]^d & \xleftarrow{\hat{\Phi}_\ell^{-1}} & \tilde{\Psi}_x \in \mathcal{H}^{(\ell)} \end{array} \quad (9.3)$$

Der im Rahmen dieser Arbeit für die SVDD verwendete Gaußkern hat die Eigenschaft stationär zu sein. Stationäre Kerne sind solche Kerne, die nur vom Differenzvektor zweier Punkte abhängen. Sie haben damit auch die Eigenschaft invariant gegenüber Translationen zu sein (Genton 2002). Durch die Verwendung eines stationären Kerns ist es möglich, eine geschlossene Lösung zu erhalten (Kwok & Tsang 2004). Im Folgenden werden die einzelnen Schritte näher erläutert.

### 9.3.1. Voraussetzungen

Ausgangspunkt ist das Supportvektor-Modell, welches mithilfe der SVDD den Suchraum beschreibt. An dieser Stelle sei davon ausgegangen, dass dieser Schritt wie im vorigen

Kapitel erläutert bereits durchgeführt wurde und somit die folgenden Informationen bereits vorliegen:

- Eine Menge von Supportvektoren  $\{\mathbf{s}_1, \dots, \mathbf{s}_\ell\} \in \mathbb{R}^d$ , welche eine Teilmenge skaliertes Fahrpläne darstellen.
- Ein Gewichtungsvektor  $\mathbf{w} = \{\beta_1, \dots, \beta_\ell\} = \{\beta | \beta \neq 0\}$ , welcher für jeden Supportvektor die Gewichtung vorgibt. Dieser Gewichtungsvektor stellt eine reduzierte Version der Lagrangemultiplikatoren dar, welcher nur noch die relevanten Werte größer 0 enthält.

Somit lässt sich die Darstellung des Bildes eines Fahrplans  $\Phi(\mathbf{x}) \in \mathcal{H}$ , welches durch die Expansion  $\Psi_x = \sum_{i=1}^n \beta_i \phi(\mathbf{x}_i)$  repräsentiert wird vereinfachen zu  $\Psi_x = \sum_{i=1}^{\ell} \mathbf{w}_i \phi(\mathbf{s}_i)$ , wobei  $\mathbf{w}$  nur noch die Gewichtungswerte größer 0 enthält. Die Zahl der Supportvektoren  $\ell$  ist deutlich kleiner als die Anzahl  $n$  der Fahrpläne in der Trainingsmenge.

Dies sind die einzigen Informationen, die für die Konstruktion des Dekoders benötigt werden.

### 9.3.2. Schritt 1: Abbildung in den Supportvektor RKHS $\mathcal{H}^{(\ell)}$ mittels empirischer Kernabbildung

Das Training der Supportvektor-Beschreibung der Daten findet in der Praxis zwangsläufig immer auf einer endlichen Menge von Trainingsdaten statt. Alle Supportvektor-Algorithmen arbeiten daher in einem Unterraum des Feature-raums, der durch die Bilder der Trainingsvektoren aufgespannt wird. Dieser Unterraum kann in einen Euklidischen Raum abgebildet werden, sodass paarweise Abstände und Winkel zwischen den Vektoren erhalten bleiben. Diese Einbettung kann durch die so genannte empirische Kernabbildung erfolgen (Schölkopf et al. 1999a; Xiong et al. 2005). Die geometrische Struktur der Trainingsdaten ist im empirischen Feature-raum die gleiche wie im unendlichdimensionalen (durch  $\Phi$  aufgespannten) Feature-raum (Xiong et al. 2005). Analog kann der Unterraum betrachtet werden, der durch die Bilder der Supportvektoren aufgespannt wird.

Sei nun

$$\begin{aligned} \Phi_\ell : \mathbb{R}^d &\rightarrow \mathbb{R}^\ell, \\ \mathbf{x} &\mapsto k(\cdot, \mathbf{x})|_{\{\mathbf{s}_1, \dots, \mathbf{s}_\ell\}} \\ &= (k(\mathbf{s}_1, \mathbf{x}), \dots, k(\mathbf{s}_\ell, \mathbf{x})) \end{aligned} \quad (9.4)$$

die empirische Kernabbildung bezogen auf die Menge der Supportvektoren  $\{\mathbf{s}_1, \dots, \mathbf{s}_\ell\}$ . Wird  $\Phi_\ell$  mittels der Kern Gram-Matrix  $K_{ij} = k(\mathbf{s}_i, \mathbf{s}_j)$  modifiziert zu

$$\begin{aligned} \hat{\Phi}_\ell : \mathbb{R}^d &\rightarrow \mathcal{H}^{(\ell)} \\ \mathbf{x} &\mapsto K^{-\frac{1}{2}}(k(\mathbf{s}_1, \mathbf{x}), \dots, k(\mathbf{s}_\ell, \mathbf{x})), \end{aligned} \quad (9.5)$$

so bildet diese Abbildung (9.5) beliebige Punkte  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$  so nach  $\mathcal{H}^{(\ell)}$  ab, dass gilt (vgl. Schölkopf et al. 1999a):

$$k(\mathbf{x}, \mathbf{y}) = \hat{\Phi}_\ell(\mathbf{x}) \cdot \hat{\Phi}_\ell(\mathbf{y}). \quad (9.6)$$

Somit könnten auch hier alle durch Skalarproduktbildung ausdrückbaren Operationen wieder über die Kernfunktion im Eingaberaum ausgeführt werden. Die Kern Gram-Matrix,

die Matrix aus den Skalarprodukten der unter  $\Phi$  abgebildeten Vektoren,  $\mathbf{K}$  ist mit  $\mathbf{K}_{ij} = k(\mathbf{s}_i, \mathbf{s}_j)$  leicht zu bestimmen, da  $k$  bekannt ist. Die inverse Wurzel  $\mathbf{K}^{-\frac{1}{2}}$  der Matrix wird über eine Spektraldekomposition von  $\mathbf{K}$  bestimmt.

Seien  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$  die Eigenwerte von  $\mathbf{K}$  und  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$  die zugehörigen Eigenvektoren, dann lässt sich die inverse Wurzel von  $\mathbf{K}$  bestimmen als

$$\mathbf{K}^{-\frac{1}{2}} = \sum_{i=0}^k \frac{1}{\sqrt{\lambda_i}} \mathbf{u}_i \mathbf{u}_i^\top.$$

Dies ist möglich, da  $\mathbf{K}$  symmetrisch ist und somit alle Eigenwerte reell sind.

*Beweis.*  $\mathbf{K}_{ij} = k(\mathbf{x}_i, \mathbf{x}_j) = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle = \langle \Phi(\mathbf{x}_j), \Phi(\mathbf{x}_i) \rangle = k(\mathbf{x}_j, \mathbf{x}_i) = \mathbf{K}_{ji}$  □

Die Berechnung von  $\mathbf{K}^{-\frac{1}{2}}$  muss nur einmal vorab durchgeführt werden.

Mit  $\hat{\Phi}_\ell$  ist es nunmehr einfach möglich, beliebige Punkte (insbesondere aus  $[0, 1]^d \subseteq \mathbb{R}^d$ ) in einen  $\ell$ -dimensionalen Raum  $\mathcal{H}^{(\ell)}$  abzubilden, der eine  $\ell$ -dimensionale Projektion der Hypersphäre enthält. Auch hier gilt: die geometrische Struktur bleibt erhalten (Xiong et al. 2005), d.h.

- Punkte aus  $\mathcal{F}_{[0,1]}$  werden in die projizierte Hypersphäre abgebildet,
- Punkte von außerhalb liegen auch in  $\mathcal{H}^{(\ell)}$  außerhalb der Sphäre.

Für die Durchführung des ersten Schrittes ist nunmehr die Möglichkeit gegeben. Das in diesem Schritt erzeugte Bild  $\hat{\Psi}_x$  muss im nächsten Schritt modifiziert werden, um zu einem Bild eines gültigen Fahrplans zu werden.

### 9.3.3. Schritt 2: Rejustierung im RKHS

Im Allgemeinen gilt, dass das Bild der mittels SVDD kodierten Region im RKHS  $\mathcal{H}$  als Hypersphäre  $\mathcal{S}$  mit Mittelpunkt  $a$  und Radius  $R_S$  (siehe Gleichung 5.27) dargestellt wird. Punkte außerhalb der Sphäre sind keine Bilder aus  $\mathcal{X}$ , d.h. im hier betrachteten Fall keine Bilder aus  $\mathcal{F}_{[0,1]}$ . Punkte aus  $\mathcal{F}_{[0,1]}$  werden unter  $\Phi$  in die Sphäre abgebildet (oder auf die Oberfläche), Bilder von Punkten von außerhalb  $\mathcal{F}_{[0,1]}$  befinden sich in  $\mathcal{H}$  auch außerhalb der Sphäre.

Durch die Verwendung eines Gaußkerns bildet  $\Phi$  genau genommen in eine maximal  $n$ -dimensionale Mannigfaltigkeit (mit  $n$ : Stichprobenumfang) ab, welche eingebettet ist in einen unendlichdimensionalen Raum  $\mathcal{H}$ . Diese Mannigfaltigkeit wird aufgespannt von den Bildern der Stichprobe. Dies gilt analog auch für die niedriger dimensionalen, von den  $\ell$  Supportvektoren aufgespannten Unterraum und der niedriger ( $\ell$ -) dimensionalen Projektion der Sphäre darin, wenn  $\hat{\Phi}_\ell$  (Gleichung 9.5) für die Abbildung verwendet wird.

Ziel dieses Schrittes ist es, Punkte von außerhalb  $\mathcal{F}_{[0,1]}$  auf möglichst direktem Weg in den zulässigen Bereich zu ziehen und ungültige Lösungen so gültig zu machen. Da SVDD eine Beschreibung der Region liefert welche auf dem Bild der Region und damit auf den Bildern der ursprünglichen Punktmenge beruht, werden somit auch die Bilder der zuvor mit  $\hat{\Phi}_\ell$  abgebildeten Punkte in das Bild der Region  $\Phi(\mathcal{F}_{[0,1]})$  hineingezogen; genauer: in die durch die Supportvektoren aufgespannte  $\ell$ -dimensionale Projektion der Region.

Zu diesem Zweck kann beispielsweise die folgende Vektorfunktion genutzt werden:

$$\tilde{\Psi}_x = \Gamma_a(\hat{\Psi}_x) = \hat{\Psi}_x + \mu \cdot (a - \hat{\Psi}_x) \cdot \frac{R_x - R_S}{R_x}, \quad (9.7)$$

um das Bild  $\hat{\Psi}_x = \hat{\Phi}_\ell(x)$  aus Schritt 1 zu transformieren und in ein Bild aus der zulässigen Region  $\tilde{\Psi}_x \in \hat{\Phi}_\ell(\mathcal{F}_{[0,1]})$  umzuwandeln. Dies wird erreicht, indem  $\hat{\Psi}_x$  in die Sphäre hinein verschoben wird. Der wählbare Parameter  $\mu \in \mathbb{R}$  beeinflusst als Skalierungsfaktor die Länge des Verschiebungsvektors. Abbildung 9.1 auf Seite 119 verdeutlicht diesen Zusammenhang.

Die einfachere Version

$$\tilde{\Psi}_x = a + \frac{(\hat{\Psi}_x - a) \cdot R_S}{R_x} \quad (9.8)$$

wie sie in ähnlicher Form auch beim Entrauschen (dort allerdings durch geodätische Projektion mit Hilfe einer Bestimmung der trennenden Hyperebene) Verwendung findet (vgl. beispielsweise (Park et al. 2007)) würde das Bild lediglich auf die Oberfläche der Sphäre verschieben. Dafür muss bei der vereinfachten Gleichung (9.8) der Parameter  $\mu \in [1, R_x]$  nicht abgeschätzt werden, welcher kontrolliert wie stark ein Punkt im Bildraum in die Sphäre hineingeschoben wird. Im Fall  $\mu = 1$  ist Gleichung (9.7) äquivalent zur einfacheren Version aus Gleichung (9.8); im Falle  $\mu = R_x$  werden alle Punkte im Bildraum auf den Mittelpunkt der Sphäre abgebildet (vgl. auch Abbildung 9.5(d) auf Seite 128). Durch Gleichung (9.7) können alle Punkt im Bildraum proportional zu ihrer ursprünglichen Entfernung vom Mittelpunkt der Sphäre verschoben werden in Richtung des Mittelpunktes – auch wenn sie bereits innerhalb der zulässigen Region liegen.

Wäre die Entfernung des am weitesten entfernt gelegenen Bildes  $R_{max}$  bekannt, so könnte Gleichung (9.7) ersetzt werden durch

$$\tilde{\Psi}_x = a + \frac{R_S}{R_{max}} \cdot (\hat{\Psi}_x - a). \quad (9.9)$$

Für die praktische Anwendung der Wirkleistungsplanung, kann jedoch die Einschränkung gemacht werden, dass lediglich skalierte Fahrpläne verwendet werden und somit Punkte, die mit  $\gamma$  abgebildet werden sollen immer aus dem Einheitshyperwürfel  $[0, 1]^d$  stammen.  $R_{max}$  kann somit in erster Näherung abgeschätzt werden als  $\max_i \{R(x_i)\}$  für eine Menge zufällig generierter Fahrpläne  $x_i \in [0, 1]^d$ .

Punkte innerhalb der Sphäre werden somit ebenfalls verschoben und weiter in die Mitte gerückt, um – bildlich gesehen – Platz zu machen für Punkte, welche von weiter außerhalb in den äußeren Bereich der Sphäre gezogen werden. Auf diese Weise wird das Abbild des gesamte Einheitshyperwürfels in  $\mathcal{H}^{(\ell)}$  quasi in die Form der Sphäre gedrückt ohne die Dichte der Punkte am Rand zu sehr zu erhöhen. Für die nachgelagerte Optimierung wird es von Vorteil sein, wenn nicht alle Punkte lediglich auf den Randbereich abgebildet werden. Dennoch, ist die Region zulässiger Lösungen im Verhältnis zum gesamten Einheitshyperwürfel klein, so erhöht sich die Dichte der abgebildeten Punkte am Rand der Hypersphäre. Dieser Effekt ist abhängig von der Wahl des Parameters  $\mu$  und kann durchaus gewollt sein, wenn beispielsweise aufgrund der Problemstellung bekanntermaßen das eigentliche Optimum außerhalb des zulässigen Bereichs und somit gute zulässige Lösungen am Rand zu erwarten sind. Tatsächlich liegt in vielen Real-World-Problemen die Lösung am Rand der zulässigen Lösungen (Simon 2013; Leguizamón & Coello 2009).

Beim Rechnen im Feature Raum  $\mathcal{H}$  gibt es einige Besonderheiten zu beachten. Alle Bilder  $\Psi$  in  $\mathcal{H}$  liegen vor als Expansion  $\sum_i \beta_i \cdot \Phi(x_i)$ . Daher wird beispielsweise die Addition zweier Punkte  $\Psi_p = \sum_i \alpha_i \cdot \Phi(x_i)$  und  $\Psi_q = \sum_i \beta_i \cdot \Phi(x_i)$  auch durch Addition der Gewichtungsvektoren ausgeführt (vgl. z. B. auch Herbrich et al. 1999):

$$\Psi_p + \Psi_q = \sum_i \alpha_i \cdot \Phi(x_i) + \sum_i \beta_i \cdot \Phi(x_i) = \sum_i (\alpha_i + \beta_i) \cdot \Phi(x_i). \quad (9.10)$$

Im Anschluss des hier vorgestellten Berechnungsschritts für die Rejustierung liegt somit das modifizierte Bild  $\tilde{\Psi}_x$  des Ausgangspunktes  $x$  in Form des veränderten Gewichtungsvektors  $\tilde{w}^{\Gamma_\alpha}$  vor und repräsentiert bereits das Bild eines Punktes aus  $\mathcal{F}_{[0,1]}$ .

### 9.3.4. Schritt 3: Urbildbestimmung

Als letzter Schritt muss noch das Urbild des Bildes  $\tilde{\Psi}_x$  aus Schritt 2 bestimmt werden, um einen Punkt aus  $\mathcal{F}_{[0,1]}$  zu erhalten. Ein wesentliches Problem bei der Bestimmung eines Urbildes zu einem Punkt aus dem RKHS ist, dass nicht jeder Punkt aus der linearen Hülle  $\mathcal{L}(\Phi)$  das Bild eines unter  $\Phi$  abgebildeten Punktes ist (Schölkopf et al. 1999a). Im Falle des hier verwendeten Gaußkerns kann sogar zu keinem Punkt aus  $\mathcal{H}$  ein exaktes Urbild gefunden werden (ausgenommen für triviale Fälle mit nur einem Term) (Kwok & Tsang 2004; Mika et al. 1999a).

Stattdessen wird derjenige Punkt  $\mathbf{y}^* \in [0, 1]^d$  bestimmt, für den der Abstand zum tatsächlichen Urbild  $\|\mathbf{y}^* - \Phi^{-1}(\mathbf{y})\|$  minimal ist. Dieser Umstand kann auch formuliert werden als  $\mathbf{y}^*$  ist approximiertes Urbild von  $\mathbf{y}$ , falls  $\|\Psi_{\mathbf{y}} - \Phi(\mathbf{y}^*)\|$  klein ist (Schölkopf et al. 1999a), wobei die Größenordnung von *klein* vom konkreten Problem abhängt. Aus diesem Grund bietet es sich an, ein Näherungsverfahren zu verwenden, wie es beispielsweise in (Mika et al. 1999b) bzw. für die Verwendung von Gaußkernen vorgeschlagen wird.

Im Falle eines Gaußkerns kann folgende Iterationsvorschrift als Näherung  $\Phi_\ell^{-1}$  verwendet werden (vgl. z. B. Schölkopf et al. 1999a):

$$\mathbf{x}_{n+1}^* = \frac{\sum_{i=1}^{\ell} (\tilde{w}_i^{\Gamma_\alpha} e^{-\|\mathbf{s}_i - \mathbf{x}_n^*\|^2 / 2\sigma^2} \mathbf{s}_i)}{\sum_{i=1}^{\ell} (\tilde{w}_i^{\Gamma_\alpha} e^{-\|\mathbf{s}_i - \mathbf{x}_n^*\|^2 / 2\sigma^2})}. \quad (9.11)$$

Da die Möglichkeit besteht, dass dieses Verfahren in lokalen Minima steckenbleibt, ist es empfehlenswert, das Verfahren mit verschiedenen Startwerten auszuführen und am Ende die beste gefundene Lösung auszuwählen. Um ein deterministisches  $\Phi_\ell^{-1}$  zu erhalten, müssen die Startwerte reproduzierbar gewählt werden. In der Praxis hat es sich empirisch als hinreichend erwiesen, den ursprünglichen Punkt  $x$  als Startwert zu nehmen. Sollte das Verfahren dennoch in einem Punkt  $\mathbf{x}^* \notin \mathcal{F}_{[0,1]}$  steckenbleiben, so muss gegebenenfalls nach festem Schema ein neuer Startwert gewählt werden. In den in dieser Arbeit betrachteten Szenarien konnte dieses Problem bisher aber nicht beobachtet werden.

Mit Ausführung dieses dritten Schrittes ist das Ziel erreicht: Der ursprüngliche, ungültige Punkt  $x$  wurde abgebildet auf einen nunmehr gültigen Punkt  $\mathbf{x}^* = \gamma(x)$ . Damit ist der Dekoder fertig konzipiert. Der nachfolgende Abschnitt wird sich mit der Illustration des Konzeptes anhand von anschaulichen Beispielen befassen.

## 9.3.5. Illustration

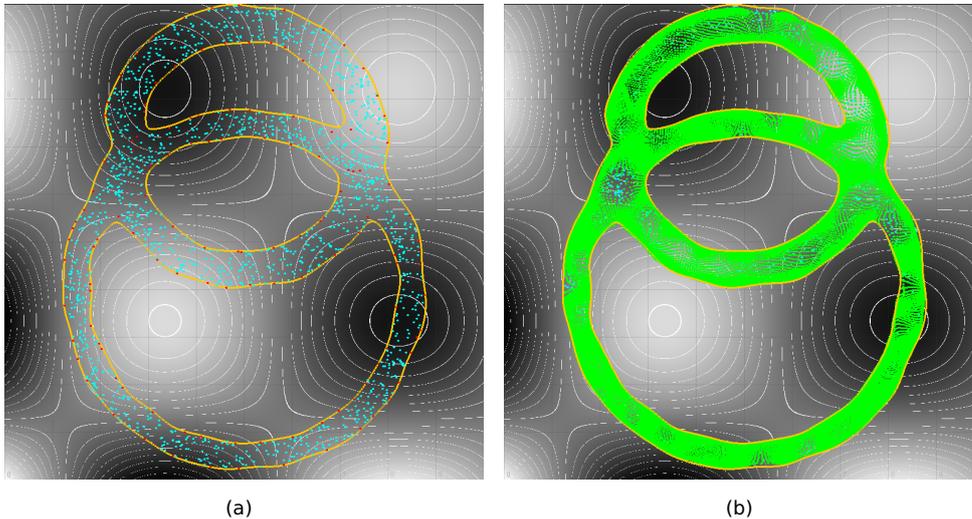


Abbildung 9.2.: Dekoderausgabe für das künstliche Beispiel eines Doppelrings.

Abschließend soll auch dieses Verfahren zunächst anhand illustrativer Beispiele rekapituliert werden. Abbildung 9.2 zeigt die Ausgabe eines Dekoders für ein 2-dimensionales, künstliches Beispiel aus (Bremer & Sonnenschein 2013b) zunächst jedoch noch ohne Bezug zum Anwendungsfall. Hierfür wurde eine Trainingsmenge mit Punkten generiert, die alle innerhalb einer Doppelringstruktur liegen. Die türkisfarbenen Punkte in der linken Abbildung 9.2(a) zeigen die Trainingsmenge, auf die das SVDD-Verfahren angewendet wurde. Als Ergebnis wurde die orangefarbene Linie als Entscheidungsgrenze zwischen Zugehörigkeit und Nichtzugehörigkeit zur Trainingsmenge erkannt. Die roten Punkte stellen diejenigen Punkte dar, die zu Supportvektoren wurden. Anschließend wurde das Dekoderverfahren, wie zuvor beschrieben auf zufällig generierte Punkte angewendet. Eine Menge von Punkten  $x_i \sim U(0, 1)^2$  wurde mit Hilfe des Dekoders abgebildet auf die zuvor gelernte Doppelringstruktur. Die rechte Abbildung zeigt in grün die Menge der abgebildeten Punkte  $\gamma(x_i)$ . Der Hintergrund codiert farblich eine künstliche Zielfunktion.

Die nächste Abbildung visualisiert die durch den Dekoder verursachte Verschiebung. Die Pfeile in Abbildung 9.3 zeigt die Verschiebung einzelner Punkte (an einem Bananen-Trainingsatz) und somit deren Spur bei der Verschiebung in den Lösungsraum hinein. Die Abbildung zeigt auch, wie einzelne, bereits innen liegende Punkte noch ein wenig weiter Richtung Kugelmittelpunkt verschoben werden, um Platz für von außen kommende Punkte zu machen. Dunkle Regionen (blau außen und oliv innen) kennzeichnen eine größere Entfernung von der Entscheidungsgrenze (orange). Man erkennt auch hier wie marginal die Unterschiede (gemessen als Entfernung des Bildes vom Kugelmittelpunkt) in größerer Entfernung außerhalb vom Lösungsraum sind. Ein Penaltyansatz könnte hier (insbesondere in höherdimensional Fällen) nicht unterscheiden und könnte somit keine Richtung erkennen. Der Dekoder jedoch bildet korrekt in den Lösungsraum hinein ab. Insbesondere auch weiter entfernt liegende Punkte. Die Abbildung erfolgt dabei in einem

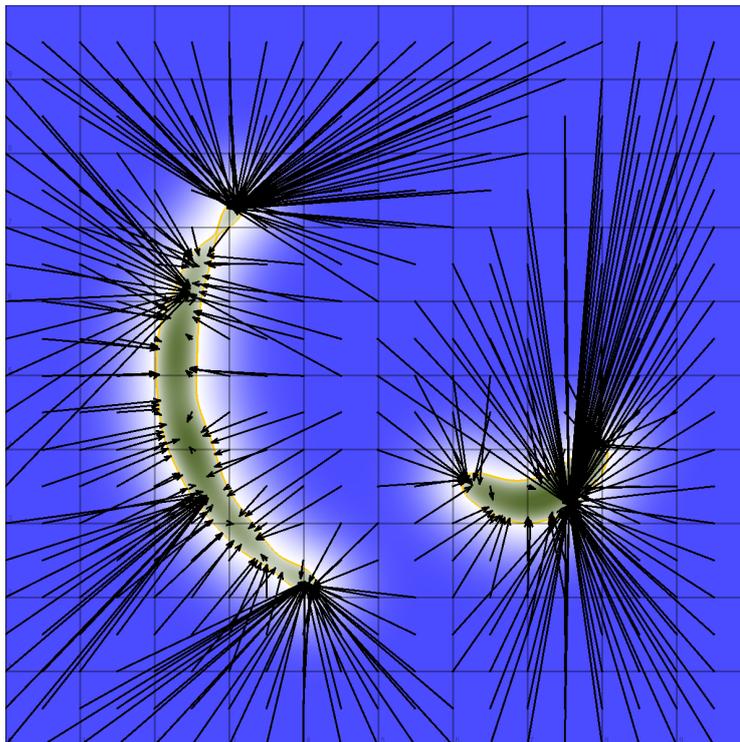


Abbildung 9.3.: Dekoderabbildung am künstlichen Beispiel. Die Pfeile repräsentieren in den Lösungsraum verschobene Punkte.

Schritt und braucht keine Zwischenschritte, die auch ungültige Lösungen betrachten würden, wie z. B. ein Penaltyansatz.

Das gleiche Bild lediglich für höherdimensionale Fälle zeigt [Abbildung 9.4 auf der nächsten Seite](#). Dargestellt sind für zwei Fälle (8-dimensionale Fahrpläne links bzw. 32-dimensionale rechts) jeweils Punkte aus dem Lösungsraum und Pfeile, welche die Abbildung ungültiger Punkte repräsentieren. Es wurden zum Zweck der Darstellung lediglich alle Punkte mit Hilfe eines Relative Distance Plane Mappings (RDP) auf eine zweidimensionale Einbettung abgebildet. Das RDP (Somorjai et al. [2004](#)) beruht auf einer intrinsisch nicht-linearen Projektion. Sie basiert auf der Tatsache, dass bei drei beliebigen hochdimensionalen Punkten die drei Abstände bei einer Darstellung im zweidimensionalen Koordinatensystem erhalten bleiben. Daher werden bei diesem Verfahren zwei beliebige Referenzpunkte gewählt, die eine Achse definieren auf die alle weiteren Punkte projiziert werden. Auf diese Weise werden bei der Projektion alle Distanzen relativ zu den beiden Referenzpunkten exakt erhalten. Der Vorteil der RDP-Methode ist die hervorragende Performance bei großen Datenmengen. Eine systematische Variation der Referenzpunkte (und somit ein systematisches Verschieben der Projektionsebene) ermöglicht auch ein zügiges Betrachten verschiedener 2-dimensionaler Einbettungen. Im vorliegenden Fall wurden aufgrund der Vergleichbarkeit verschiedener separat berechneter Einbettungen jedoch 0 und 1 fixe Referenzpunkte verwendet.

In dieser Abbildung zeigt sich auch bereits eine Limitierung, die durch zu kleine Trainingsmengen entsteht. Zu kleine Trainingsmengen können die Struktur des Lösungsraums

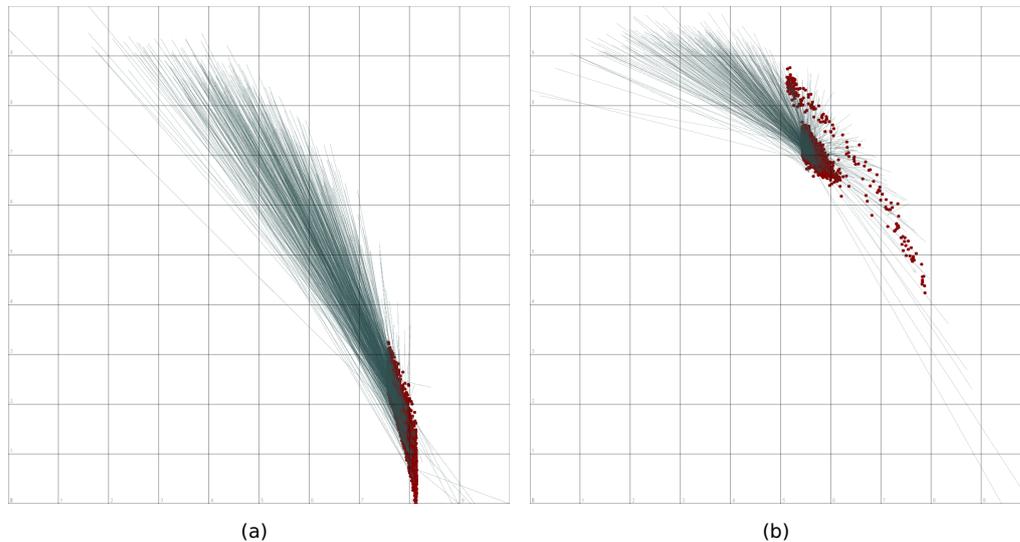


Abbildung 9.4.: Hochdimensionale Dekoderabbildung an zwei Beispielen. Abbildung 9.4(a) zeigt die Abbildung eines Dekoders für einen 8-dimensionalen Beispielsuchraum (durch ein Wärmepumpenmodell erzeugt) am Beispiel von 500 abgebildeten Punkten und ihrem Mapping; Abbildung 9.4(b) zeigt analog ein 32-dimensionales BHKW-Beispiel.

naturgemäß nicht vollständig wiedergeben. Das Trainingsverfahren kann aber auch nur die Teile erfassen, die auch durch die Trainingsmenge wiedergespiegelt werden. In Abbildung 9.4 stellen die roten Punkte eine Menge von jeweils 20000 gültigen Fahrplänen (als Näherung des nicht kontinuierlich darstellbaren Lösungsraums) dar, zum Training wurden jedoch nur 800 verwendet. Wie man leicht sieht, bildet der aus den unvollständigen Informationen gelernte Dekoder nicht in alle Regionen ab, sondern spart offensichtlich einige Randregionen aus, die von der Trainingsmenge nicht erfasst wurden. Dieses Problem wird immer auftreten, da eine endliche, diskrete Trainingsmenge die kontinuierlichen Räume immer unvollständig abbilden wird. Hier wird letztlich eine Trade-off zwischen Rechengeschwindigkeit und benötigter Genauigkeit zu finden sein. Mögliche Erweiterungen zur Adressierung dieses Problems werden auch in Kapitel 17 ab Seite 231 noch behandelt.

Wie bereits in Abschnitt 9.3.3 auf Seite 122 erläutert, kann die Stärke mit der Punkte in das innere der Kugel gezogen werden parametrisiert werden. Abbildung 9.5 zeigt die Auswirkungen am Beispiel des Bananen-Trainingsatzes. Hierbei wurde jeweils ein Dekoder mit künstlicher zusätzlicher Verkürzung des resultierenden Bildvektors (der vom Kugelmittelpunkt aus auf das jeweils verschobene Bild eines Punktes zeigt) durch Verkleinerung von  $\mu$  in Gleichung (9.7), sodass die Verschiebung jeweils weiter in die Kugelmitte hinein erfolgt. Dunkle Farben kennzeichnen wieder eine weitere Entfernung vom Kugelrand und somit auch Positionen zur Kugelmitte hin.

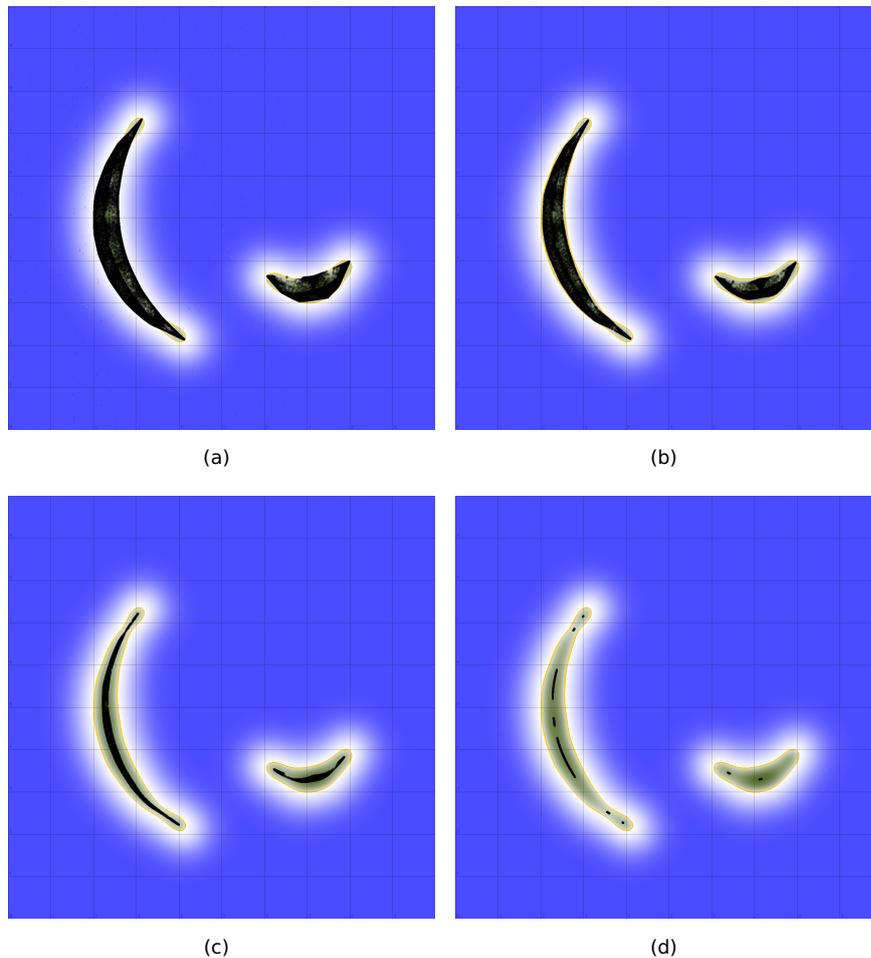


Abbildung 9.5.: Darstellung des Skalierungseffektes bei der Rejustierung im RKHS. Von links oben nach rechts unten: 100%, 10%, 1% bzw. 0.1% von  $\mu$  aus Gleichung (9.7) ziehen die Bilder abgebildeter Punkte kontinuierlich weiter zur Mitte der Hypersphäre.

## 9.4. Integration in Optimierung

Bildet man ein regelmäßiges Gitter mit Hilfe des Dekoders ab, so entsteht beispielsweise die in Abbildung 9.6 dargestellte Figur. Hierbei wurden alle Kreuzungspunkte des Gitters mithilfe eines Dekoders abgebildet. Die Beziehungen (bzw. Verbindungslinien) zwischen den Punkten wurden beibehalten. Man kann sich nunmehr vorstellen, dass ein Optimierungsverfahren, welches sich auf der Suche nach einem Optimum am ursprünglichen Gitter entlang bewegt hätte, bei Verwendung des Dekoders tatsächlich die neu entstandene Figur absucht und sich damit automatisch nur im Lösungsraum bewegt. Wie die Abbildung deutlich zeigt, werden Punkte an unterschiedlichen Rändern des Lösungsraums unter dem Dekoder zu Nachbarnpunkten, sodass ein Suchverfahren an solchen Stellen quasi den Zwischenraum ungültiger Lösungen automatisch überspringen würde.

Abbildung 9.6 zeigt ebenfalls, dass bei dem hier vorgeschlagenen Supportvektor-

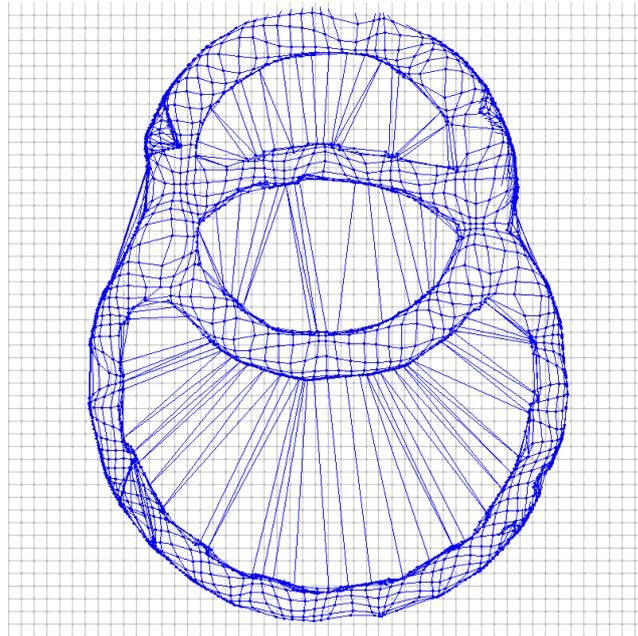


Abbildung 9.6.: Dekoderabbildung eines regelmäßigen Gitters aus (Bremer & Sonnenschein 2013b).

Dekoder ggf. das Lokalitätskriterium für gute Dekoder (Dasgupta & Michalewicz 1997) verletzt sein kann. Dies ist bei konkaven oder disjunkten Suchräumen jedoch kaum vermeidbar. Der gleiche Effekt würde auch bei anderen Dekodern (z. B. Koziel & Michalewicz 1999) in nicht-konvexen Suchräumen) auftreten.

Man sieht an dieser Abbildung jedoch auch, dass die Region am Rand des Lösungsraums aufgrund einer offenbar höheren Punktdichte stärker durchsucht würde als das Innere. Dies gilt insbesondere, wenn der Lösungsraum im Vergleich zum umgebenden Hyperwürfel (im Bild die Ringstruktur im Vergleich zur gesamten Rechteckfläche) eher klein ist, da in einem solchen Fall besonders viele ungültige Punkte von außen in die Figur hinein abgebildet werden. Andererseits wächst in diesem Fall auch die Wahrscheinlichkeit, dass das wahre Optimum außerhalb der Figur liegt und das zulässige Optimum somit am Rand zu finden ist.

Ein weiteres Beispiel zeigt Abbildung 9.7. Wie auch schon in dem Fall aus Abbildung 9.2 wurde anhand einer diesmal bananenförmigen Trainingsmenge 9.7(a) die Entscheidungsgrenze gelernt sowie ein Dekoder abgeleitet. Abbildung 9.7(b) zeigt wieder die Abbildung einer Menge zufälliger Punkte. Als Zielfunktion (wieder farblich kodiert als Hintergrund dargestellt) wurde hier die bekannte Testfunktion für Optimierungsansätze Six-Hump-Camel-Back

$$f_{CB}(x_1, x_2) = (4 - 2.1x_1^2 + \frac{x_1^4}{3})x_1^2 + x_1x_2 + (-4 + 4x_2^2)x_2^2 \quad (9.12)$$

nach (Szegö 1978) verwendet, mit dem Wertebereich  $-1.9 \leq x_1 \leq 1.9$ ,  $-1.1 \leq x_2 \leq 1.1$  skaliert zu  $[0, 1]^2$ . Die beiden Kreuze in 9.7(c) markieren die mit Hilfe von Partikel Swarm Optimierung (siehe auch Kapitel 12 ab Seite 177) gefundenen Optima. Das Szenario wurde so gewählt, dass ein Optimum im Inneren des zulässigen Bereiches und eines außerhalb

liegt. In beiden Fällen war der Dekoder erfolgreich und hat das Optimierungsverfahren korrekt zu gültigen Lösungen gelöst.

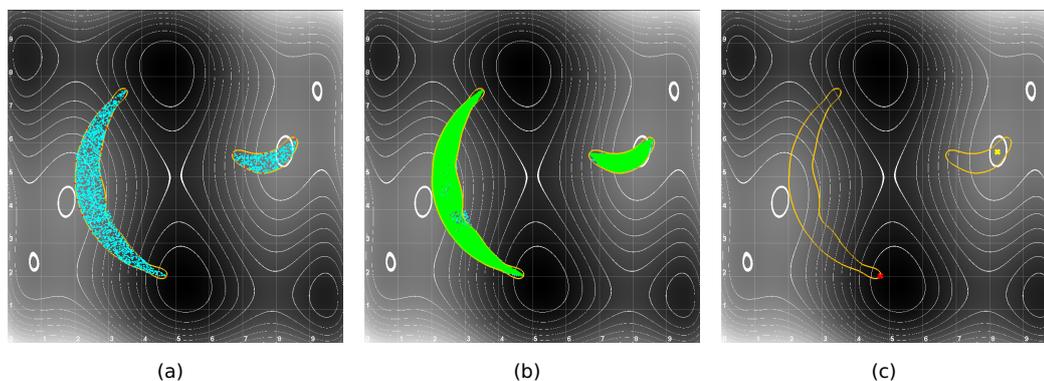


Abbildung 9.7.: Ein erstes, artifizielles Dekoder-Optimierungsbeispiel (Bremer & Sonnenschein 2013b).

An den Beispielen in diesem Abschnitt lässt sich ebenfalls bereits erkennen, dass der Dekoder auf einer Reihe von Näherungen beruht. Bereits das SVDD-Modell beruht naturgemäß auf unvollständigen Informationen, da bereits die Trainingsmenge das mathematische Modell einer Anlage nur näherungsweise wiedergeben kann. Hinzu kommt, dass immer vereinfachend mit einer endlichen Menge diskreter Trainingsdaten gearbeitet werden muss. Zudem beruht die Bestimmung der Flexibilität einer Anlage oft auf Prognosen wie beispielsweise dem zukünftigen Wärmebedarf, der durch ein BHKW zu decken ist. Zum anderen wird bei der Berechnung der Supportvektoren auf eine Heuristik zur schnelleren Berechnung zurückgegriffen (vgl. Abschnitt 5.2.2 auf Seite 57). Als dritte Approximation kommt die Urbildbestimmung (Schritt 3 im Dekoder) ins Spiel, welche aufgrund der Verwendung eines Gaußkerns wie bereits erläutert immer nur Näherungen liefern kann. Zusätzlich wird – zumindest hier für die Wirkleistungsplanung – auch für den nachgeschalteten, eigentlichen Optimierungsschritt eine Heuristik Verwendung finden. Mit dem Auftreten kleiner Fehler wird bei der Anwendung des Dekoders immer zu rechnen sein. Wie groß diese Fehler sind und ob diese hinreichend gut erkannt und (falls notwendig) behoben werden können ist immer eine Frage des konkreten Anwendungsfalls und der Parametrisierung des Verfahrens.

Treten bei der hier betrachteten Wirkleistungsplanung geringe Fehler in den späteren Fahrplänen auf (im Bereich von wenigen Watt bei einer Maximalleistung von mehreren kW) so könnten diese vermutlich relativ einfach durch die Anlagensteuerung ausgeglichen werden. Im Falle von Aufgabenstellungen, die eine höhere Genauigkeit verlangen, müsste gegebenenfalls auf den Einsatz einiger Heuristiken zugunsten langsamerer aber exakterer Verfahren verzichtet werden. An anderer Stelle wäre bei der Parametrisierung ein Trade-Off zugunsten der Genauigkeit anstatt zugunsten höherer Performanz zu wählen. Für den Fall der Wirkleistungsplanung wird dieser Aspekt im Evaluationskapitel 11 auf Seite 147 noch einmal aufgegriffen.

## 9.5. Zusammenfassung

Mit dem Supportvektor-Dekoder konnte in diesem Kapitel der letzte Baustein für die eingangs skizzierte Prozesskette zur Integration von Constraints in ein Optimierungsproblem wie das Wirkleistungsplanungsproblem geschaffen werden.

Dezentrale Erzeuger sollen sich dynamisch mit Speichern und steuerbaren Verbrauchern zusammenschließen können, um gemeinsam einen definierten (am Markt ausgeschriebenen) Wirkleistungsverlauf für einen gegebenen Zeitraum nach Fahrplan zu realisieren. Betrachtet werden sollen also Day-Ahead-Szenarien, wobei der Zeitraum auch kürzer als 24 Stunden sein kann. Auch soll das Verfahren unabhängig von der Länge der heute in der Elektrizitätswirtschaft immer noch üblichen 15-Minuten Planungsintervallen sein. Essentielle Voraussetzungen für die Erstellung eines optimalen Gesamtfahrplans für eine Gruppe von dezentralen Energieerzeugern und -verbrauchern ist die Kenntnis der aktuellen, individuellen Handlungsspielräume (individuelle Suchräume mit realisierbaren Alternativlastgängen) der verschiedenen Anlagen. Diese müssen durch die Anlagen selbst prognostiziert und kommuniziert werden. Die Prognose muss in den Anlagen selbst erfolgen, da hierfür neben den eigenen anlagenspezifischen Betriebsparametern auch die Kenntnis des Bedarfs sowie des aktuellen Betriebszustandes notwendig ist. Mit dem im vorigen Kapitel beschriebenen Klassifikatormodell ist es bereits möglich diese individuellen Fahrplanmengen effizient anhand exemplarischer Fahrpläne zu beschreiben.

Für die Planung im Rahmen des gesamten Verbundes ergibt erst die Kombination der Suchräume aller Anlagen den Gesamtsuchraum, in dem nach dem günstigsten Einsatzplan (als Kombination von Einzelfahrplänen) gesucht werden muss. Jede Anlage verfügt über individuelle Constraints und der dynamische Charakter der Verbünde verbietet eine statische Verbundmodellierung. Mit dem in diesem Kapitel beschriebenen Dekoder jedoch wird es möglich, beliebige Anlagentypen einheitlich in ein Optimierungsverfahren einzubinden. Das Verfahren benötigt lediglich für jede Anlage den Dekoder. Die Suche findet immer in einer Menge von Einheitshyperwürfeln  $[0, 1]^d$  (einer je Anlage) statt. Die Dekoder bilden diese Einheitshyperwürfel individuell auf die zulässigen Bereiche der individuellen Lösungsräume der Anlagen ab, ohne dass das Optimierungsverfahren hierfür eine konkrete mathematische Modellierung der Anlagen kennen müsste. Das Optimierungsverfahren generiert systematisch Fahrpläne  $x \in [0, 1]^d$  und der passende Dekoder erzeugt hieraus einen realisierbaren Fahrplan bevor dieser durch eine gegebene Zielfunktion bewertet wird. Die Darstellung der Fahrpläne in Vektorform erlaubt ebenfalls eine Abstraktion von Anzahl und Länge der Zeitintervalle über die der Fahrplan gilt.

Auf diese Weise kann das Optimierungsmodell (im zentralen Fall) für einen neu gebildeten Verbund ebenfalls dynamisch erzeugt werden, indem einem Optimierungsverfahren die einzelnen Dekoder bekannt gemacht werden. Im verteilten Fall behält ggf. jede Anlage ihren Dekoder für sich, kann diesen jedoch ebenfalls zur systematischen Generierung realisierbarer Fahrplanalternativen während der Teilnahme an einer verteilten Optimierung nutzen – eine Fähigkeit, welche eine mathematische Modellierung einer Anlage allein üblicherweise nicht leisten kann.

Zentrale und dezentrale Optimierungsverfahren als mögliche Anwendung für den Dekoder werden in den Kapiteln Kapitel 12 ab Seite 177 und Kapitel 13 ab Seite 199 noch getrennt betrachtet. In der Praxis wird in der Regel eine Mehrziel-Optimierung angestrebt

## 9. Supportvektor-Dekoderansatz

---

werden. Im nachfolgenden Kapitel wird daher zunächst noch die Frage betrachtet, wie das bisherige Modell genutzt werden kann, um die Bewertung einzelner Fahrpläne als Lösungskandidaten hinsichtlich verschiedener Optimierungskriterien zu erweitern.

# 10. Modellierung von Kosten und Kennzahlen

*I can make more generals,  
but horses cost money.*

---

*(Abraham Lincoln)*

## 10.1. Einleitung

In vielen Optimierungsszenarien sind unterschiedliche Lösungskandidaten oft auch mit unterschiedlichen Kennzahlen zur Bewertung hinsichtlich verschiedener Zielfunktionen verbunden. Kennzahlen (üblicherweise als Kenngröße bestehend aus Kennzahl und Einheit) sind Maßzahlen zur Quantifizierung bestimmter Eigenschaften und (in der Praxis) ursprünglich im Unternehmens-Controlling beheimatet (ZVEI 1970). Aber auch im Öko-Controlling und im Risikomanagement sind sie seit langem beheimatet (Rautenstrauch 1999; Davies & Haubenstock 2002). In dieser Arbeit soll unter einer Kennzahl allgemein ein Indikator verstanden werden, der die Bewertung einer (Optimierungs-)Lösung im Hinblick auf bestimmte Eigenschaften und Auswirkungen eines Lastgangs erlaubt. Beispiele hierfür sind:

- Kosten für die Erzeugung eines Lastgangs,
- Zuverlässigkeit mit der ein Lastgang realisiert werden kann,
- Primärenergieeinsatz oder andere Umweltkennzahlen,
- verbleibende Flexibilität für nachfolgende Planungszeiträume.

Auf konzeptioneller Ebene werden generell zwei Typen von Kennzahlen unterschieden:

- zu maximierende Performanzindikatoren (z. B. verbleibende Flexibilität), und
- zu minimierende Intensitätsindikatoren (z. B.  $CO_2$ -Ausstoß oder der Primärenergieeinsatz)

In klassischen Mehrziel-Optimierungsszenarien ist normalerweise eine Annotation von Lösungskandidaten mit Kennzahlen zur Bewertung hinsichtlich verschiedener Kriterien nicht notwendig. Hier liegen die Bewertungskriterien direkt als mathematische Formulierung in Form eines Satzes von Zielfunktionen vor. Ein solches Vorgehen ist in dem hier vorgeschlagenen Ansatz bei dem die Optimierung auf einem Satz von dezentral erstellten Suchraummodellen mit Dekoder arbeitet nicht immer sinnvoll möglich. Zwar wäre es technisch möglich zusammen mit dem Suchraummodell und dem Dekoder auch einen Satz von Bewertungsfunktionen zu kommunizieren (im zentralen Fall), jedoch sind

die funktionalen Zusammenhänge zur Bewertung einzelner Fahrpläne gegebenenfalls Betriebsgeheimnis und damit nicht außer Haus zu geben. Wie in diesem Kapitel gezeigt werden wird, gibt es aber eine elegante Alternative.

Die Motivation des Themas Kennzahlenintegration in das Suchraummodell soll zwar zunächst unabhängig von einem konkreten Optimierungsszenario und somit auch unabhängig von einem gegebenen Anwendungsszenario geschehen, dennoch werden hier zur Verdeutlichung wie bisher auch Praxisbeispiele aus dem Smart Grid-Umfeld gewählt.

Der Rest dieses Kapitels befasst sich zunächst mit der konzeptionellen Umsetzung der Integration von Kennzahlen in das bisherige Suchraummodell (einschließlich Dekoder). Anschließend werden die im Rahmen dieser Arbeit verwendeten Kennzahlen vorgestellt sowie erste Simulationsergebnisse zur anwendungsfallunabhängigen Evaluation (und Illustration) der Kennzahlenintegration vorgestellt.

## 10.2. Integrationsansatz

Wie bereits in Abschnitt 6.4 auf Seite 84 beschrieben, ist im Kontext dieser Arbeit eine Kennzahl ein Zahlenwert, welcher einen Fahrplan im Hinblick auf die Erreichung bestimmter (Neben-)Ziele charakterisiert. Im Zuge der Einsatzplanung werden diese Kennzahlen herangezogen um verschiedene Fahrpläne hinsichtlich dieser Bewertung gegeneinander abzuwägen.

In Definition 6.13 auf Seite 86 wurde die skalierte Kennzahl zur Bewertung eines Fahrplans  $x$  in Abhängigkeit vom Anlagenstartbetriebszustand  $z_0$  definiert als

$$\mathcal{I}_{[0,1]}^{(Q)}(\mathbf{x}, z_0) : \mathbb{R}^d \times Z_U \rightarrow [0, 1],$$

$$x \mapsto \frac{\mathcal{I}^{(Q)}(\mathbf{x}, z_0)}{\max_{\mathbf{x} \in \mathbb{R}^d, z \in Z_U} \{\mathcal{I}^{(Q)}(\mathbf{x}, z)\}}. \quad (10.1)$$

Gegeben seien die Indikatorfunktion  $\mathcal{I}_{[0,1]}^{(Q)}$  für die Kenngröße  $Q$  und der Startzustand  $z_0$  einer Anlage. Gegeben sei ferner eine Trainingsmenge  $\mathcal{X} = \{\mathbf{x}_i\}_n \in \mathbb{R}^d$  von  $n$  Beispielfahrplänen. Jedes Element eines Fahrplanvektors  $\mathbf{x} \in \mathcal{X}$  beschreibt die mittlere Wirkleistung im entsprechenden Zeitintervall. Für die Kennzahlenintegration werden jetzt neue Trainingsvektoren konstruiert in dem an jeden Fahrplanvektor der Kennzahlenwert (gegebenenfalls auch mehrere) als zusätzliches Element angehängt wird. Für jede Kennzahl kommt ein Element zum Fahrplanvektor hinzu:

$$\mathcal{X}^* = \{(x_1, \dots, x_d, \mathcal{I}_{[0,1]}^{(Q_1)}([x_1, \dots, x_d], z_0), \dots, \mathcal{I}_{[0,1]}^{(Q_m)}([x_1, \dots, x_d], z_0))\}_n. \quad (10.2)$$

Die Elemente der modifizierten Trainingsmenge  $\mathcal{X}^*$  stellen nunmehr Featurevektoren dar deren erste  $d$  Elemente nach wie vor die mittlere Wirkleistung kodieren. Die letzten  $m$  Elemente kodieren jeweils eine den zugehörigen Fahrplan bewertende Kennzahl, wobei die Position eindeutig für die zugehörige Kenngröße steht (Bremer & Sonnenschein 2012b).

Auf diese modifizierte Trainingsmenge kann nun das SVDD-Verfahren für die Erstellung

des Suchraummodells unverändert angewandt werden. Analog kann der Dekoder erstellt werden, der dann aber beliebige Featurevektoren  $x \in [0, 1]^{d+m}$  abbildet:

$$\begin{aligned} [0, 1]^{d+m} &\rightarrow \mathcal{F}_{[0,1]} \times [\mathcal{I}_{[0,1]}]^m \\ \gamma(x) &\mapsto (\mathbf{y}, \mathcal{I}_{[0,1]}^{Q_1}(\mathbf{y}, z_0), \dots, \mathcal{I}_{[0,1]}^{Q_m}(\mathbf{y}, z_0)). \end{aligned} \quad (10.3)$$

Der Vorteil bei diesem Ansatz ist, dass alle Verfahren für die Erstellung des Suchraummodells und des Dekoders unverändert verwendet werden können. Im Rahmen einer Wirkleistungsplanung kann ein beliebiger Vektor aus  $[0, 1]^d$  als Kandidat für einen Wirkleistungsfahrplan betrachtet werden. Wird dieser ergänzt um  $m$  Werte aus  $[0, 1]$  als angenommene Kennzahlenwerte (diese können auch falsch sein), so bildet der Dekoder ab auf einen Vektor mit einem durch die zugehörige Anlage realisierbaren Wirkleistungsverlauf und (näherungsweise) korrekten Kennzahlen für diesen Fahrplan. Hierbei sind zwei Aspekte zu beachten:

1. Die Kennzahlen, die durch den Dekoder erzeugt werden, bewerten den ebenfalls durch den Dekoder erzeugten, realisierbaren Fahrplan. Der Dekoder erzeugt keine Bewertung für den ursprünglich gegebenen Fahrplan. Da im Rahmen einer Optimierung aber auch der vom Dekoder modifizierte Fahrplan als Lösungskandidat betrachtet wird, ist dieses Verhalten wünschenswert. Es ist jedoch nicht möglich, Kennzahlen zu einem gegebenen Fahrplan zu bestimmen.
2. Die Belegung der Elemente des Featurevektors für die Kennzahlen vor der Abbildung durch den Dekoder kann prinzipiell beliebig gewählt werden. Jedoch hat die Wahl auch einen Einfluss auf den Wirkleistungsfahrplan, der durch den Dekoder erzeugt wird, da sowohl Leistungswerte als auch Kennzahlen immer gemeinsam durch das Modell abgebildet werden.

Nachfolgend werden zunächst einige Kennzahlen für die Evaluation definiert.

## 10.3. Verwendete Kennzahlen

Wie bereits diskutiert ist im praktischen Einsatz eine Vielzahl verschiedener Kennzahlen für unterschiedlichste Optimierungsziele denkbar. Im Rahmen der Evaluation des hier vorgestellten Ansatzes sollen konkret die nachfolgend vorgestellten Kennzahlen zum Einsatz kommen.

### 10.3.1. Produktionskosten

Ein unmittelbar nachvollziehbarer Kostenfaktor sind die monetären Kosten, die direkt durch das Abarbeiten eines Fahrplans durch eine Einheit entstehen. Es soll an dieser Stelle keine auf Grenzkosten basierende Preisgestaltung durch den Markt betrachtet werden. Vielmehr sollen die variablen Betriebskosten einer Einheit betrachtet werden, die als Kosten den Gewinn einer Koalition schmälern. Wenn wie in dem betrachteten Konzept die einzelnen Kosten aller Einheiten vom Erlös der Koalition abgezogen werden bevor der verbleibende Gewinn auf die einzelnen Einheiten aufgeteilt werden kann (Bremer

& Sonnenschein 2013c), sollten diese minimiert werden, um einen möglichst großen Gewinn verteilen zu können. Nicht alle Typen von dezentralen Energieanlagen haben variable Kosten, die direkt von einem konkreten Fahrplan abhängen. Stellvertretend seien hier zwei Beispiele für ein BHKW gegeben:

- **Brennstoffkosten** sind bei einem BHKW meist der dominierende Kostenfaktor. Die direkte Abhängigkeit der Höhe von einem konkreten Fahrplan und damit von der insgesamt verrichteten Arbeit ist evident. Zu berücksichtigen ist jedoch auch ein variabler Wirkungsgrad der Anlage. Dieser hängt stark vom konkreten Typ der Anlage ab. So variiert beispielsweise der Gesamtwirkungsgrad bei einem SOLO-Stirling BHKW zwischen 94.2% und 98.2% (Thomas 2007).
- **Spezifische Wartungskosten** sind zumindest bei modulierenden BHKW ebenfalls als durch den Fahrplan bestimmte variable Kosten zu betrachten. Bei Betrieb im Teillastbereich sinkt die mittlere erzeugte Leistung pro Betriebsstunde (Thomas 2007). Hierdurch erhöhen sich die spezifischen Wartungskosten. Im Falle eines ecoPOWER BHKW steigen laut (Thomas 2007) die Kosten von  $1.84 \frac{ct}{kW_{hel}}$  auf  $2.3 \frac{ct}{kW_{hel}}$  bei Absenken der Last von 100% auf 80%. Für ein SOLO-Stirling BHKW wurden ähnlich Werte berechnet. Dieselbe Energiemenge unter Vollast in einem kürzeren Zeitraum zu erzeugen ist in diesem Fall also günstiger.

Für die variablen spezifischen Wartungskosten sowie für die Brennstoffkosten liegen leider keine verwertbaren Daten vor. Für den Fall eines SOLO-Stirling BHKW finden sich in (Thomas 2007) Messstandsdaten für den Gesamtwirkungsgrad in Abhängigkeit von der aktuell erzeugten elektrischen Leistung. Vereinfachend soll o.B.d.A angenommen werden, dass diese Kennlinie auf die in dieser Arbeit zur Evaluation verwendeten BHKW übertragbar ist. Zur Überprüfung der Ergebnisse des Dekoders ist eine exakte Übereinstimmung des verwendeten Simulationsmodells mit dem Verhalten einer realen Anlage auch nicht erforderlich, da der Dekoder mit dem verwendeten Simulationsmodell verglichen wird, nicht das Simulationsmodell mit der Anlage. Nun kann ferner angenommen werden, dass die Brennstoffkosten umso geringer sind je effizienter das BHKW arbeitet. Daher wird die Effizienz als Ersatz für die Produktionskosten verwendet. Statt die Kosten zu minimieren kann auch die Gesamteffizienz maximiert werden. Zusätzliche Kosten durch Kaltstarts etc. werden hier zunächst nicht betrachtet.

Aus den Daten aus (Thomas 2007) lässt sich folgender Zusammenhang (für das Regressionsmodell siehe auch Abbildung 10.1) für die mittlere Gesamteffizienz des BHKW bei der Abarbeitung eines Fahrplans  $x$  ableiten:

$$\mathcal{I}_{[0,1]}^{(eff)}(x, z_0) = \frac{1}{d} \cdot \sum_{i=1}^d 0.0345x_i^5 - 0.4982x_i^4 + 2.6164x_i^3 - 6.0189x_i^2 + 6.0096x_i - 1.9783. \quad (10.4)$$

Der Startzustand  $z_0$  kennzeichnet hierbei einen beliebigen Betriebszustand mit ausreichender Betriebstemperatur des BHKW-Motors (Kaltstarts werden hier nicht betrachtet). Dieser Zusammenhang hat eigentlich nur sehr speziell für den in (Thomas 2007) durchgeführten Teststandslauf Gültigkeit, mangels weiterer Daten wird dieser Zusammenhang nachfolgend jedoch o.B.d.A für alle Anlagen verwendet. Da das hier vorgestellte Verfahren unabhängig von dem kodierten Kennzahlenzusammenhang ist und die Kennzahl als solche

nicht evaluiert werden soll, ist hiermit dennoch eine Aussage über die Integration dieser Kennzahl in Suchraummodell und Dekoder möglich.

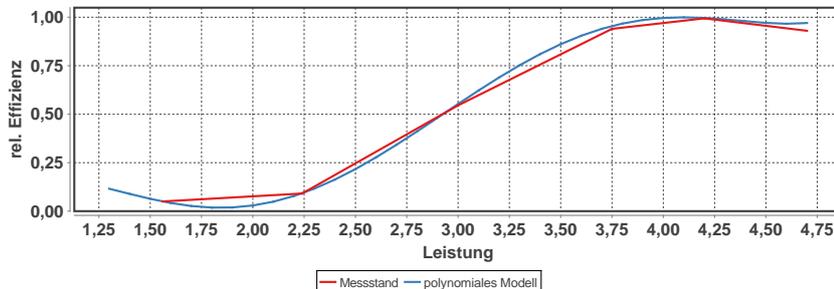


Abbildung 10.1.: Messstandsdaten aus (Thomas 2007) skaliert auf ein Ecopower BHKW mit 4.7 kW Nennleistung und die jeweils minimale und maximale absolute Effizienz sowie das hieraus abgeleitete polynomiale Modell.

### 10.3.2. Umweltkennzahlen

Umweltkennzahlen sind Maßzahlen für die Umweltauswirkungen (z. B.  $CO_2$  Emissionen, Wasser- oder Energiefußabdruck, o. ä.) unternehmerischen Handelns. Sie erfassen bisher üblicherweise die Umweltwirkung von Unternehmen, Organisationen oder Kommunen. So gesehen kennzeichnen sie die Performanz bei der Erreichung von Unternehmenszielen unter Berücksichtigung der Wirkung auf die Umwelt (Hřebíček et al. 2007). Üblicherweise soll durch die Erhebung von Umweltkennzahlen Wirkung in zweierlei Richtung erreicht werden: nach innen, um im Rahmen von Entscheidungsprozessen die Performanz zu verbessern und nach außen, um im Rahmen einer Berichterstattung die erreichte Performanz auch belegen zu können (Jamous & Müller 2013).

Bis dato (und insbesondere im Unternehmenskontext) werden solche Kennzahlen meist einmalig (z. B. jährlich und auch erst im Nachgang) betrachtet, wodurch die Anwendung solcher Kennzahlen begrenzt ist. Aktuell ist jedoch ein Wandel bei der Einschätzung der Bedeutung von Umweltkennzahlen zu beobachten und diverse Forschungsbemühungen (Rapp et al. 2011; Meyerholt et al. 2010) zielen darauf ab, Umweltkennzahlen auch für die strategische und täglich operationale Ebene einsetzbar zu machen (Teuteberg & Marx Gómez 2010).

Für Gruppen kleinerer Unternehmen mit gemeinsamen Zielen trotz unternehmerischer Unabhängigkeit wird die Möglichkeit eines gemeinsamen Berichtens bezüglich der Umweltperformanz aktuell in der Forschung diskutiert (Solsbach et al. 2011).

Bis dato gibt es auch kein allgemeines Beschreibungsmodell für den standardisierten Austausch von Umweltkennzahlen auf semantischer Ebene (Rapp & Bremer 2013). Für die Ebene von Unternehmen bzw. Organisationen wurde ein erstes auf einer speziell entwickelten Ontologie beruhendes Konzept für einen Beschreibungsstandard jedoch bereits vorgeschlagen (Löschner 2013). Eine mögliche Erweiterung für den Anwendungsfall Smart Grid wurde in (Bremer 2012) vorgeschlagen.

Welche Bedeutung hat der Einsatz von Umweltindikatoren in dem hier betrachteten

Szenario? Untersuchungen betrachten bisher üblicherweise die Umweltwirkung des Smart Grid als Ganzem (Huber et al. 2011), erneuerbarer Energien im Allgemeinen (Nitsch et al. 2004) oder aber die einer Technologie als solcher, unabhängig von konkreter Erzeugung einzelner Anlagen. Ein Beispiel für letzteres sind Untersuchungen zum Vogelschlag bei Windenergieanlagen (Peters et al. 2008). Effekte des Smart Grid als solches werden z. B. bei der Erhöhung des Anteils Erneuerbarer oder bei der Unterstützung von mehr Energiebewusstsein seitens der Verbraucher gesehen (Huber et al. 2011). Dies sind Effekte, die über die Vielzahl der Anlagen und über eine längere Zeitdauer gemittelt sind. Sie erlauben keine individuelle Verbesserung durch wiederholte Wahl jeweils kurzfristig bester Alternativen und sollen daher hier nicht betrachtet werden.

Im Kontext dieser Arbeit interessant sind Ansätze aus dem Bereich nachhaltiges Produktdesign (Müller 2013) bzw. aus dem Bereich der nachhaltigen Beschaffung (Dada 2013). Diese Anwendungsfälle für Umweltkennzahlen zielen darauf ab, ein Produkt schon im Entwurf so zu gestalten, dass spätere Umweltauswirkungen minimiert werden. Dies gilt dann auch für die Berücksichtigung von Lieferketten während der Produktion. Nun kann aber ein Virtuelles Kraftwerk oder auch ein dynamisch gebildeter Verbund von elektrischen Einheiten als Produzent eines Wirkleistungsproduktes (oder eines Systemdienstleistungsproduktes wie Regelleistung o. ä.) gesehen werden, sodass eine grundsätzliche konzeptionelle Verwandtschaft der Vorgänge gegeben ist. Produkt ist im Smart-Grid-Fall ein Lastgang, der aus Einzelteilen (Einzellastgänge) von einer Menge von Zulieferern zusammengesetzt wird. Ein Algorithmus hierfür kann sich ähnliche Fragen bzgl. umweltfreundlicherer Alternativen bezüglich der Einzelteile stellen wie die Beschaffungsabteilung bei der Produktplanung im Rahmen eines Designprozesses beim produktionsintegrierten Umweltschutz. Gehen Umweltindikatoren in die Bewertung eines Verbundes ein, so können auch verschiedene Zulieferer hinsichtlich ihrer Wirkung verglichen und bei der Beschaffung berücksichtigt werden. Eine ausführlichere Diskussion dieser Fragestellungen findet sich u. a. in (Bremer & Sonnenschein 2013d). Beispiele für Umweltkennzahlen zur Bewertung individueller Fahrplanalternativen sind

- Statische Verluste durch unvollständige Isolierung vorzeitig gefüllter Speicher,
- Umwandlungsverluste durch häufiges Füllen und Leeren eines Speichermediums, sowie
- Variable Effizienz durch unterschiedlich effektive Betriebsweisen.

Sie alle haben letztlich Einfluss auf den Primärenergieverbrauch und damit auf den Gesamtwirkungsgrad einer Anlage bezogen auf einen bestimmten Fahrplan. Im Rahmen dieser Arbeit sollen im Folgenden statische Energieverluste durch Warmwasserspeicher als Beispiel für eine Umweltkennzahl exemplarisch integriert werden. Weitere Wirkungen wie unterschiedlicher  $CO_2$ -Ausstoß durch Verwendung verschiedener Primärenergieträger werden hier nicht einbezogen.

Die Verluste  $\mathcal{I}_{[0,1]}^{(loss)}$  bestimmen sich anlagenindividuell aus den Isolierungseigenschaften des Pufferspeichers und dem Temperaturgradienten zwischen Pufferspeichertemperatur und Umgebung.

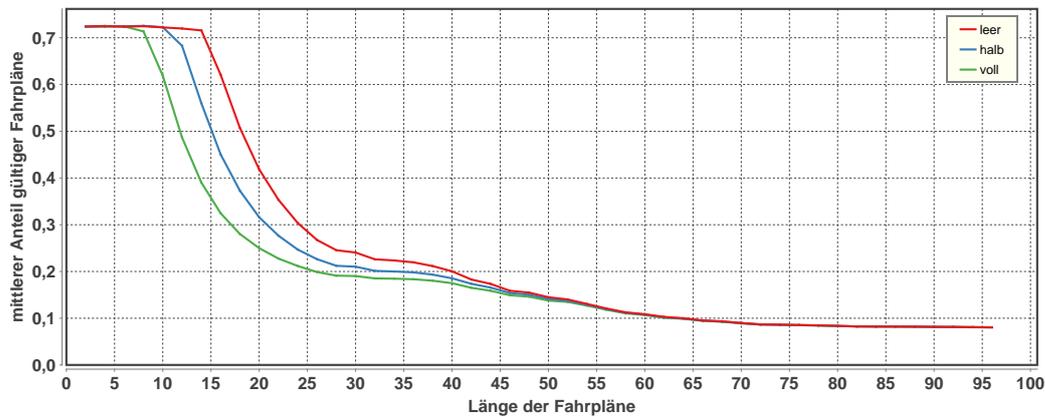


Abbildung 10.2.: Vergleich der Suchraumgrößen für BHKW mit unterschiedlich gefülltem Pufferspeicher. Dargestellt ist jeweils der mittlere Anteil der realisierbaren Fahrpläne je Zeitintervall für verschiedene Fahrplanlängen (jeweils der Mittelwert aus 50 Simulationen).

### 10.3.3. Erhalt von Freiheitsgraden

Da der hier skizzierte Planungsprozess in regelmäßigen Abständen (üblicherweise täglich) für den jeweils nachfolgenden Planungshorizont durchgeführt werden muss, ist es vorteilhaft bereits bei der aktuellen Planung zu berücksichtigen, dass für nachfolgende Vorgänge genügend Variabilität im Suchraum vorhanden bleibt. Hierfür ist zu beachten, in welchen Betriebszustand das Abarbeiten eines Fahrplans führt. Als Beispiel mag ein BHKW mit thermischem Pufferspeicher dienen.

Abbildung 10.2 zeigt wie die Anzahl der Möglichkeiten je Zeitintervall in Abhängigkeit vom anfänglichen Füllstand des Pufferspeichers verhält. Offensichtlich sind mit leerem Pufferspeicher eine größere Anzahl von Fahrplänen zu realisieren als mit einem vollen. Dies ist unmittelbar einsichtig, wenn im Wesentlichen Wärme produziert wird. Abbildung 10.2 zeigt jedoch auch, dass sich die Effekte für länger Zeiträume aufheben, sobald in den frühen Morgenstunden (um Zeitintervall 32 herum) größerer Wärmebedarf u. a. zum Duschen benötigt wird.

Größere Unterschiede lassen sich beispielsweise bei einer Redox-Flow-Batterie beobachten. Abbildung 10.3 zeigt die Möglichkeiten in Abhängigkeit vom initialen Ladezustand (state of charge: SOC); diesmal als absoluten Anteil realisierbarer Fahrpläne für den jeweiligen Zeithorizont. Wichtig ist hier insbesondere, dass im Gegensatz zum BHKW bei einer Batterie ein mittlerer Ladezustand als Ausgangslage zu einer Vergrößerung des Suchraums führt. Bei mittlerem Ladezustand sind sowohl Lade- als auch Entlade-Fahrpläne möglich.

Eine Kennzahl zur Bewertung des Erhalts von Freiheitsgraden ist also im Wesentlichen anlagenspezifisch und somit individuell für jeden Anlagentyp zu definieren. Für die Batterie bietet es sich wie gesehen an, auf einen mittleren Ladezustand abzielen. Es wird also eine Fehlerfunktion benötigt, welche die Abweichung des sich nach Abarbeitung eines Fahrplans einstellenden Ladezustands vom mittleren Ladezustand (50%, bzw. 0.5) bewertet. Hierzu kann beispielsweise folgende Fehlerfunktion verwendet werden:

$$\mathcal{I}_{[0,1]}^{(flex)}(\mathbf{x}, z_0) = e^{-20 \cdot (SOC_{batt}[d] - 0.5)^2}, \quad (10.5)$$

## 10. Modellierung von Kosten und Kennzahlen

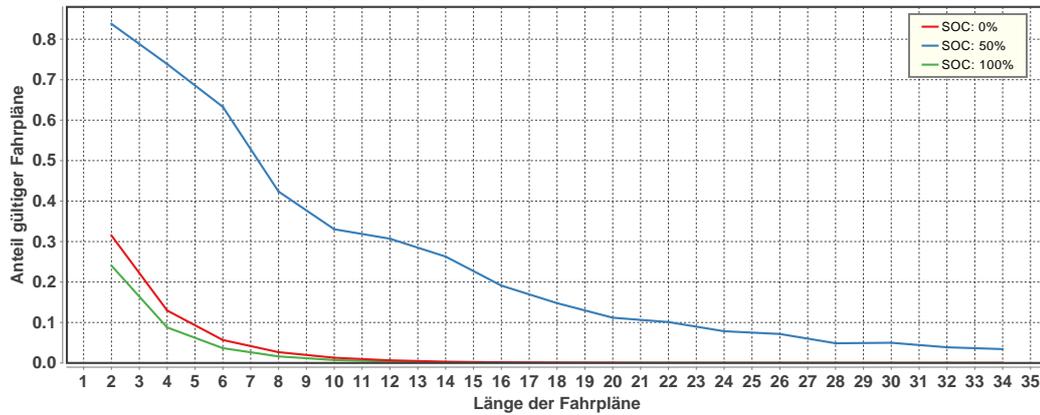


Abbildung 10.3.: Vergleich der Suchraumgrößen für Redox-Flow-Batterien mit unterschiedlichem Startladezustand (SOC). Dargestellt ist jeweils der Anteil der realisierbaren Fahrpläne an der Gesamtmenge der Fahrpläne für verschiedene Fahrplanlängen (jeweils der Mittelwert aus 50 Simulationen).

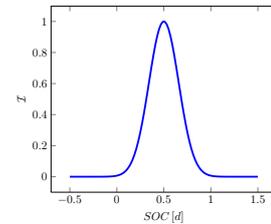


Abbildung 10.4: Indikatornäherungsfunktion für die Bewertung des Batterieendzustandes.

die in Abhängigkeit vom Ladezustand  $SOC_{batt}[d] \in [0, 1]$  nach Beendigung des letzten Zeitintervalls  $d$  zu dem die Ausführung des Fahrplans  $x \in \mathbb{R}^d$  ausgehend vom Startzustand  $z_0$  führt. Abbildung 10.4 zeigt die Fehlerfunktion. Der Faktor  $-20$ , der die Basisbreite der Kurve 10.4 bestimmt, wurde empirisch ermittelt.

Beim BHKW kann die Bewertung anhand des finalen Ladezustands (nach Beendigung von Zeitintervall  $d$ ) des Speichers vorgenommen werden:  $\mathcal{I}_{[0,1]}^{(flex)}(x, z_0) = 1 - SOC_{chp}[d]$ .

### 10.4. Simulationsergebnisse

Um das generelle Vorgehen bei der Kennzahlenintegration zu evaluieren und zu illustrieren, sollen an dieser Stelle zunächst einige Simulationsergebnisse auf anwendungsfallunabhängiger Ebene präsentiert werden.

Ein erstes Beispiel liefert Abbildung 10.5 aus (Bremer et al. 2011b). Hierfür wurden mit Hilfe eines BHKW-Modells analog zu den Constraintbeispielen aus Abschnitt 6.3.2 auf Seite 80 Trainingsfahrpläne für je zwei Zeitintervalle generiert. Als drittes Element wurde diesen Trainingsvektoren eine artifizielle Kennzahl hinzugefügt, die nach Bohachevsky's Testfunktion

$$f_{B_2}(x) = x_1^2 + 2x_2^2 - 0.3 \cdot \cos(3\pi x_1) \cdot \cos(4\pi x_2) + 0.3 \quad (10.6)$$

aus den Leistungswerten berechnet wurde. Die Kennzahl ist farblich codiert als Hintergrund in Abbildung 10.5 dargestellt. Diese artifizielle Kennzahl hat keinerlei Bezug zum Thema

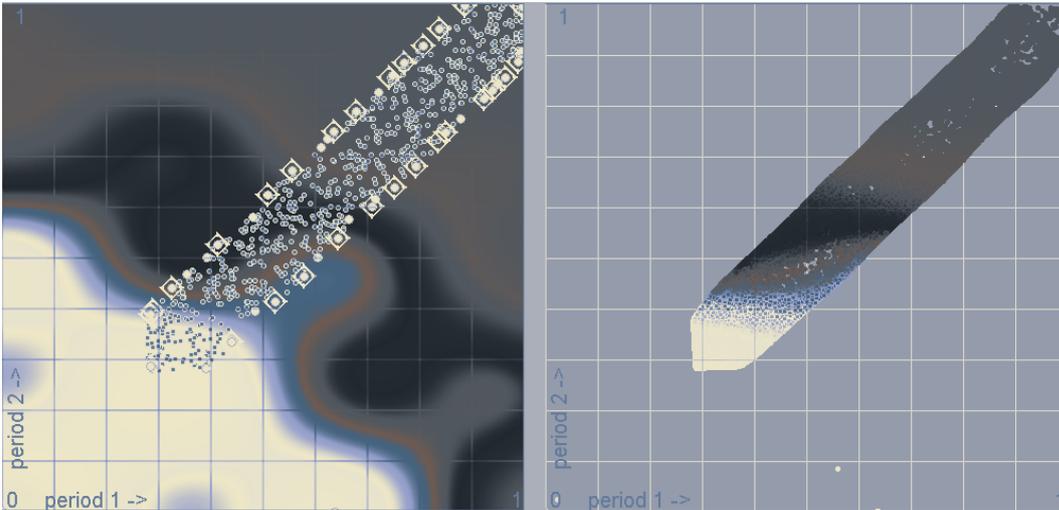


Abbildung 10.5.: Rekonstruktion einer künstlichen Kennzahl (verwendet wurde Bohachevsky's Funktion; Gleichung 10.6) aus (Bremer et al. 2011b), nachcoloriert.

Smart Grid, stellt aber eine Testkennzahl mit gewisser Komplexität dar. In der rechten Hälfte der Abbildung ist die Rekonstruktion durch den Dekoder dargestellt, die durch Erzeugen einer Menge von Zufallsvektoren  $x \in [0, 1]^3 \sim U(0, 1)^3$  und durch Abbilden dieser Zufallsvektoren mit Hilfe des zugehörigen Dekoders erstellt wurde. Nach dem Abbilden wurden die ersten beiden Elemente des Dekoder-Outputs als Leistungswerte und damit in der 2-dimensionalen Darstellung als Position interpretiert, das dritte Element entsprechend als Farbcode (Helligkeitswert).

Abbildung 10.6 zeigt analog zwei weitere Beispiele mit den Testfunktionen Ripple aus der CILib-Bibliothek (Cloete et al. 2008) und Salomon's Funktion (Salomon 1995) zusammen mit der bereits aus dem Kapitel 8 bekannten Doppelringstruktur.

Eine wichtige Frage im Zusammenhang mit der Technik zur Integration der Kennzahlen stellt sich im Zusammenhang mit der Anzahl der zusätzlich benötigten Supportvektoren zur Modellbeschreibung. Diese ist von Bedeutung für die benötigte Informationsmenge (und damit für den Speicherbedarf) für die Modellbeschreibung. Andererseits geht sie quadratisch (wegen der Matrix-Vektor-Multiplikation in Gleichung (9.5)) in die Berechnungskomplexität des Dekoders ein.

Abbildung 10.7 zeigt den durch eine bzw. zwei Kennzahlen verursachten Zuwachs in Prozent bei den benötigten Supportvektoren in Abhängigkeit von der Fahrplanlänge. Es ist unmittelbar einsichtig, dass dieser Zuwachs bei kurzen Fahrplänen stärker ausgeprägt ist, da die Kennzahlen hier einen größeren Anteil am gesamten Featurevektor haben. Bereits ab einer Fahrplanlänge von etwa 32 Zeitintervallen ist praktisch kein nennenswerter Zuwachs mehr festzustellen. Für die meisten praktischen Anwendungsfälle dürfte sich die Integration von Kennzahlen nicht merklich auf die Performanz des Gesamtsystems auswirken.

## 10. Modellierung von Kosten und Kennzahlen

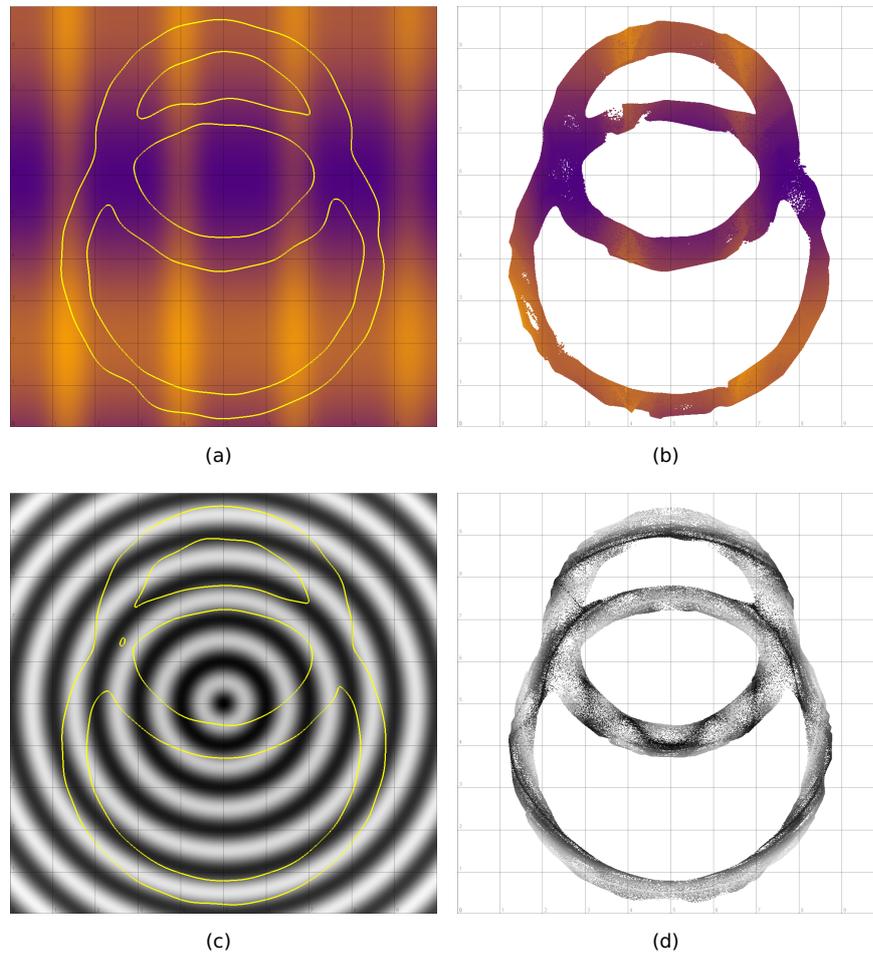


Abbildung 10.6.: Beispiele für die Rekonstruktion artifizierlicher Kennzahlen.

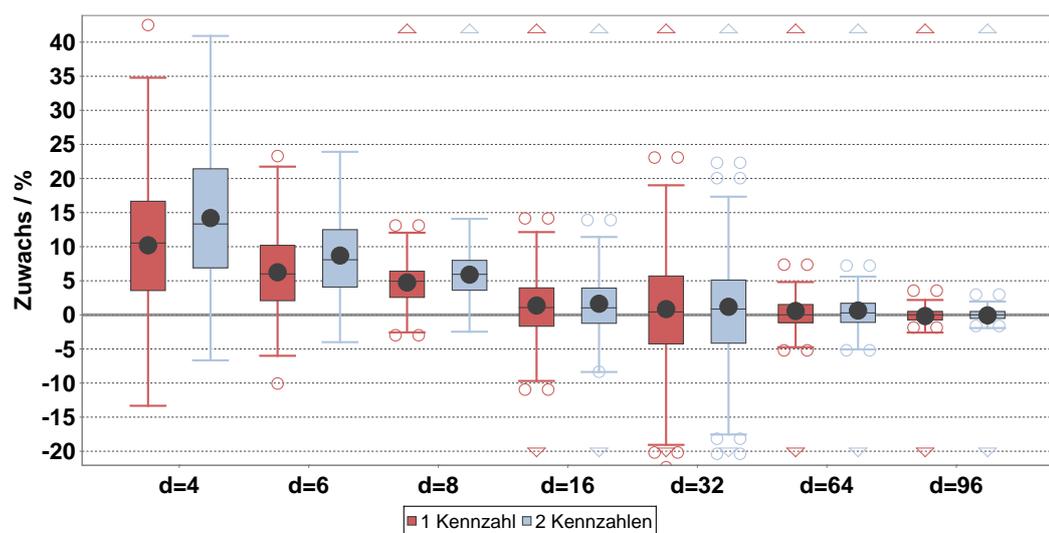


Abbildung 10.7.: Durch Kennzahlenintegration zusätzlich benötigte Supportvektoren.

## 10.5. Zusammenfassung

Mehrzieloptimierung bedarf einer adäquaten Bereitstellung von Kennzahlen zur Bewertung eines Lösungskandidaten hinsichtlich der individuellen Optimierungsziele. In dem hier verwendeten Smart Grid Anwendungsfall setzt sich ein Lösungskandidat aus vielen einzelnen Fahrplänen zusammen, die alle in die Bewertung hinsichtlich der verschiedenen Ziele einbezogen werden müssen. Einige Ziele ergeben sich unmittelbar (Minimierung der Differenz zwischen Summenlastgang und Produkt), für andere ist eine differenzierte Bewertung bezüglich einzelner Fahrpläne notwendig. Die Erfüllung hinsichtlich einiger Ziele kann durch einfache Aggregation der Bewertung einzelner Fahrpläne beurteilt werden (z. B. Produktionskosten), andere erfordern komplexere Betrachtungen (wenn beispielsweise ein Ökocontrolling gewünscht ist). In jedem Fall ist es notwendig, die Annotation mit Kennzahlen auf Fahrplanebene vorzunehmen, da jeder Fahrplan individuellen Einfluss auf die Bewertung hat.

Der Modellierungsansatz zur Beschreibung eingeschränkter Suchräume mithilfe des SVDD-Verfahrens bietet einen eleganten Weg, zusätzliche Informationen für die individuelle Annotation einzelner Alternativen in das bis dato beschriebene Suchraummodell einzubinden. Durch einfache Erweiterung der Dimensionalität der Trainingsdaten kann dies bewerkstelligt werden, ohne an dem eigentlichen Verfahren etwas ändern zu müssen.

Den großen Vorteil spielt diese Vorgehensweise im Zusammenspiel mit dem Dekoder aus. Während das Suchraummodell als Klassifikator lediglich in der Lage ist, die korrekte Zuweisung von Indikatoren zu gegebenen Fahrplänen zu überprüfen, können mit Hilfe des Dekoders die Kennzahlen zu den Fahrplänen auch automatisiert rekonstruiert werden. In dieser Weise können die zugeordneten Werte von einem Planungsalgorithmus in die Lösungsbewertung einbezogen werden, wenn eine zulässige Lösung generiert wird. Hierdurch wird einem Planungsalgorithmus nicht nur der Hinweis gegeben wo nach gültigen Lösungen zu suchen ist, sondern auch wo gute Lösungen zu finden sind.



**Teil III.**

# **Anwendung und Evaluation**



# 11. Evaluation

*What happens if a big asteroid hits Earth?  
Judging from realistic simulations involving a sledge hammer and  
a common laboratory frog,  
we can assume it will be pretty bad.*

---

*(Dave Barry)*

## 11.1. Einleitung

In diesem und den nachfolgenden Kapiteln soll nunmehr die Tauglichkeit des hier vorgestellten Konzepts anhand von Implementierungsartefakten untersucht werden. Ziel wird es hier nicht sein, die Tauglichkeit verschiedener Optimierungsverfahren für das Lösen des Wirkleistungsplanungsproblems zu betrachten oder zu bewerten. Ziel wird es vielmehr sein, zu zeigen, dass das Konzept mit der Umsetzung als Supportvektor-Dekoder in Kombination mit bekannten Standardverfahren für die Optimierung realisierbar ist und als Ganzes einen möglichen Lösungsansatz darstellt. Es wird gezeigt, wie verschiedene Standard-Optimierungsverfahren durch den Dekoderansatz und die automatisierte Transformation des Optimierungsproblems in die Lage versetzt werden, trotz der durch unbekannte Constraints eingeschränkten Suchräume zu einer gültigen Lösung zu kommen, obwohl diese Verfahren von Haus aus keine Constraintintegration vorgesehen haben.

In Abschnitt [6.5 auf Seite 86](#) wurden bereits die besonderen Schwierigkeiten durch die verteilten Informationen über die Struktur der Lösungsräume diskutiert. An dieser Stelle sollen zunächst noch einige Betrachtungen bezüglich der Komplexität des Optimierungsproblems durchgeführt werden, bevor konkrete Umsetzungen des Optimierungsansatzes vorgestellt werden.

Aufgrund einer Relaxation diskreter Fahrplanräume (vgl. Abschnitt [11.3.3 auf Seite 159](#)) ist das Argument der NP-Vollständigkeit bei der constraintfreien Formulierung im Allgemeinen nicht mehr zutreffend. Das Problem wird so transformiert, dass

1. der Lösungsraum nun kontinuierlich ist und
2. keine Constraints mehr zu beachten sind.

Das transformierte Problem kann als nicht-lineares Programm (NLP) betrachtet werden (vgl. auch (Bremer & Sonnenschein [2013b](#))). Auch das Finden des globalen Optimums eines NLP kann NP-vollständig sein (Murty & Kabadi [1987](#)). Wichtiger für die Umsetzung in einen konkreten Optimierungsansatz ist jedoch eine gewisse Vorstellung von der Struktur der Fehler- (bzw. je nach Definition der Fitness-) -funktion über die optimiert wird.

Die Größe des zu durchsuchenden Lösungsraums ergibt sich durch die zu erwartende

Anzahl der Anlagen in einem Verbund und durch die Länge (und die zeitliche Auflösung) der Planungszeiträume. Ausgehend von der derzeitigen betrieblichen Übung im heutigen Energiesystem kann von 96-dimensionalen Fahrplänen mit Viertelstundenauflösung ausgegangen werden. Da Systeme mit sich dynamisch rekonfigurierenden Verbänden bisher nicht existieren, ist es schwierig eine Aussage über die Anzahl der beteiligten Anlagen zu machen. Derartige Systeme sind derzeit Gegenstand der wissenschaftlichen Forschung. Als Beispiel sei hier der Forschungsverbund Smart Nord (<http://smartnord.de>) genannt, der auch als Anregung für die Definition von Testszenarien für diese Arbeit dient.

Für die konkrete Lösung der Wirkleistungsplanung muss später noch eine Auswahl von geeigneten Optimierungsansätzen getroffen werden. Hier können entweder exakte Verfahren oder Näherungsverfahren zum Einsatz kommen. Es ist zu vermuten, dass exakte Verfahren bei dem vorliegenden Problem nicht performant zu einem Ergebnis führen werden. Insbesondere kann vermutet werden, dass die Fehler- oder Fitnessfunktion deren Optimum gesucht wird, durch die Transformation in einen constraintfreien Suchraum unter Umständen schwieriger zu durchsuchen sein wird. Es besteht die Möglichkeit, dass hierdurch zusätzliche lokale Optima eingeführt werden.

Um Aussagen über die Struktur der Lösungsräume des transformierten Optimierungsproblems und der Fehlerfunktion treffen zu können, sollen nachfolgend zunächst verschiedene Betrachtungen aus der Fitnesslandschaftanalyse (engl. Fitness Landscape Analysis) durchgeführt werden. Für die statistische Analyse der Struktur von Fitnesslandschaften sind verschiedene Techniken bekannt (Weinberger 1990; Hordijk 1996; Vassilev et al. 2000), von denen hier drei betrachtet werden, um die Notwendigkeit der Verwendung von Heuristiken zu begründen.

Nach dieser Betrachtung der Problemkomplexität werden in diesem Kapitel zunächst die verwendeten Ansätze und Szenarien sowie die Modellierungen der beteiligten Anlagen vorgestellt. Abschließend werden die Komponenten des Suchraummodell (Klassifikator und Dekoder) in ausgewählten Einzelszenarien getestet, bevor die Integration des Gesamtkonzeptes anhand ausgewählter Optimierungsansätze in den nachfolgenden Kapiteln zur zentralen bzw. dezentralen Optimierung betrachtet wird.

### 11.1.1. Strukturanalyse

Eine Fitnesslandschaft  $\mathcal{L}$  definiert sich über den Suchraum, der alle Lösungen enthält, eine Nachbarschaftsbeziehung zwischen den Lösungen und eine Fitnessfunktion  $\nu$ , welche jeder Lösung eine Bewertung zuordnet. Trägt man die Lösungsgüte als räumliche Dimension über dem Suchraum auf, so ergibt sich im übertragenen Sinne das Bild einer Landschaft auf der sich die Lösungen beispielsweise eines evolutionären Algorithmus auf der Suche nach Optima (Täler oder Berge je nach Definition des Optimierungsziels und der Bewertungsfunktion als Fitness oder Fehlergröße) im Rahmen der Suche bewegen. Die Struktur dieser Landschaft kann nun beispielsweise analysiert werden, indem die Korrelation benachbarter Punkte, die von den Unterschieden in der Bewertung abhängt, untersucht wird (Vassilev et al. 2000). Flache Regionen mit wenig Steigung weisen eine stärkere Korrelation benachbarter Punkte auf als zerklüftete. Die Korrelation dient somit als Maß für die Rauheit der Fitnesslandschaft. Nach Weinberger (1990) kann hierzu die

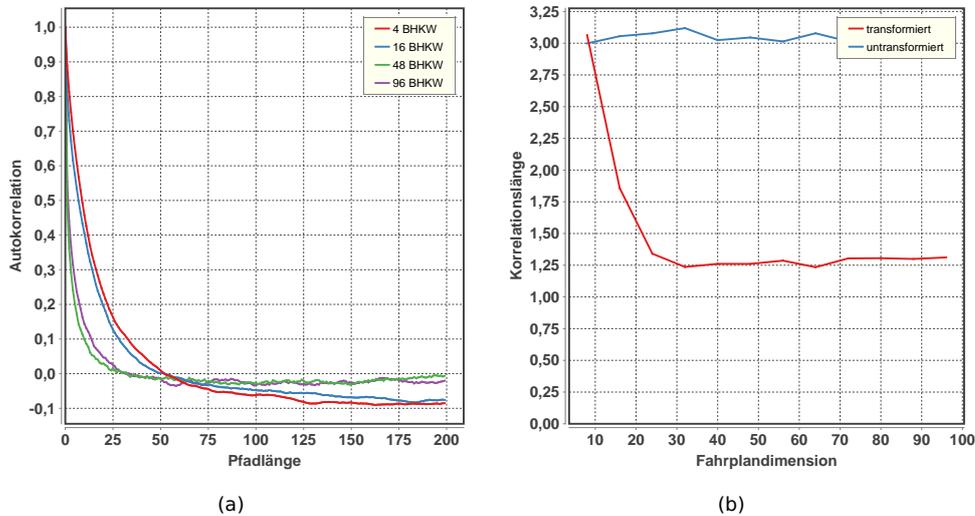


Abbildung 11.1.: Korrelationscharakteristik der Fitnesslandschaften: Autokorrelation der Pfade für verschiedene Problemgrößen (Anzahl beteiligter Anlagen) 11.1(a) und ein Vergleich der Korrelationslängen von transformiertem und nicht transformiertem Wirkleistungsplanungsproblem für unterschiedlich dimensionierte Planungshorizonte (Fahrplanlängen) 11.1(b).

Autokorrelation eines zufällig gewählten Pfades (Random Walk) auf der Fitnesslandschaft betrachtet werden.

Sei  $X = (x_{ij})$  eine Lösung des Wirkleistungsplanungsproblems, wobei  $x_{ij}$  die mittlere skalierte Wirkleistung von Anlage  $i$  im Zeitintervall  $j$  bezeichne. Sei  $\{f_t\}_{t=1}^n$  eine Sequenz von  $n$  Fitnesswerten, die wie folgt gesampelt werden: ausgehend von einer zufälligen Startlösung  $X_1$  werden nachfolgende, benachbarte Lösungen  $X_{t+1}$  nach (Merkuryeva & Bolshakovs 2011) aus  $X_t$  generiert indem jeder Leistungswert mit einer Wahrscheinlichkeit von je  $1/3$  um  $0.1$  erhöht, erniedrigt oder unverändert gelassen wird sofern sie dadurch den zulässigen Lösungsraum nicht verlässt. Es wird sichergestellt, dass mindestens eine Änderung stattfindet. Die Reihe  $F = \{f_t\}_{t=1}^n$  enthält dann die Werte  $f_t = \nu(X_t)$ , mit  $\nu(X) = \delta(\sum_{x_i \in X} x, z)$  im Falle der Wirkleistungsplanung ohne Mehrzieloptimierung (vgl. Gleichung 6.14 auf Seite 88). Nun kann die Autokorrelation

$$\rho(\sigma) = \frac{E[f_t f_{t+\sigma}] - E[f_t]E[f_{t+\sigma}]}{V[f_t]} \quad (11.1)$$

für die Pfadlänge  $\sigma$  betrachtet werden, wobei  $E[f_t]$  und  $V[f_t]$  Erwartungswert und Varianz der Reihe bezeichnen. Weinberger (1990) definiert zudem die Korrelationslänge

$$\tau = -\frac{1}{\ln(\rho(1))}.$$

Abbildung 11.1 zeigt einige Versuchsreihen zur Korrelationscharakteristik des in dieser Arbeit betrachteten Wirkleistungsplanungsproblems. In Abbildung 11.1(a) wird die Stärke der Korrelation für verschieden weit entfernte Punkte der Fitnesslandschaft miteinander verglichen. Hierbei werden verschiedene Problemgrößen betrachtet, die sich durch die unterschiedliche Anzahl beteiligter Anlagen ergeben. Das schnelle Abfallen der

Autokorrelationsfunktionen auch bereits bei kleinen Problemgrößen lässt auf zerklüftete Fitnesslandschaften mit vielen lokalen Optima schließen (Vassilev et al. 2000; Merkurjeva & Bolshakovs 2011). Der Vergleich der Korrelationslängen, also dem mittlerer Abstand auf einem Pfad ab dem der Großteil der Punkte nicht mehr korreliert ist (Vassilev et al. 2000), zeigt zudem, dass durch die Transformation des Wirkleistungsplanungsproblems in eine constraintfreie Formulierung eine größere Zerklüftung der Fitnesslandschaft entsteht. Dieser Effekt ist aber nicht unerwartet. Anzumerken ist noch, dass alle Kennzahlen nicht für einzelne Pfade bestimmt wurden, sondern über jeweils 1000 zufällige Pfade gemittelt wurden. Dies gilt auch für die nachfolgenden Betrachtungen.

### 11.1.2. Informationsanalyse

Die Korrelationsanalyse gibt zunächst eine vage Vorstellung von der Struktur der Fitnesslandschaft. In (Vassilev et al. 2000) wird eine erweiterte Analyse basierend auf dem Informationsgehalt als entropisches Maß der Zeitreihe  $\{f_t\}_{t=1}^n$  vorgestellt, welche einen schrittweisen, zufälligen Weg durch die Fitnesslandschaft nimmt. Basierend auf dem Entropiebegriff von Shannon (1948) und der algorithmischen Informationstheorie (Chaitin 1987) wird in (Vassilev et al. 2000) eine Charakterisierung der Verteilung und der Anzahl der Optima auf einem zufälligen Weg über die Landschaft erarbeitet. Diese Konzepte über den Informationsgehalt (als Maß für die Komplexität der Landschaft) und den partiellen Informationsgehalt (als Maß für die Modalität der Bewertungsfunktion) werden im Folgenden auf Instanzen des Wirkleistungsproblems exemplarisch angewandt, um eine Abschätzung der Komplexität des Problems zu bekommen. Die nachfolgenden Kennzahlen basieren darauf, den Pfad durch die Fitnesslandschaft als Ensemble von Basiselementen zu betrachten. Drei Klassen werden betrachtet: flache Gebiete (benachbarte Punkte mit ähnlicher Fitness), isolierte Punkte (umgeben ausschließlich von besseren oder schlechteren Punkten) und Punkte, die weder isoliert noch flach sind. Zu diesem Zweck wird zunächst jeder Pfad in eine Abfolge von Token  $S(\epsilon) = s_1 s_2 \dots s_n$  aus dem Alphabet  $\{\bar{1}, 0, 1\}$  umgewandelt durch

$$S_i = \psi_{f_i}(i, \epsilon) = \begin{cases} \bar{1}, & \text{falls } f_i - f_{i-1} < -\epsilon \\ 0, & \text{falls } |f_i - f_{i-1}| \leq \epsilon \\ 1, & \text{falls } f_i - f_{i-1} > \epsilon \end{cases} \quad (11.2)$$

für ein gegebenes  $\epsilon \in [0, \max f_i]$  (vgl. Vassilev et al. 2000)). Der String  $S(\epsilon)$  enthält nun Informationen über die Struktur der Fitnesslandschaft entlang des zufällig gewählten Pfades. Ein Objekt entlang des Pfades wird nun durch zwei aufeinanderfolgende Buchstaben des Strings definiert. So kennzeichnet beispielsweise die Sequenz  $\bar{1}1$  den Wechsel von Gefälle zu Steigung und somit eine Senke. Das Entropiemaß für dieses Ensemble von Objekten lässt sich wie folgt bestimmen (nach Vassilev et al. 2000):

$$H(\epsilon) = - \sum_{p \neq q} P_{[pq]} \log_6 P_{[pq]}, \quad (11.3)$$

mit der Häufigkeit  $P_{[pq]}$  des Auftretens der Sequenz  $pq$  in  $S(\epsilon)$ .

Von Bedeutung für die Lösungskomplexität ist auch die Modalität der Zielfunktion, also die Anzahl (lokaler) Optima eines Pfades. Diese kann ebenfalls aus dem String  $S(\epsilon)$

Tabelle 11.1.: Komplexitätsbetrachtungen für drei ausgewählte Problemszenarien.

Szenario	Formulierung	Entropie	part. Inform.	# Optima	Stabilität
Szenario 1	untransformiert	0.773 ± 0.034	0.410 ± 0.039	40.460 ± 03.908	0.124 ± 0.026
	transformiert	0.705 ± 0.049	0.528 ± 0.034	52.328 ± 3.393	0.299 ± 0.080
Szenario 2	untransformiert	0.752 ± 0.041	0.427 ± 0.039	42.148 ± 3.920	0.145 ± 0.035
	transformiert	0.773 ± 0.032	0.435 ± 0.039	43.020 ± 3.857	0.207 ± 0.050
Szenario 3	untransformiert	0.792 ± 0.028	0.381 ± 0.037	37.566 ± 3.643	0.102 ± 0.018
	transformiert	0.349 ± 0.122	0.107 ± 0.051	10.424 ± 5.058	0.928 ± 0.101

herausgelesen werden. Im Gegensatz zur Entropie, welche die Diversität der Objekte auf einem Pfad misst, muss für die Bestimmung eines Maßes für die Modalität eine Klassifikation der Objekte auf dem Pfad vorgenommen werden, um die Anzahl der Optima entlang des Pfades zu bestimmen. Zunächst wird als Maß der partielle Informationsgehalt bestimmt. Nach (Vassilev et al. 2000) wird hierfür der String  $S(\epsilon)$  umgewandelt in einen String  $S'(\epsilon) = o_1 o_2 \dots o_\mu$  aus dem Alphabet  $\{\bar{1}, 1\}$ , sodass sich der kürzeste String ergibt, welcher die Wechsel von Steigung und Gefälle entlang des Pfades repräsentiert. Vassilev et al. (2000) gibt hierfür eine rekursive Vorschrift. Der partielle Informationsgehalt ist nun definiert als

$$M(\epsilon) = \frac{\mu}{n} \in [0, 1], \quad (11.4)$$

wobei ein Wert von 1 eine maximale Modalität kennzeichnet. Die absolute Anzahl an Optima (für ein gegebenes  $\epsilon$ ) entlang eines Pfades lässt sich bestimmen als

$$\lfloor \frac{n \cdot M(\epsilon)}{2} \rfloor. \quad (11.5)$$

All diese Betrachtungen sind sehr sensitiv gegenüber der Wahl des Parameters  $\epsilon$ . Kleine Werte führen zu einer großen Sensitivität gegenüber Unterschieden in benachbarten Fitnesswerten. Der kleinste Wert für  $\epsilon$  bei dem keine Unterschiede mehr festgestellt werden und bei dem somit die Fitnessfunktion vollständig flach wird, wird als Informationsstabilität bezeichnet (Vassilev et al. 2000). Auch diese Informationsstabilität lässt sich mit der maximalen Fitnessdifferenz in einem Pfad skalieren, um vergleichbare prozentuale Werte zu bekommen.

Tabelle 11.1 zeigt beispielhaft die Ergebnisse für drei Szenarien. Szenario 1 umfasst eine Wirkleistungsplanung für einen Verbund aus 10 BHKW, die alle unterschiedliche Gradienten fahren können, 10 Boilern unterschiedlicher Leistungsklasse und 10 Wärmepumpen mit unterschiedlichen Temperaturhaltevorgaben. Szenario 2 besteht aus 10 BHKW, 40 Boilern und 10 Redoxflow-Batterien. Szenario 1 und 2 umfassen je einen Zeitrahmen von zwei Stunden. Szenario 3 enthält lediglich eine Redoxflow-Batterie für ein Ganztagesprodukt. Alle Kennzahlen wurden mit einem sehr sensitiven Wert für  $\epsilon$  von 1% der maximalen Fitnessdifferenz bestimmt.

Wie Tabelle 11.1 zeigt, hängt es auch von dem gewählten Szenario ab, ob die Komplexität (Modalität) der Zielfunktion durch die Transformation zu einer constraintfreien Formulierung zunimmt oder nicht. In jedem Fall sind die Zuwächse eher marginal (wenngleich vorhanden), sodass sie den Vorteil durch die Constraintfreiheit im Allgemeinen nicht durch einen zu großen Komplexitätszuwachs aufzehren.

Sowohl Struktur als auch Größe des (zusammengesetzten) Gesamtlösungsraums des Optimierungsproblems, das in allen Teillösungsräumen der Anlagen eines Verbundes gleichzeitig suchen muss, lassen die Verwendung von Heuristiken (insbesondere mit

Techniken zum Entkommen aus lokalen Optima) für die Bestimmung guter Näherungslösungen vielversprechend erscheinen. Die Bestimmung einer exakten Lösung wird im allgemeinen Fall nicht effizient berechenbar sein. Zudem beruhen die Transformation des Lösungsraums und auch schon die Modellierung der Anlagen selbst auf Prognosen und Näherungen, sodass durch kein Verfahren eine exakte globale Lösung garantiert werden kann.

Die Lösung des in dieser Arbeit betrachteten Optimierungsproblems mit exakten Verfahren dürfte allein aufgrund der Größe des Suchraums (insbesondere bei größeren Verbänden) im Allgemeinen nicht effizient lösbar sein. Deshalb und aufgrund der zerklüfteten Struktur der Fehlerfunktion mit einer Vielzahl an lokalen Optima in denen ein Optimierungsverfahren steckenbleiben kann, werden im Rahmen der Evaluation dieser Arbeit daher Heuristiken für eine näherungsweise Lösung herangezogen.

### 11.2. Verwendete Evaluationsansätze

Die Evaluation des in dieser Arbeit vorgestellten Konzepts und dessen Anwendbarkeit für die Wirkleistungsplanung wird anhand von Simulationen durchgeführt. Zur Anwendung kommt hier in erster Linie ein Design-Science-Ansatz. In diesem Forschungskonzept werden Artefakte geschaffen, die klar definierte organisatorische Teilprobleme lösen und einzeln untersucht werden können (Hevner et al. 2004). Bei solchen Artefakten handelt es sich um strukturierte Entitäten, die u. a. auch Software-Prototypen, rein mathematische Modellierungen oder Konstrukte der Logik umfassen. Als wichtiger Bestandteil des Design Science sind diese Artefakte Gegenstand rigoroser Evaluation durch sowohl analytische Methoden als auch auf funktionaler Ebene. Da eine Anbindung an reale Anlagen nicht zur Verfügung steht, kommen diese Artefakte zunächst im Rahmen von Simulationsstudien zum Einsatz.

Anleihen zum Vorgehen auf konzeptioneller Ebene stammen auch aus dem Prozessmodell zum Smart Grid Algorithm Engineering (Nieße et al. 2013). Die hierin vorgeschlagene starke Einbindung von Domänenwissen über das Smart Grid in die Problemdefinition während einer vorgeschalteten Konzeptionsphase wird durch die Projekteinbindung in Smart Nord eingebracht; ebenso wie der anschließende, evolutionäre Designzyklus. Experimente werden auch im Rahmen dieser Arbeit mittels Simulationen basierend auf einer Bibliothek wiederverwendbarer Anlagensimulationskomponenten (vgl. Nieße et al. 2013) durchgeführt.

Als Teilprobleme des Konzepts für zu erstellende Artefakte bieten sich die Komponenten für Sampling (mathematisches Konzept), Klassifikationsmodell und Dekoder (Softwareartefakte) an (vgl. Abschnitt 11.4 auf Seite 160). Das Sampling wurde analytisch bereits in Abschnitt 8.2 auf Seite 102 betrachtet. Auf funktionaler Ebene ist es Bestandteil der Modellerstellung. Modell und Dekoder werden nachfolgend (Abschnitt 11.4 auf Seite 160) getestet. Die Integration als Gesamtkonzept wird anhand von Simulationsszenarien evaluiert, welche nach Verbundbildung und Produktzuweisung ansetzen. Betrachtet wird also nicht der gesamte Ablauf, wie er in Abschnitt 6.5 auf Seite 86 vorgestellt wurde, sondern vornehmlich der Optimierungsteil. Da es nicht das Ziel ist, die Optimierungsverfahren selbst zu bewerten, wird die Integration mit verschiedenen Optimierungsverfahren un-

tersucht, ohne die erreichten Optimierungsergebnisse zu bewerten oder hinsichtlich der Verfahren zu vergleichen. Hierfür wären insbesondere noch ausführliche und für jedes Verfahren individuelle Parameterstudien erforderlich.

Die betrachteten Szenarien werden im nachfolgenden Abschnitt zunächst definiert.

## 11.3. Modellierung der Szenarien

Derzeit ist kein in der Praxis eingesetztes System bekannt, welches eine Wirkleistungsplanung marktorientiert und mit dynamisch gebildeten Verbänden vornimmt. Wie die Zusammensetzung solcher Verbände in der Praxis aussähe ist Gegenstand aktueller Forschung und somit heute noch Spekulation. Gleiches gilt für die Ausgestaltung zukünftiger Wirkleistungsprodukte. Andererseits ist die Kenntnis der konkreten Zusammensetzung der Verbände oder der Produktgestaltung auch nicht notwendig, um das in dieser Arbeit vorgeschlagene Verfahren testen zu können. Schließlich ist gerade die Unabhängigkeit von Anlagentyp, Constraints, Verbundzusammensetzung, Optimierungsziel oder Bewertungssystem eine Stärke des Verfahrens. Daher werden für die Evaluation der Anwendbarkeit artifizielle Szenarien verwendet.

### 11.3.1. Szenarien

Für die in der Optimierung betrachteten Szenarien werden zwei Arten von Verbund einbezogen: reine Erzeugerverbände bestehend aus einer heterogenen Gruppe von unterschiedlich spezifizierten BHKW sowie Verbände mit einem Anteil an Verbrauchern zur Produktveredelung. Somit ist es auch bei gemischten Verbänden das Ziel, einen gewünschten Produktlastgang nachzufahren. Es ist nicht das Ziel, verbundintern einen Lastausgleich herzustellen. Derartige Szenarien wären mit dem in dieser Arbeit vorgestellten Supportvektor-Dekoder-Ansatz zwar umsetzbar, werden hier jedoch nicht betrachtet. Für die gemischten Verbände wird empirisch folgende Zusammensetzung festgesetzt: 70% BHKW, 20% Wärmepumpen und 10% Warmwasserbereiter. Alle Anlagen werden jeweils mit für die gesamte Parametrisierung (vgl. auch den nachfolgenden Abschnitt über die Anlagenmodelle) zufällig gewählten Startbelegungen initialisiert. Als mögliche Produkte werden zwei Varianten betrachtet:

1. Als nachgefragtes Wirkleistungsprofil für die Erzeugung (auch für gemischte Verbände mit Verbrauchern zur Produktveredelung) werden Standardlastprofile (ggf. nicht für den ganzen Tag) von (Stromnetz Hamburg 2014) verwendet. Es wird jeweils ein Tag des Jahres 2014 zufällig ausgewählt. Das Profil wird von der Größe her so skaliert, dass die erforderliche Gesamtenergiemenge dem Leistungsvermögen des Verbundes entspricht. Es ist jedoch unwahrscheinlich dass der Verbund das Profil über den gesamten, betrachteten Zeitraum exakt abfahren kann.
2. Um ein Produkt mit bekanntem Optimum zu haben wird ein artifizielles Produkt verwendet, welches durch den Verbund garantiert exakt gefahren werden könnte. Dieses wird erzeugt durch Aufsummieren zufällig generierter Anlagenfahrpläne (vom

jeweiligen Simulationsmodell durch Sampling erzeugt, aber nicht im Trainings-sample vorhanden).

Als Verbundgröße werden außer in Sonderfällen Verbünde mit 10, 100 bzw. 1000 Anlagen betrachtet.

Diese Szenarienauswahl umfasst Szenarien mit verschiedenen Anlagentypen, die sowohl reine Erzeugung als auch die gleichzeitige Optimierung von Erzeugung und Verbrauch umfassen. Auf diese Weise wird es möglich, das Verfahren für Fahrpläne nicht nur mit unterschiedlichem Leistungsspektrum sondern auch mit unterschiedlichem Vorzeichen für Erzeugung bzw. Verbrauch gleichzeitig zu benutzen, um die Integrationsfähigkeit für diese Anlagentypen zu zeigen. Die genaue Zusammensetzung eines Verbundes ist für die Evaluation des Supportvektor-Dekoder-Verfahrens nicht von Belang, da das Konzept die Integrierbarkeit prinzipiell beliebiger Anlagen vorsieht (vgl. Voraussetzungen in Abschnitt [7.3 auf Seite 96](#)). Die Beschränkung erfolgt daher auf vorhandene Simulationsmodelle. Aussagen über die Eigenschaften speziell zusammengesetzter Verbünde sollen hier nicht getroffen werden.

Um die Skalierbarkeit des Verfahrens zu betrachten, müssen Verbünde unterschiedlicher Größe simuliert werden. Zuverlässige Aussagen über die zu erwartenden Größen (Anlagenzahl) im Falle dynamischer Verbünde sind derzeit nicht möglich. Daher wurden 1000 Anlagen als pragmatische Obergrenze festgesetzt, da auch bei dezentralen Optimierungsverfahren der gesamte Prozess auf einem einzelnen Arbeitsplatzrechner noch simuliert werden soll.

Somit deckt der Pool an betrachteten Szenarien alle wesentlichen Aspekte zur Beurteilung der Durchführbarkeit der Wirkleistungsplanung unter Zuhilfenahme des in dieser Arbeit vorgeschlagenen Konzeptes ab. Allgemeine Aussagen über erreichbare Lösungsgüten, günstigste Zusammensetzungen von Verbänden o. ä. sind nicht Gegenstand der hier durchgeführten Betrachtungen.

### 11.3.2. Anlagenmodelle

Für die Umsetzung der zuvor beschriebenen Szenarien in eine simulative Untersuchung werden verschiedene Anlagenmodelle benötigt. Diese sollen hier kurz beschrieben werden.

#### **BHKW**

Ein Anlagentyp von dem erwartet werden darf, eine wichtige Rolle in den Verbänden zu spielen, ist die Gruppe der Blockheizkraftwerke, die bereits Gegenstand verschiedener Forschungsarbeiten auf dem Gebiet der Virtuellen Kraftwerke sind (vgl. z. B. Steck [2013](#); Nikonowicz & Milewski [2012](#)). Solche Kraft-Wärme-Kopplungsanlagen sind bereits heute in verschiedenen Leistungsklassen am Markt – auch in der erweiterten Version als Kraft-Wärme-Kälte-Kopplung. Der große Vorteil liegt in der einfachen Steuerbarkeit.

Für die Simulationen im Rahmen dieser Arbeit wird eine eher kleine aber modulati-onfähige Anlage exemplarisch betrachtet werden, da diese eine gute Flexibilität für die Optimierung aufweisen. Einen Überblick über verschiedene solcher Mini-BHKW inklusive Prüfstandsmessungen gibt beispielsweise (Thomas [2007](#)).

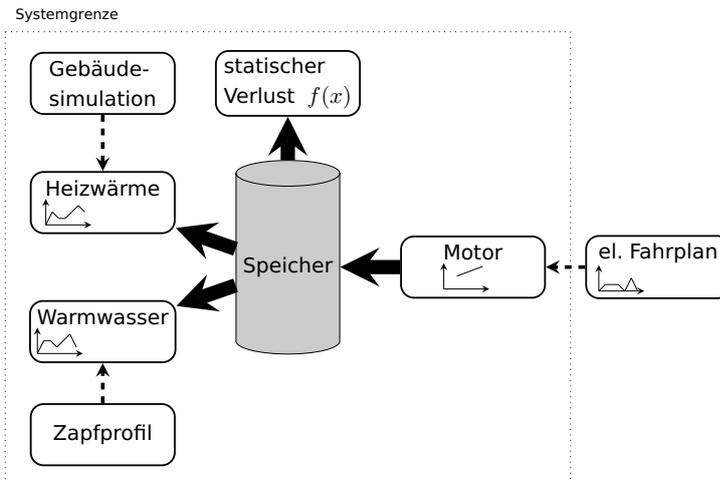


Abbildung 11.2.: Modell und Modellkomponenten für das verwendete BHKW-Modell.

Im Rahmen dieser Arbeit kommt ein Modell für Mini-BHKW zum Einsatz, welches bereits für (Bremer et al. 2010) konzipiert worden ist. Es handelt sich hierbei um ein BHKW-Modell welches von der Auslegung einem Vaillant EcoPower 4.7 (<http://www.vaillant.de/ecopower/>) nachempfunden ist. Dieses Gerät ist in der Lage, seine elektrische Erzeugung zwischen 1.3 und 4.7 kW zu modulieren. Die gleichzeitige Wärmeerzeugung schwankt dann zwischen 4 und 12.5 kW. Eine Statik für die Umrechnung von elektrischer in thermischer Leistung und umgekehrt wurde nach Teststanddaten aus (Thomas 2007) abgeleitet. Das Anlagenmodell abstrahiert von technischen Details und modelliert im Wesentlichen die Wärmeflüsse eines Gesamtsystems aus BHKW, Wärmepufferspeicher und Heizwärmebedarf eines Einfamilienhauses. Abbildung 11.2 zeigt das Modell. Berechnet wird die Wärmebilanz des Pufferspeichers aus Erzeugung durch das BHKW, Heizwärmeverluste (Berechnung der zeitvariablen Wärmeverluste anhand eines Gebäudemodelles), Warmwasserzapfung (gegeben durch statische Zapfprofile) und funktional gegebenen, statischen Verlusten durch unvollständige Isolierung. Die Modellierung der Warmwasserzapfung richtet sich nach den Zapfprofilen ZP 120, 3 DIN 4702, Teil 8 bzw. DIN EN 13203-2 (vgl. hierzu (Mühlbacher 2007)). Aus der Wärmebilanz (für jedes Zeitintervall  $t$ ) ergibt sich die jeweilige Temperaturänderung im Pufferspeicher (ohne Berücksichtigung einer Schichtung) im zeitlichen Verlauf. Äquivalent hierzu könnte analog auch die Energiebilanz bestimmt werden. Der Temperaturverlauf erlaubt jedoch die Einführung eines anschaulichen Constraints in Form eines zulässigen Temperaturbereichs innerhalb dessen sich der Pufferspeicher bewegen darf. Als weitere konfigurierbare Constraints können dem Modell eine minimale Up- bzw. Down-Time sowie ein maximaler Leistungsgradient zwischen benachbarten Zeitintervallen vorgegeben werden.

### Wärmepumpe

In letzter Zeit beginnen Wärmepumpen sich verstärkt zur Gebäudeheizung durchzusetzen, obwohl das Grundprinzip und die technische Umsetzung bereits seit längerem bekannt sind. Die Studie von Nabe et al. (2011) geht davon aus, dass bis zum Jahr 2030 etwa 4% aller Wohneinheiten mit Wärmepumpen beheizt werden.

Wärmepumpen bestehen in der Regel aus einem (flüssigkeitsbasierten) Umwälzsystem, das Wärme aus einem Reservoir geringer Wärme (beispielsweise aus der Umgebungsluft oder dem Grundwasser) aufnimmt und als Heizwärme in ein System mit höherer Temperatur (z. B. die Raumheizung) abgibt (Chua et al. 2010). Technisch gesehen gibt es verschiedene Systeme für unterschiedlichste Anwendungsfälle. Im Rahmen dieser Arbeit soll eine Wärmepumpe auf einem abstrakten Level modelliert werden. Dies geschieht analog zu (Treur 2013). Wärmepumpen haben (wenn man die Heizleistung betrachtet) einen ungewöhnlich hohen Wirkungsgrad, da die Wärme nicht erzeugt sondern lediglich transportiert wird. Ist der Wirkungsgrad der Wärmepumpe bekannt, so kann aus der zugeführten elektrischen Energie auf die an das Heizsystem abgegebene Wärmemenge geschlossen werden, sodass eine ähnliche Modellierung wie beim BHKW möglich wäre. Allerdings ist der Wirkungsgrad nicht konstant, sondern hängt von Faktoren wie Wasser- oder Umgebungstemperatur ab. Der jeweils aktuelle Wirkungsgrad wird daher auch als saisonaler Wirkungsgrad bezeichnet. In (Treur 2013) wurde ein Regressionsmodell für die Näherung des saisonalen Wirkungsgrades vorgeschlagen. Dieses wird in dieser Arbeit verwendet, um basierend auf den jeweils aktuellen Bedingungen den jeweils geltenden Wirkungsgrad zu bestimmen. Mit diesem kann sodann ein gegebener elektrischer Fahrplan in einen thermischen Lastgang umgerechnet werden.

Modelliert wird die Wärmepumpe in dieser Arbeit ohne Pufferspeicher. Stattdessen wird der Wärmebedarf  $P_{demand}$ , der von der Wärmepumpe zu decken ist, dem Modell als Zeitreihe vorgegeben. Aufgrund eines gegebenen elektrischen Fahrplans kann basierend auf den aktuellen Umwelt ( $\vartheta_{out}(t)$ ) und Heizsystembedingungen ( $\vartheta_{heater}(t)$ ) die resultierende Wärmeerzeugung im Zeitintervall  $t$  berechnet werden (Treur 2013):

$$P_{th}(t) = P_{el}(t) \cdot spf(t),$$

mit  $spf(t) = 7.45 - 0.1 \cdot (\vartheta_{heater}(t) - \vartheta_{out}(t)) + 0.004 \cdot \vartheta_{out}(t)^2$ . (11.6)

Die elektrische Leistung ist üblicherweise ein gerätetypischer fester Wert, sodass hier eigentlich nur die Werte 0 oder  $P_{max}$  vorkommen können. Eine Relaxation dieser Einschränkung wird wie bei anderen Geräten auch nach der Methode aus Abschnitt 11.3.3 auf Seite 159 vorgenommen. Als Constraint für die Zulässigkeit der Wärmeerzeugung wird folgende Bedingung eingeführt:

$$\int_{t=0}^{\tau} P_{th}(t) - P_{demand}(t) dt \leq \epsilon_{th} \quad \forall \tau \leq d. \quad (11.7)$$

Die aufintegrierte Abweichung zwischen Wärmebedarf und -erzeugung darf einen vorgegebenen Schwellwert nicht überschreiten. Als zusätzliches Constraint ist es technisch notwendig, eine Mindestlaufzeit vorzugeben. Je nach Gerätetyp finden sich hier sehr unterschiedliche Angaben. Da im Sinne dieser Arbeit ein Gerät mit möglichst großer Flexibilität modelliert werden soll, dienen Wärmepumpen der Firma Dimplex (<http://www.dimplex.de/>) als Orientierung. Hier gilt nach dem Einschalten eine Mindestlaufzeit von 6 Minuten.

Abbildung 11.3 zeigt exemplarisch einige mit diesem Modell generierte Fahrpläne. Als Vorgabe für die maximale integrierte Abweichung vom Wärmebedarf wurden 0.2 kWh vorgegeben. Die mittlere Wärmeerzeugung (über den ganzen Tag) in dem Beispiel beträgt  $80.84875 \pm 0.11508$  kWh bei einem angenommenen Wärmebedarf von 80.84873 kWh (genügt z. B. einem Niedrigenergiehaus nach der Energieeinsparverordnung (Deutsche Bundesregierung (Bundeskabinett) 2009) mit 150 m<sup>2</sup> Grundfläche). Die resultierenden

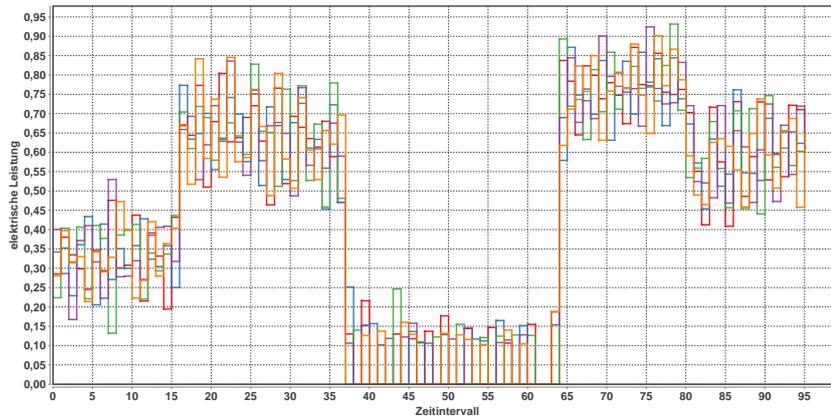


Abbildung 11.3.: Exemplarische elektrische Fahrpläne für ein Wärmepumpenmodell zur Deckung des thermischen Profils aus Abbildung 11.4.

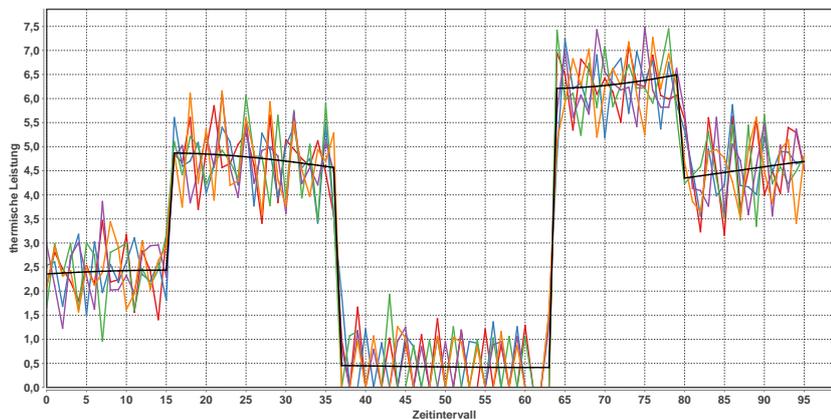


Abbildung 11.4.: Thermische Heizleistung resultierend aus den elektrischen Fahrplänen aus Abbildung 11.3. Das zu deckende thermische Profile ist in schwarz eingezeichnet.

Wärmeerzeugungsprofile mitsamt dem vorgegebenen Wärmebedarf sind in Abbildung 11.4 gegeben.

**Untertischwarmwassergerät** Eine häufig in Arbeiten zum Demand-Side Management betrachtete Gerätegruppe ist die Gruppe der Kühl- und Gefriergeräte (Stadler et al. 2007; Bigler et al. 2011). Hier sind insbesondere Anlagen in Industriegröße von Interesse. Da diese Art Gerät zumindest im Haushaltsbereich jedoch vergleichsweise wenig Potenzial bietet (Dena 2010; Seebach et al. 2009) sofern nicht eine große Zahl zu einem so genannten virtuellen Gerät zusammengefasst wird (Lünsdorf & Sonnenschein 2010; Lünsdorf 2012a), soll an dieser Stelle ein von der Modellierung her vergleichbares Gerät mit entsprechend größerem Potenzial betrachtet werden: der Untertischwarmwassergerät. Ein solches Gerät hat die Aufgabe, einen gegebenen Vorrat an Wasser innerhalb bestimmter Temperaturgrenzen zu halten. Üblicherweise funktioniert ein solches Gerät analog zu einem Kühlschrank über eine Zweipunktregelung. Anders als ein Kühlgerät, welches über

eine Wärmepumpe mit Kompressor arbeitet, kann ein Warmwasserboiler relativ hochfrequent ein- und ausgeschaltet werden, da hier lediglich ein Heizstab Verwendung findet. Hierdurch wird eine Emulation auch von Leistungszwischenstufen möglich. Ein solches Gerät lässt sich über folgende Bilanzgleichung beschreiben:

$$\frac{d}{dt}\vartheta(t) = \frac{\alpha(T_a - \vartheta(t)) + (P_{el} - P_{demand}(t))}{m \cdot C_p}, \quad (11.8)$$

mit einer mittleren spezifischen Wärmekapazität des Wassers  $C_p = 4185 \text{ J} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}$ , der Masse des Wassers  $m$ , der zeitabhängigen Wassertemperatur  $\vartheta(t)$  und der als konstant angenommenen Umgebungstemperatur  $T_a$ , der Heizleistung  $P_{el}$  und einem Faktor  $\alpha = \lambda \cdot \frac{A}{d}$ , welcher die Isolierung über Oberfläche, Dicke und Wärmeleitfähigkeit beschreibt.  $P_{demand}(t)$  gibt die zeitabhängige Warmwasserzapfung als mittlere Leistung im Zeitintervall  $t$  an.

Für die Beschreibung des aktuellen Betriebszustandes des Gerätes ist die aktuelle Temperatur hinreichend. Der Vorteil dieses Gerätes liegt darin, dass der Einfluss einer nicht beeinflussbaren Nutzung auf den Betrieb sehr leicht durch Änderung des Warmwasserbedarfs simuliert werden kann und dass derartige Änderungen einen deutlichen, auch auf kurzen Zeiträumen beobachtbaren Effekt haben. Somit eignet sich dieses Gerätemodell auch sehr gut für die Betrachtung von Prognoseunsicherheiten (vgl. Kapitel 14 ab Seite 209).

**Batteriespeicher** Die Verwendung von Batteriespeichern scheint sich als ein unverzichtbares Mittel für zukünftige Smart Grid Anwendungen herauszukristallisieren (Macary et al. 2012; Vytelingum et al. 2010; Fenn et al. 2010). Für die Modellierung von Speichern im Rahmen dieser Arbeit wurde auf ein Modell zurückgegriffen, das im Rahmen des Projektes Smart Nord entwickelt wurde (Psola et al. 2013).

Hier wurde ein allgemeines Speichermodell entwickelt, welches die Batterie im Wesentlichen über den aktuellen, minimalen und maximalen Energieinhalt, minimale und maximale Leistung, den Wirkungsgrad und die Selbstentladung charakterisiert. Bei Vorgabe einer Sollleistung kann die Zulässigkeit entschieden werden über einen Vergleich, ob die Leistung im zulässigen Rahmen liegt und ob der resultierende Energieinhalt den Vorgaben entspricht.

Alle Modelle wurden in der Programmiersprache Java umgesetzt. Im Rahmen der Evaluation kamen jedoch auch Modelle aus anderen Quellen zum Einsatz. So wurden zum Beispiel im Rahmen des Projektes Smart Nord auch Modelle verwendet, die in der Programmiersprache Python umgesetzt sind. Dies ist im Rahmen der Erstellung des Simulators im Projekt Smart Nord (<http://smarnord.de/>) geschehen. Von Christian Hinrichs wurden im Rahmen einer Integration (vgl. (Hinrichs et al. 2013b)) des Suchraummodells und des Dekoders ebenfalls in Python umgesetzte Modelle und teilweise auch Modelle aus dem Framework GridLAB-D (<http://www.gridlabd.org/>) zusammen mit den in dieser Arbeit entwickelten Artefakten eingesetzt.

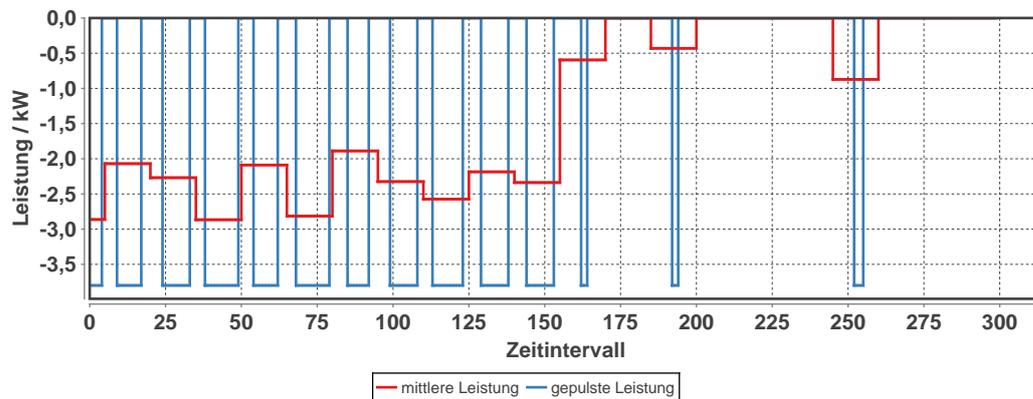


Abbildung 11.5.: Vergleich der Fahrpläne bei der Relaxation zu kontinuierlichen, gemittelten Leistungswerten am Beispiel einer Wärmepumpe.

Angemerkt sei an dieser Stelle noch, dass im Rahmen des Projektes Smart Nord derzeit Suchraummodell und Dekoder für eine Vielzahl weiterer Anlagenmodelle verwendet (Blank et al. 2013) und evaluiert werden.

### 11.3.3. Relaxation diskreter Fahrplanräume

Das Suchraummodell ist so wie es in Abschnitt Kapitel 8 ab Seite 101 präsentiert wurde zunächst geeignet für kontinuierliche Suchräume. Viele Anlagen haben jedoch eigentlich einen diskreten Definitionsbereich für ihre Fahrpläne. Ein Beispiel hierfür ist die Wärmepumpe. Diese ist entweder eingeschaltet (bei relativ konstanter Leistungsaufnahme) oder ausgeschaltet. Weitere Beispiele sind das Warmwassergerät oder die Kühlaggregate der meisten Kühlhäuser. Zwar variiert die Leistungsaufnahme eines Kühlaggregats geringfügig mit der Betriebsdauer, als ansteuerbare Leistungsstufen steht jedoch auch hier zunächst nur die Wahl zwischen an und aus oder gegebenenfalls wenigen Zwischenstufen. Aber es gibt auch weitere Anlagen mit mehr als zwei diskreten Leistungsstufen. Bei einigen modulierenden BHKW ist es beispielsweise nicht möglich, jeden Leistungswert im Modulationsbereich anzusteuern. Stattdessen können in der Praxis meist nur wenige diskrete Werte aus diesem Bereich angesteuert werden (Tröschel 2010).

Eine Erweiterung des Suchraummodells auf diskrete Suchräume soll in dieser Arbeit nicht mehr betrachtet werden. Stattdessen kommt für solche Anlagenmodelle ein Vorgehen für die Relaxation dieses Problems durch Abbilden solcher Fahrpläne auf kontinuierliche Fahrplanräume zum Einsatz.

Im Rahmen der Einsatzplanung werden immer Fahrpläne betrachtet, die über Zeitintervalle definiert sind. Für jedes Zeitintervall (z. B. 15 Minuten) wird entweder die mittlere Wirkleistung oder – äquivalent dazu – die Energiemenge betrachtet. Eine Anlage kann sofern keine Constraints dem entgegenstehen in einem Zeitintervall der Länge  $\Delta t$  für irgendeine Zeitdauer von  $t \in [0, \Delta t]$  in Betrieb sein und läuft somit mit einer mittleren Leistung zwischen 0% und 100% der Maximalleistung (über das ganze Intervall gesehen). Wird auf diese Weise ein Fahrplan mit kontinuierlichen Leistungswerten für eine solche Anlage erzeugt, so müsste dieser vor einer Ausführung von einem Anlagencontroller durch

eine geeignete Pulsweitenmodulation zurückkonvertiert werden, damit er von der Anlage umgesetzt werden kann. Abbildung 11.5 zeigt ein Beispiel für eine Wärmepumpe. Wie gut eine solche Umsetzung technisch möglich ist, hängt stark von der jeweiligen Anlage ab. Während der elektrische Heizstab eines Warmwassergeräts sehr schnell geschaltet werden kann, ist dies beispielsweise bei Kühlkompressoren nicht der Fall. Eine hohe Taktung ist meist nicht möglich, sodass hier weitere Vereinfachungen (zusammenfassen benachbarter Pulse) des Fahrplans notwendig werden würden. Daher wird im Rahmen der Evaluation in Smart Nord auch betrachtet werden, inwieweit sich eine solche Rückdiskretisierung des Fahrplans, die zwangsläufig mit Näherungen behaftet ist, auf die Qualität der erreichten Optimierungsergebnisse auswirkt.

### 11.4. Komponententests

Nachdem nun die Anlagenmodelle für die Evaluationsszenarien abgesteckt sind, soll zunächst überprüft werden, ob und wie gut diese bzw. deren Flexibilität im Einzelnen durch das Suchraummodell dargestellt werden können. Da das Suchraummodell je nach Anwendungsfall auch separat von dem Dekoder verwendet werden kann, werden hier beide Komponenten getrennt voneinander betrachtet.

Für den Klassifikator ist zu zeigen, dass die Vorhersagen (realisierbar oder nicht) über bisher ungesehene Fahrpläne in hinreichendem Maße zutreffend sind. Hierzu werden die SVDD-Klassifikatoren für verschiedene, simulierte Anlagen trainiert. Die Vorhersagen des Klassifikators über die Realisierbarkeit neuer Fahrpläne werden dann mit dem Simulatormodell verglichen. Hierbei erfolgt eine Beschränkung der verwendeten Simulationsmodelle auf die zuvor beschriebenen Anlagentypen, sodass an dieser Stelle zunächst auch nur Aussagen über diese Modelle gesichert möglich sind. Da zudem die Berechnung der Bewertungskennzahlen recht ineffektiv ist, können nicht für jede mögliche Parametrierung der jeweiligen Anlagen eigene Simulationsreihen durchgeführt werden. Hier ergäben sich aufgrund der Freiheitsgrade bei der Parametrierung zu viele Kombinationsmöglichkeiten. Stattdessen sind die jeweiligen Versuchsreihen über verschiedene zufällige Parametrierungen bei den jeweiligen Ablagen gemittelt.

Für den Dekoder muss gezeigt werden, dass er bei hinreichend kleinem Fehler zuverlässig beliebige Fahrpläne auf einen jeweils gültigen Fahrplan abbildet. Hierzu werden für die verschiedenen Anlagentypen Dekoder erzeugt und Mengen von zufällig erzeugten Fahrplänen mit dem Dekoder abgebildet. Die abgebildeten Fahrpläne werden mit Hilfe des ursprünglichen Simulationsmodells auf ihre Realisierbarkeit hin überprüft, um den jeweiligen Fehler zu bestimmen. Zudem ist zu betrachten, wie gut die vom Dekoder erzeugten (durch Abbilden zufälliger Fahrpläne) Fahrplanmengen sich räumlich mit der ursprünglichen Trainingsmenge überdecken. Hierdurch kann festgestellt werden, ob alle Bereiche des Suchraums vom Dekoder bedient werden, oder ob bestimmte Bereiche ausgespart bleiben. Die Einschränkungen bezüglich der Anlagenmodelle sind die gleichen wie beim Klassifikator. Als zusätzliche Einschränkung kommt hinzu, dass der Vergleich der Suchräume nur über die Näherung durch endliche (und vergleichsweise kleine) Mengen von Fahrplänen möglich ist.

### 11.4.1. Klassifikator

Für die Bewertung von Klassifikatoren im Allgemeinen sind bereits eine ganze Reihe von Kennzahlen etabliert worden. Ein Übersicht über solche Kennzahlen und Bewertungsmöglichkeiten findet sich beispielsweise in (Powers 2008; Witten et al. 2011).

Die Bestimmung dieser Kennzahlen für Klassifikatoren beruht üblicherweise auf einem wiederholten Vergleich von Vorhersagen des Klassifikators mit der Realität. In dem Anwendungsfall hier kann das Klassifikationsergebnis eines gegebenen Fahrplans mithilfe des Anlagensimulationsmodells verglichen werden. Hierdurch können zunächst die Häufigkeiten für die vier möglichen Fälle bestimmt werden:

**True Positive ( $n_{tp}$ )** Einstufung vom Klassifikator: realisierbar, Einstufung vom Simulationsmodell: realisierbar.

**True Negative ( $n_{tn}$ )** Einstufung vom Klassifikator: nicht realisierbar, Einstufung vom Simulationsmodell: nicht realisierbar.

**False Positive ( $n_{fp}$ )** Einstufung vom Klassifikator: realisierbar, Einstufung vom Simulationsmodell: nicht realisierbar.

**False Negative ( $n_{fn}$ )** Einstufung vom Klassifikator: nicht realisierbar, Einstufung vom Simulationsmodell: realisierbar.

Aus diesen Häufigkeiten können im Nachgang diverse bedingte Wahrscheinlichkeiten als Kennzahlen zur Beurteilung des Klassifikators bestimmt werden. Folgende Kennzahlen sind üblich:

**Klassifikationsfehler:** Der Klassifikationsfehler  $\frac{n_{fp}+n_{fn}}{m}$  für  $m$  klassifizierte Fahrpläne definiert sich über den Anteil der falsch klassifizierten Fahrpläne. Analog lässt sich die komplementäre Wahrscheinlichkeit für die Vertrauenswahrscheinlichkeit definieren.

**Trefferquote (engl. recall):** Die Trefferquote oder Sensitivität  $\frac{n_{tp}}{n_{tp}+n_{fn}}$  gibt den Anteil der richtig als ausführbar klassifizierten Fahrpläne an und kennzeichnet somit, wie viel Prozent des Lösungsraums tatsächlich bei der Optimierung einbezogen werden können.

**Falsch-Negativ-Rate:** Die Falsch-Negativ-Rate (engl. miss)  $\frac{n_{fn}}{n_{tp}+n_{fn}}$  gibt den Anteil fälschlicherweise als nicht ausführbar klassifizierten Fahrpläne an. Sie liefert als komplementäre Wahrscheinlichkeit zur Trefferquote keine zusätzliche Information zur Beurteilung.

**Ausfallrate (engl. fallout):** Die Ausfallrate oder Falsch-Positiv-Rate  $\frac{n_{fp}}{n_{tn}+n_{fp}}$  gibt den Anteil der fälschlicherweise als ausführbar klassifizierten nicht ausführbaren Fahrpläne an. Sie gibt die Wahrscheinlichkeit an, dass ungültige Fahrpläne irrtümlich bei der Optimierung berücksichtigt werden

**Richtig-Negativ-Rate:** Die Richtig-Negativ-Rate oder Spezifität  $\frac{n_{tn}}{n_{tn}+n_{fp}}$  misst den Anteil richtigerweise zurückgewiesener Fahrpläne und liefert als komplementäre Wahrscheinlichkeit zur Ausfallrate keine zusätzliche Beurteilung.

**Genauigkeit (engl. precision):** Die Relevanz oder Genauigkeit  $\frac{n_{tp}}{n_{tn}+n_{fp}}$  gibt an, welcher Anteil der als ausführbar klassifizierten Fahrpläne auch tatsächlich ausführbar

## 11. Evaluation

Tabelle 11.2.: Vergleich verschiedener Klassifikatorkennzahlen für ein Untertischwarmwassergerät für 2 bis 8-dimensionale Fahrpläne ( $d = 2 \dots 8$ ). Die Parametrierung für das Training des Klassifikators wurde jeweils auf einen maximalen  $F_1$ -Wert eingestellt.

Kenngröße	$d = 2$	$d = 4$	$d = 8$
Trefferquote	$0.9852 \pm 0.0051$	$0.8851 \pm 0.0149$	$0.8043 \pm 0.0916$
Genauigkeit	$0.9850 \pm 0.0053$	$0.9203 \pm 0.0117$	$0.7746 \pm 0.0916$
$F_1$	$0.9851 \pm 0.0020$	$0.9022 \pm 0.0067$	$0.7868 \pm 0.0695$
Ausfallrate	$3.4791 \times 10^{-4} \pm 1.2491 \times 10^{-4}$	$2.7850 \times 10^{-5} \pm 4.8204 \times 10^{-6}$	$1.2300 \times 10^{-6} \pm 4.3176 \times 10^{-7}$
Segreganz	$0.9996 \pm 1.1793 \times 10^{-4}$	$0.9999 \pm 5.1669 \times 10^{-6}$	$0.9999 \pm 4.5549 \times 10^{-7}$
Fehler	$6.7456 \times 10^{-4} \pm 9.0577 \times 10^{-5}$	$6.9370 \times 10^{-5} \pm 4.2293 \times 10^{-6}$	$2.5999 \times 10^{-6} \pm 6.9615 \times 10^{-7}$

Tabelle 11.3.: Vergleich verschiedener Klassifikatorkennzahlen für eine exemplarische Wärmepumpe für 2 bis 8-dimensionale Fahrpläne. Die Parametrierung für das Training des Klassifikators wurde jeweils auf einen maximalen  $F_1$ -Wert eingestellt.

Kenngröße	$d = 2$	$d = 4$	$d = 8$
Trefferquote	$0.8545 \pm 0.0263$	$0.7588 \pm 0.0209$	$0.6855 \pm 0.0508$
Genauigkeit	$0.9989 \pm 0.0011$	$0.9256 \pm 0.0232$	$0.8139 \pm 0.0601$
$F_1$	$0.9209 \pm 0.0154$	$0.8336 \pm 0.0140$	$0.7406 \pm 0.0165$
Ausfallrate	$3.1350 \times 10^{-4} \pm 3.5515 \times 10^{-4}$	$2.5422 \times 10^{-3} \pm 8.8630 \times 10^{-4}$	$8.8704 \times 10^{-4} \pm 2.6686 \times 10^{-4}$
Segreganz	$0.9499 \pm 8.5811 \times 10^{-3}$	$0.9901 \pm 8.4952 \times 10^{-4}$	$0.9983 \pm 2.6686 \times 10^{-4}$
Fehler	$0.0389 \pm 0.0070$	$0.0120 \pm 9.8064 \times 10^{-4}$	$0.0026 \pm 2.1118 \times 10^{-4}$

sind. Da alle als ausführbar klassifizierten Fahrpläne theoretisch bei einer Optimierung berücksichtigt werden können, gibt die Relevanz eine Schätzung dafür, dass das Optimierungsergebnis eine gültige Lösung darstellt. Die Relevanz entspricht allerdings nicht direkt der Wahrscheinlichkeit, dass die Lösung gültig ist, da das Optimierungsproblem über mehrere Suchräume definiert ist und zudem nicht jeder Fahrplan die gleiche Wahrscheinlichkeit hat Teil der Lösung zu werden.

**Segreganz (engl. Negativ predictive value):** Die Segreganz – auch Trennfähigkeit genannt –  $\frac{n_{tn}}{n_{tn} + n_{fn}}$  gibt an, welcher Anteil der während der Optimierung als ungültig abgelehnten Fahrpläne auch tatsächlich ungültig sind. Interessant für den hier vorliegenden Anwendungsfall wäre eher die komplementäre Wahrscheinlichkeit zur Segreganz, welche angibt, wie groß der Anteil der zu Unrecht abgelehnten Fahrpläne ist, da diese fälschlicherweise von der Lösungsfindung ausgeschlossen werden.

Anzumerken ist, dass die Werte von Genauigkeit und Segreganz nur dann statistisch aussagekräftig sind, wenn die Häufigkeit der ausführbaren Fahrpläne in der Stichprobe zur Berechnung der Bewertungskriterien mit der Häufigkeit in der Gesamtheit aller Fahrpläne übereinstimmt. Dies ist nur gegeben, wenn für die Berechnung eine unabhängige Stichprobe zugrunde liegt, die aber die in Abschnitt 8.2 auf Seite 102 bereits beschriebenen Probleme der Berechenbarkeit mit sich bringt. Daher können die hier vorgestellten Kennzahlen auch nur für verhältnismäßig kurze Fahrpläne von bis zu ca. 8 Zeitintervallen in akzeptabler Zeit bestimmt werden.

Da ein Klassifikator hinsichtlich Trefferquote und Genauigkeit aufgrund wechselseitiger Abhängigkeit nicht gleichzeitig unabhängig voneinander optimiert werden kann, wird häufig auch das harmonische Mittel beider Werte verwendet. Bei gleicher Gewichtung beider Kennzahlen spricht man vom  $F_1$ -Maß.

In einer ersten Testreihe wurde ein mittels eines Untertischwarmwassergeräts trainier-

Tabelle 11.4.: Vergleich verschiedener Klassifikatorkennzahlen für einen fast vollständig geladenen Batteriespeicher (Redox-Flow) für 2 bis 8-dimensionale Fahrpläne. Die Parametrierung für das Training des Klassifikators wurde jeweils auf einen maximalen  $F_1$ -Wert eingestellt.

Kenngröße	$d = 2$	$d = 4$	$d = 8$
Trefferquote	0.9400 ± 0.0226	0.8825 ± 0.0326	0.8315 ± 0.0354
Genauigkeit	0.9509 ± 0.0106	0.6093 ± 0.0196	0.4326 ± 0.0139
$F_1$	0.9452 ± 0.0111	0.7102 ± 0.0058	0.5685 ± 0.0078
Ausfallrate	0.0283 ± 0.0067	0.1702 ± 0.0177	0.2094 ± 0.0190
Segreganz	0.9656 ± 0.0124	0.9596 ± 0.0092	0.9610 ± 0.0109
Fehler	0.0399 ± 0.0077	0.1580 ± 0.0069	0.2028 ± 0.0109

ter Klassifikator zur Identifikation realisierbarer Fahrpläne getestet. Tabelle 11.2 zeigt das Ergebnis für 2, 4 und 8-dimensionale Fahrpläne. Verglichen wurden die Realisierbarkeitseinstufungen von gleichverteilt zufallsgenerierten Fahrplänen von Klassifikator und dem mathematischen Modell selbst. Während Trefferquote und Genauigkeit bei den 2-dimensionalen Fahrplänen mit jeweils 98 % noch recht hoch sind, nehmen diese Werte mit zunehmender Dimensionalität ab; was durchaus auch zu erwarten ist. Trennfähigkeit und Fehler bleiben jedoch auf einem guten Niveau. Dies ist dadurch zu erklären, dass gleichbleibend jeweils 500 Fahrpläne in der Trainingsmenge verwendet wurden. Mit zunehmender Dimensionalität kann dadurch der Suchraum offenbar nicht mehr in seiner vollen Strukturausprägung erfasst werden. Auch dieser Effekt ist nicht verwunderlich: je höher-dimensional der Suchraum ist, desto umfangreicher muss die Trainingsmenge sein, um einen Klassifikator gleicher Güte zu erhalten. Die Frage ist hier eher nach einem geeigneten Kompromiss zwischen einer vom Rechenaufwand her im praktischen Einsatz vertretbaren Menge an Trainingsinstanzen und dem möglichen Fehler, den eine zu kleine Trainingsmenge mitbringt. Offensichtlich ist aber der Klassifikationsfehler nicht so stark von der Stichprobengröße beeinflusst wie der erfasste Anteil des tatsächlichen Suchraums. Die Tabellen 11.3 und 11.4 zeigen die Ergebnisse für Wärmepumpen und Redox-Flow-Batterien. Auch hier ließen sich durch Erhöhung des Trainingsmengenumfangs auch für höherdimensionale Szenarien deutliche Verbesserungen erreichen (vgl. auch Abbildung 11.6).

Für längere Fahrpläne ist die Berechnung aller Kennzahlen nicht mehr möglich. Hier wurde ein anderes Vorgehen gewählt: Es wurde eine Menge von Testfahrplänen erzeugt, die sich je zur Hälfte aus zufällig geratenen Fahrplänen (mit einer Wahrscheinlichkeit einen ungültigen zu raten von nahezu 1) und aus einem weiteren Sample (nach der sukzessiven Methode) zusammensetzt. Diese Testmenge (jeweils 100.000 Fahrpläne) wurde anhand des Klassifikators und mit dem Simulationsmodell auf Realisierbarkeit geprüft, um die Rate korrekt klassifizierter Fahrpläne zu bestimmen. Gemittelt wurde über 100 Trainingsmengen desselben Startzustand. Tabelle 11.5 zeigt das Ergebnis hierzu für Fahrpläne mit 32 bis 96 Viertelstundenintervallen. Die Parametrierung wurde jeweils für 96-dimensionale Fahrpläne vorgenommen. So zeigt das Ergebnis auch, wie gut (bzw. schlecht) Parameter für ganztägige Suchraummodelle auch für kürzere Zeiträume wiederverwendet werden können, falls im Rahmen einer Umplanung das Suchraummodell für verbleibende Restzeiten neu erstellt werden muss. Die Sensitivität der Modelle diesbezüglich ist stark unterschiedlich. Dies ist unmittelbar einsichtig, da der Informationszuwachs über den Tag bei einigen Modell nicht konstant ist. Während sich für einen ungestörten Wasserboiler kaum Unterschiede ergeben, sind die Unterschiede bei BHKW entsprechend hoch, wenn

## 11. Evaluation

Tabelle 11.5.: Korrektklassifikationsrate für höher dimensionale Fahrpläne verschiedener Anlagen. Für die Berechnung wurden Mengen von je 50% gültiger (gesampelt) und ungültiger (zufällig erzeugter) Fahrpläne getestet (je 50.000). Die Parametrierung ist jeweils für 96 Dimensionen optimiert.

Einheit	$d = 32$	$d = 48$	$d = 64$	$d = 80$	$d = 96$
Boiler	$0.9819 \pm 0.0055$	$0.9778 \pm 0.0051$	$0.9749 \pm 0.0062$	$0.9722 \pm 0.0062$	$0.9694 \pm 0.0064$
BHKW	$0.5965 \pm 0.0127$	$0.6249 \pm 0.0149$	$0.8256 \pm 0.0370$	$0.9335 \pm 0.0450$	$0.9938 \pm 0.0489$
Wärmepumpe	$0.8682 \pm 0.0208$	$0.9876 \pm 0.0103$	$0.9855 \pm 0.0116$	$0.9881 \pm 0.0083$	$0.9851 \pm 0.0103$
inv. Batterie	$0.9532 \pm 0.0409$	$0.9706 \pm 0.0267$	$0.9734 \pm 0.0220$	$0.9770 \pm 0.0187$	$0.9729 \pm 0.0221$

zu bestimmten Zeiten ein erhöhten Warmwasserbedarf eine große Änderung in den möglichen Fahrplänen induziert – je nach dem, ob diese vom Zeithorizont her solche Zeitpunkte enthalten oder nicht. Kürzere Fahrpläne enthalten die späteren Spitzen nicht. Ähnliches gilt für die Wärmepumpen mit einem tagsüber stark reduzierten Wärmebedarf.

Ein besonderes Problem stellte sich bei den Batteriespeichern heraus: Es schien zunächst unmöglich Korrektklassifikationsraten von mehr als ca. 70% zu erreichen. Dies ist in dem vergleichsweise riesigen Suchraum einer Batterie, der zudem kaum Struktur aufweist, begründet. Ein hinreichend groß dimensionierter Batteriespeicher kann fast alle Fahrpläne realisieren. Hier hat es sich als vorteilhaft erwiesen, das Negativ des Suchraums (Menge der nicht realisierbaren Fahrpläne) lernen zu lassen. Im Falle der Batterie ist dieser offenbar bedeutend kleiner und strukturierter, sodass sich so bessere Resultate erzielen lassen. Bei der Benutzung müssen lediglich alle Klassifikationsanfragen bezüglich der Realisierbarkeit negiert werden. Bei allen Versuchen in Tabelle 11.5 wurde der Umfang der Trainingsmenge mit 800 Fahrplänen konstant gelassen.

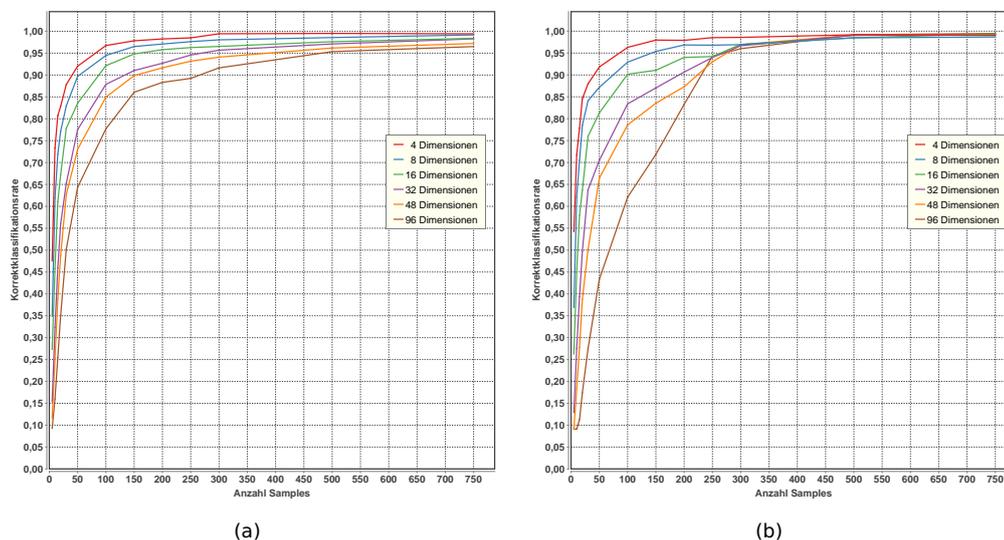


Abbildung 11.6.: Zusammenhang zwischen Samplegröße und Korrektklassifikationsrate am Beispiel Warmwassergerät 11.6(a) und Wärmepumpe 11.6(b).

Für die Darstellung des Zusammenhangs zwischen der Größe der Trainingsmenge und der erreichbaren Klassifikatorgüte wurden verschiedene Testreihen mit Warmwassergeräten und Wärmepumpen durchgeführt. Abbildung 11.6 zeigt die Ergebnisse für

verschiedene Fahrplanlängen (gleichbedeutend mit Produktlieferzeiträumen). Wie zu erwarten liefert eine größere Trainingsmenge ein besseres Klassifikationsergebnis. Dies ist nicht weiter verwunderlich. Es ist bekannt, dass bei Vorliegen nur einer Klasse in der Regel eine größere Trainingsmenge benötigt wird, um eine gleich gute Trennung bei der Erkennung zu erreichen wie bei 2-Klassenverfahren (vergleiche z. B. Khan & Madden 2010). Als Kennzahl wurde die Korrektklassifikationsrate benutzt, welche sich bestimmt als Prozentsatz korrekt klassifizierter Fahrpläne. Hierfür wurden aus Komplexitätsgründen wieder 50000 gültige Fahrpläne mit Hilfe des Simulationsmodells sowie ungültige, zufällige Fahrpläne für den Vergleich von Klassifikator und Anlagensimulation bestimmt. Das Verhältnis aus gültigen und ungültigen Fahrplänen stimmt zwar mit der Realität insbesondere bei längeren Fahrplänen nicht überein, dennoch erlaubt diese Korrektklassifikationsrate eine Abschätzung darüber, wie groß der Anteil korrekt erkannter gültiger Fahrpläne ist, da die positive Klasse bei diesem Vorgehen überproportional stark vertreten ist.

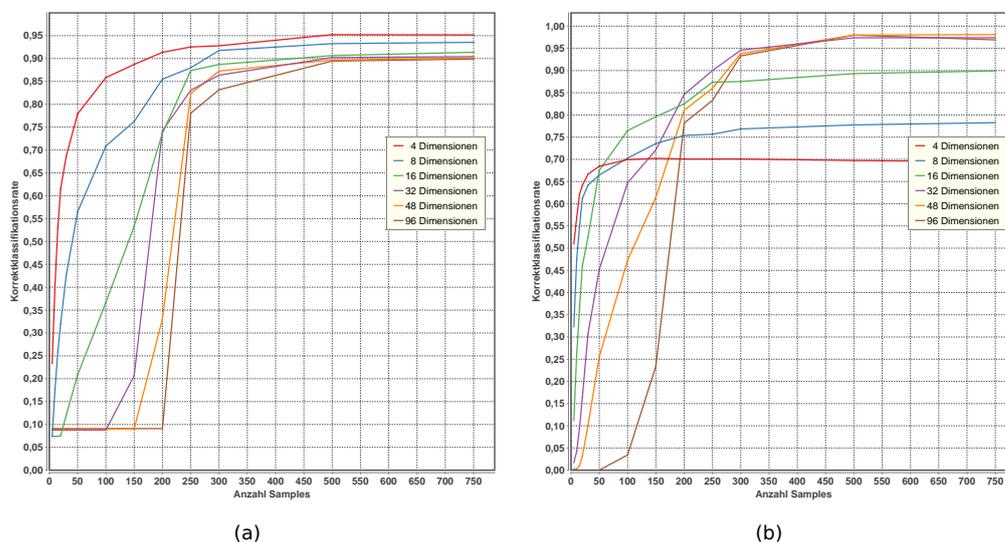
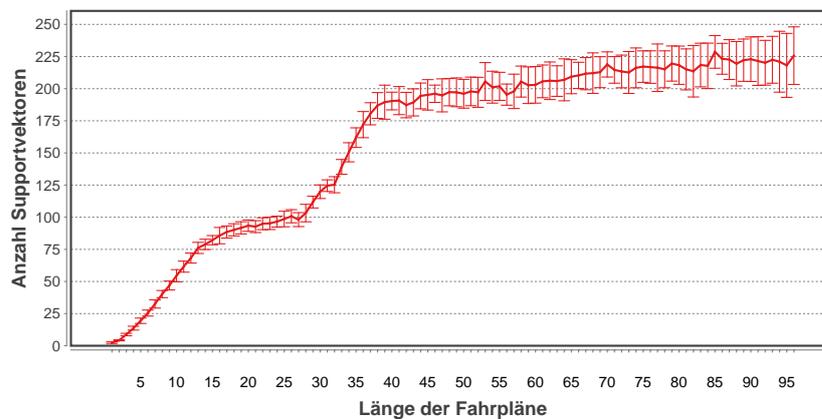


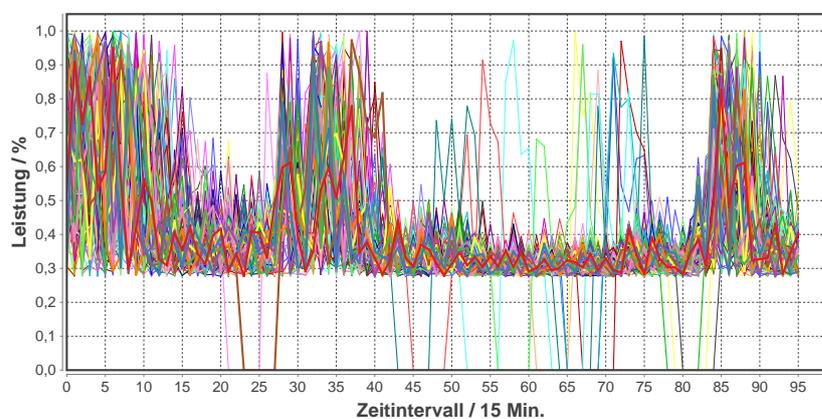
Abbildung 11.7.: Zusammenhang zwischen Samplegröße und Korrektklassifikationsrate am Beispiel einer Redox-Flow-Batterie: normales Batteriemodell links (11.7(a)) und als invertiertes Modell rechts (11.7(b)).

Bei Anlagen mit hinreichend kleinen Suchräumen (wie die beiden Anlagen in Abbildung 11.6) scheinen bereits relative kleine Trainingsmengen mit unter 1000 Samples ein brauchbares Klassifikationsergebnis zu liefern. Abbildung 11.7 zeigt jedoch, dass es auch Anlagen (in dem Beispiel Redox-Flow-Batterien) gibt, welche einen so hohen Anteil gültiger Fahrpläne haben, dass sehr große Trainingsmengen benötigt werden. Abhilfe schafft hier aber der bereits geschilderte Trick mit der Modellierung der Komplementärmenge des Suchraums. Abbildung 11.7(b) zeigt eine Redox-Flow-Batterie bei der das Negativ des Suchraums (also die Menge der nicht realisierbaren Fahrpläne) gelernt wurde. Zumindest für höherdimensionale Fahrpläne kann ein solches Vorgehen bei der Erstellung des Klassifikators eine Verbesserung erreichen. Dies ist jedoch immer im Einzelfall zu prüfen. Wie Abbildung 11.7(b) für unter 16-dimensionale Fahrpläne zeigt, kann dieses Vorgehen auch zu einer Verschlechterung führen.

## 11. Evaluation



(a)



(b)

Abbildung 11.8.: Anzahl der für eine gute Klassifikation benötigten Supportvektoren [11.8\(a\)](#) am Beispiel BHKW im Vergleich mit der Charakteristik der Fahrpläne aus der Trainingsmenge [11.8\(b\)](#). Die Fehlerbalken in [11.8\(a\)](#) zeigen jeweils die Standardabweichung der Versuchsreihe.

Eine wichtige Kenngröße ist die Anzahl der Supportvektoren, die für ein hinreichend gutes Modell mitsamt Dekoder benötigt werden. Von dieser Anzahl hängt auch die Performanz des Dekoders ab. Da die Anzahl der Supportvektoren für ein Modell sowohl von der Länge der Fahrpläne als insbesondere auch von der Charakteristik (der Form) des Suchraums einer Anlage abhängen, kann diese Kenngröße nur anhand von Beispielen betrachtet werden. [Abbildung 11.8](#) zeigt ein Beispiel für ein BHKW. Dargestellt ist die Entwicklung der Anzahl der Supportvektoren in Abhängigkeit von der Länge der Fahrpläne [11.8\(a\)](#). Für die gesamte Versuchsreihe wurden immer Fahrpläne aus demselben Startzustand und vom selben Zeitpunkt ab an verwendet. Somit reflektiert jeder Punkt der Reihe [11.8\(a\)](#) die Anzahl der Supportvektoren für den vorderen Abschnitt des Suchraums bzw. für den bis zum entsprechenden Zeitintervall reichenden vorderen Teil der Trainingsvektoren aus [11.8\(b\)](#).

Man sieht, dass die Anzahl stärker wächst wenn sich mit wachsender Fahrplanlänge der Informationsgehalt stärker ändert. Dies ist zum Beispiel im Bereich der Zeitintervalle

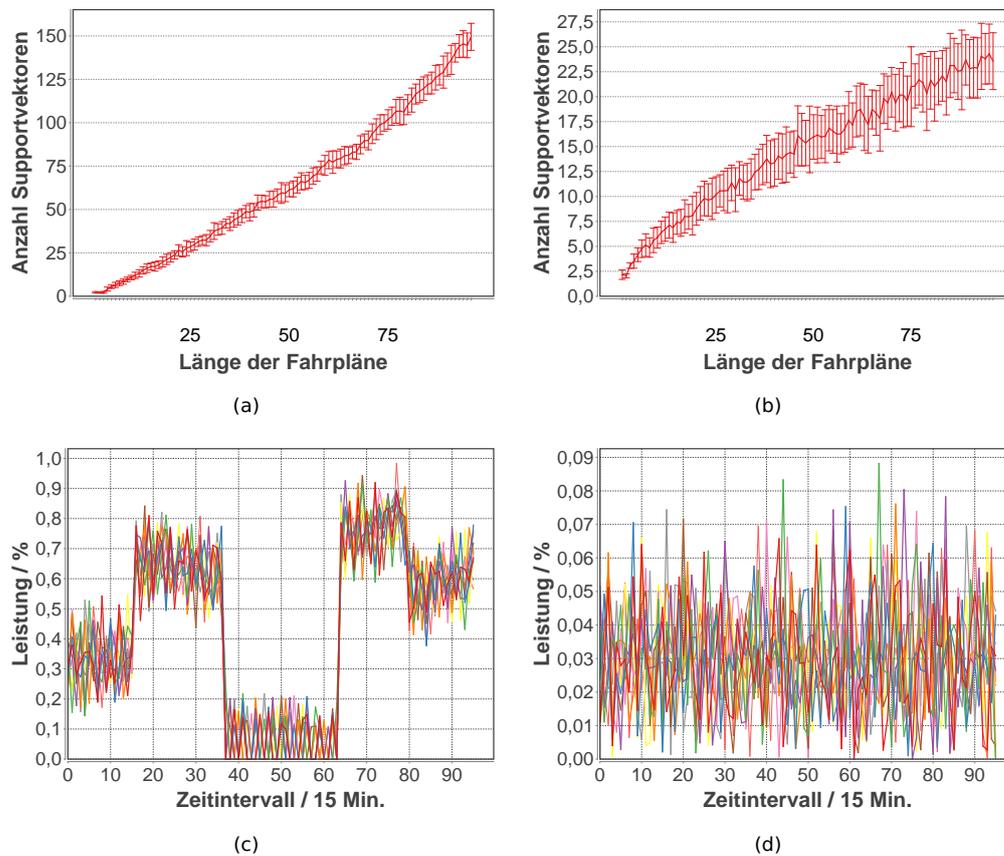


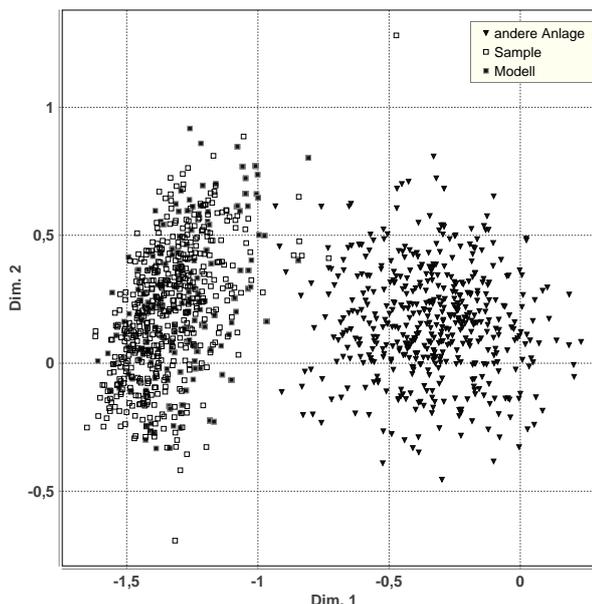
Abbildung 11.9.: Anzahl der für eine gute Klassifikation benötigten Supportvektoren am Beispiel Wärmepumpe (links) und Boiler (rechts).

etwa zwischen 30 und 40 der Fall, wo im Modell das morgendliche Duschen als thermische Belastung implementiert ist. Dieser Effekt schwächt sich allerdings mit zunehmender Fahrplanlänge ab; die Anzahl der Supportvektoren scheint zumindest für dieses Modell eine obere Schranke zu haben. Dies ist nicht immer für Fahrpläne mit einer Länge von nur einem Tag der Fall, wie das Beispiel der speicherlosen Wärmepumpe [11.9\(a\)](#) zeigt. Anscheinend fügen auch die Sprünge der Leistungswerte in den Fahrplänen durch die Änderung der Temperaturvorgabe in [11.9\(c\)](#) keinen nennenswerten Informationsgehalt zu den sich ansonsten eher wie beim Boiler in [11.9](#) verhaltenden Fahrplänen hinzu.

Man sieht, dass bereits mit handhabbar kleinen Trainingsmengen (mit weniger als 1000 Fahrplaninstanzen) akzeptable Resultate zu erzielen sind. Das bedeutet aber auch, dass im praktischen Einsatz die Berechnungskomplexität für die Erstellung des Klassifikatormodells klein genug ist, um auch auf leistungsschwächeren Rechnern ein Ergebnis in deutlich weniger als einer Minute zu berechnen. Im Falle einer vorausschauenden Wirkleistungsplanung in Day-Ahead-Szenarien ist dies sicherlich ausreichend. Im Falle von Umplanungen während der Produktauslieferung kann gegebenenfalls auch auf kleinere Trainingsmengen ausgewichen werden, da über einen kürzeren Zeithorizont geplant werden muss, wenn die Produktlieferung bereits läuft.

## 11. Evaluation

Abbildung 11.10: Visueller Vergleich von Simulationsmodell und Suchraummodell für hochdimensionale Fahrpläne eines BHKW. Dargestellt sind drei Mengen von zufällig generierten 96-dimensionalen Fahrplänen, die mithilfe von RDP in eine 2-dimensionale Darstellung eingebettet wurden. Generiert wurden die Mengen mit dem Simulationsmodell, dem trainierten Suchraummodell-Dekoder durch Abbilden gleichverteilt erzeugter, zufälliger Fahrpläne und mit dem Simulationsmodell eines weiteren BHKWs mit einem leicht abweichenden Laufzeitverhalten. Dim. 1 und 2 kennzeichnen die beiden intrinsischen Dimensionen auf die 96 Dimensionen der Fahrpläne durch die Dimensionsreduktion abgebildet wurden.



Die Anzahl der Supportvektoren, die für ein Modell benötigt wird, ist im Wesentlichen interessant für Fälle in denen das Modell kommuniziert werden soll – beispielsweise an eine zentrale Steuerungsinstanz, falls es sich um ein statisches Virtuelles Kraftwerk handelt. Bei 250 Supportvektoren in einem Modell und 4 Byte je Fließkommawert beläuft sich das zu kommunizierende Datenvolumen (inklusive Gewichtungsvektor und Anlagenparameter) auf deutlich weniger als 100 kByte je Modell. Darüber hinaus ist die Anzahl der Supportvektoren aber auch von Interesse, da sie maßgeblich den Berechnungsaufwand für den Dekoder determiniert, der im nächsten Abschnitt behandelt wird.

### 11.4.2. Dekoder

Für den visuellen Vergleich hochdimensionaler Fahrplanmengen ist es aus Gründen der Darstellbarkeit wieder notwendig, eine Dimensionsreduktion vorzunehmen. Da sich die 2-dimensionale Einbettung der Entscheidungsgrenze als schwierig erweist, werden hierfür lediglich Punktwolken miteinander verglichen. Hierzu wird der Dekoder verwendet. Abbildung 11.10 zeigt ein erstes Beispiel mit hochdimensionalen Fahrplänen eines BHKW. Hierfür wurden drei Mengen mit jeweils 96-dimensionalen Fahrplänen generiert. Die erste Menge wurde mit Hilfe des Simulationsmodells generiert, mit dem auch die ursprüngliche Trainingsmenge erzeugt wurde. Die zweite Menge wurde dann mit Hilfe des Dekoders erzeugt, der für dieses Simulationsmodell trainiert wurde. Als Vergleichsmenge schließlich wurde noch eine dritte Menge von Fahrplänen durch ein zweites Simulationsmodell für leicht anders spezifiziertes BHKW erzeugt, wobei der einzige Unterschied die Vorgabe einer geringeren Mindestlaufzeit ist. Man sieht leicht wie sich die Lage des Clusters der separaten Anlage von den relativ deckungsgleichen von Simulationsmodell und Dekoder derselben Anlage unterscheidet.

Ein zweites Beispiel zeigt Abbildung 11.11. In diesem Beispiel mit Wärmepumpenfahr-

plänen lässt sich erkennen, dass auch der Suchraum einer einzelnen Anlage in separate Bereiche zerfallen kann. Dies ist an der vergleichenden Zeitreihendarstellung nicht unmittelbar ersichtlich. Die geringe Variabilität der Fahrpläne beruht auf der Tatsache, dass ein Modell ohne Pufferspeicher gewählt wurde, um einen Fahrplanraum zu erzeugen, der sich in seiner Struktur deutlich von den BHKW-Modellen mit Pufferspeicher unterscheidet.

Abbildung 11.11 lässt aber auch erkennen, dass die Parametrierung des Dekoders noch nicht optimal gewählt ist. Offensichtlich deckt die vom Dekoder erzeugte Punktwolke mit Fahrplänen zwar denselben Bereich ab wie auch das Simulationsmodell, jedoch erzeugt der Dekoder eine eher dünn besiedelte Punktwolke, wobei viele der ungültigen Fahrpläne anscheinend auf sehr ähnliche, benachbarte gültige Fahrpläne abgebildet werden. Einen älteren Vergleich hochdimensionaler Fahrpläne aus (Bremer et al. 2010) zeigt Abbildung 11.12. Hier wurde die Parametrierung offenbar günstiger gewählt. Für den Vergleich wurden hier zwei verschiedene Dimensionsreduktionsmethoden verwendet. Zum einen die so genannte Sammon-Projektion (Sammon 1969) und zum anderen eine Hauptkomponentenanalyse (engl.: Principal Component Analysis, PCA) verwendet. Beides sind etablierte Verfahren zur Dimensionsreduktion und wurden in dieser Arbeit lediglich benutzt.

Es werden also Kennzahlen zur Bestimmung geeigneter Parameterwerte zur Erzeugung guter Dekoder benötigt.

Betrachtet man zwei nichtleere, kompakte Teilmengen eines metrischen Raums, so definiert der Hausdorff-Abstand (Hausdorff 1914) ein Maß dafür, wie weit diese beiden Teilmengen auseinander liegen. In der Geometrie wird dieser Abstand als Maß dafür benutzt, wie deckungsgleich zwei Geometrien sind. Je deckungsgleicher die beiden Mengen sind (jeder Punkt der einen Menge liegt nah bei irgendeinem Punkt der anderen Menge) desto geringer ihr Hausdorffabstand. Der Hausdorffabstand ist definiert als der größte minimale Abstand, den ein Punkt der einen Menge zu einem Punkt der anderen hat:

$$d_H(X, Y) = \max\left\{\sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y)\right\}. \quad (11.9)$$

Der Hausdorff-Abstand ist ursprünglich definiert für kompakte Geometrien, es gibt jedoch auch eine diskrete Variante zum Messen des Überdeckungsgrad von Punktwolken, wie sie beispielsweise für den Vergleich von Bildern oder die Erkennung von Formen oder Umrissen verwendet wird (siehe z. B. Mémoli & Sapiro 2004; Huttenlocher et al. 1993, sowie die Referenzen darin).

Werden zwei Geometrien durch zwei in ihrem jeweiligen Inneren zufällig generierte Punktwolken näherungsweise beschrieben und werden sodann statt der Geometrien selbst die beiden Punktwolken geometrisch mittels des Hausdorffabstands miteinander verglichen, so ist selbst bei völliger Übereinstimmung der beiden Geometrien ein Hausdorffabstand von 0 nicht zu erwarten. Offensichtlich hängt die Größe des Fehlers von der Anzahl der Punkte in der jeweiligen Punktwolke ab. Anschaulich werden die beiden konvexen Hüllen der Punktwolken miteinander verglichen die umso größer sind je weniger Punkte für die Näherung der ursprünglichen Form zur Verfügung stehen. Die Anzahl der Punkte soweit zu erhöhen, dass der verbleibende Fehler hinreichend klein wird ist keine Option, da hier die Komplexität bei der Berechnung der Hausdorffdistanz eine Schranke setzt. Es wird daher wie folgt vorgegangen, um eine Bewertungskennzahl aus dem diskreten Hausdorffabstand abzuleiten: Es wird zunächst der Erwartungswert für den Fehler aufgrund der Größe der Punktmenge bestimmt als mittlerer Hausdorffabstand von 100

## 11. Evaluation

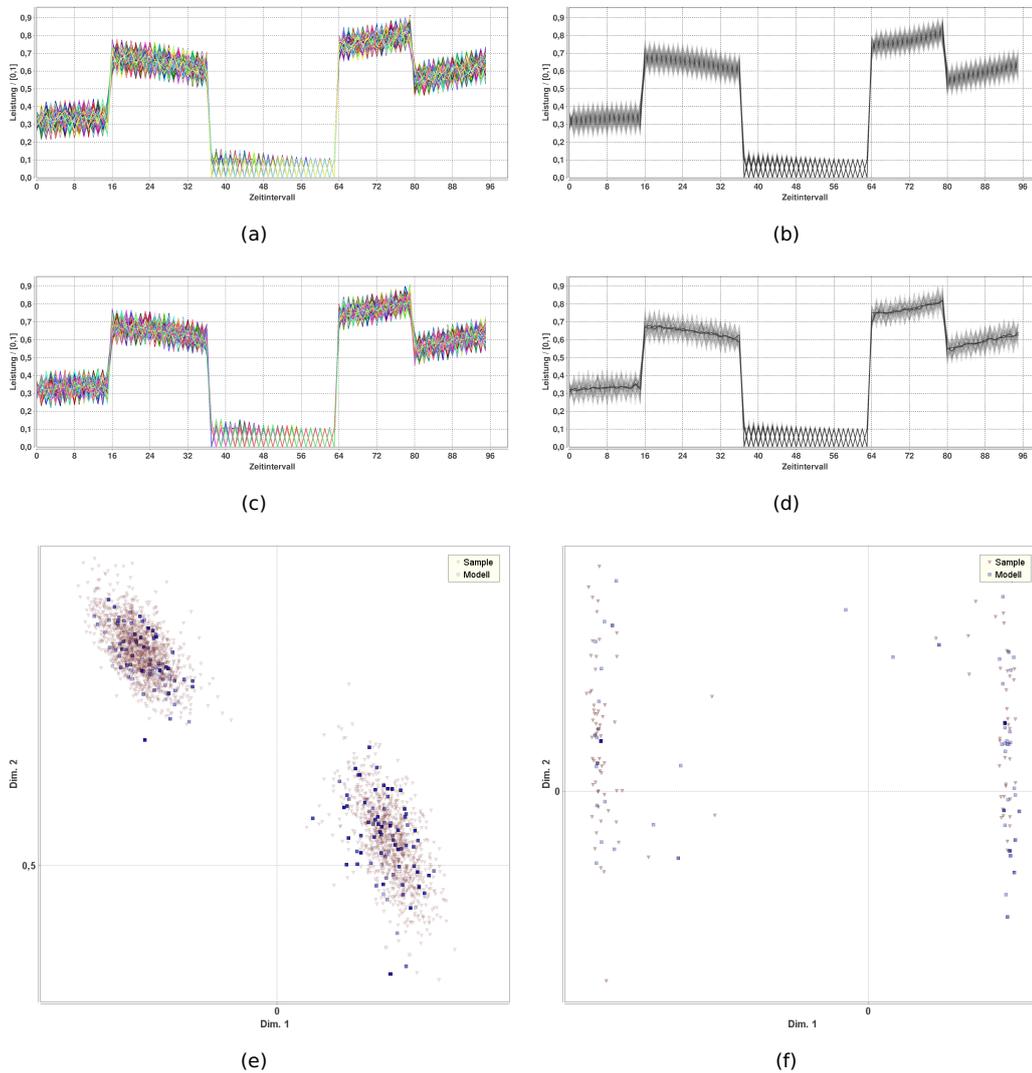


Abbildung 11.11.: Vergleich der Modelle für eine Wärmepumpe. Abbildung 11.11(a) zeigt eine durch Sampling des Simulationsmodells generierte zufällige Fahrpläne für ein Wärmepumpenmodell ohne Pufferspeicher mit eher engen Temperatureinhaltevorgaben. Abbildung 11.11(b) zeigt die gleiche Menge in halbtrensparender Darstellung (dunkle Bereiche deuten mehr Fahrpläne an dieser Stelle an). Abbildungen 11.11(c) und 11.11(d) zeigen respektive die durch den Dekoder generierten Mengen. Abbildungen 11.11(e) und 11.11(f) zeigen 2-dimensionale Einbettungen der beiden Mengen erzeugt durch PCA bzw. RDP.

paarweise miteinander verglichenen Samples  $S$  entsprechender Größe. Der mittlere Fehler ist

$$\overline{E}_{d_H}^{(n)} = \frac{\sum_{i,j;i \neq j} d_H(S_i, S_j)}{\binom{m}{2}}, \quad (11.10)$$

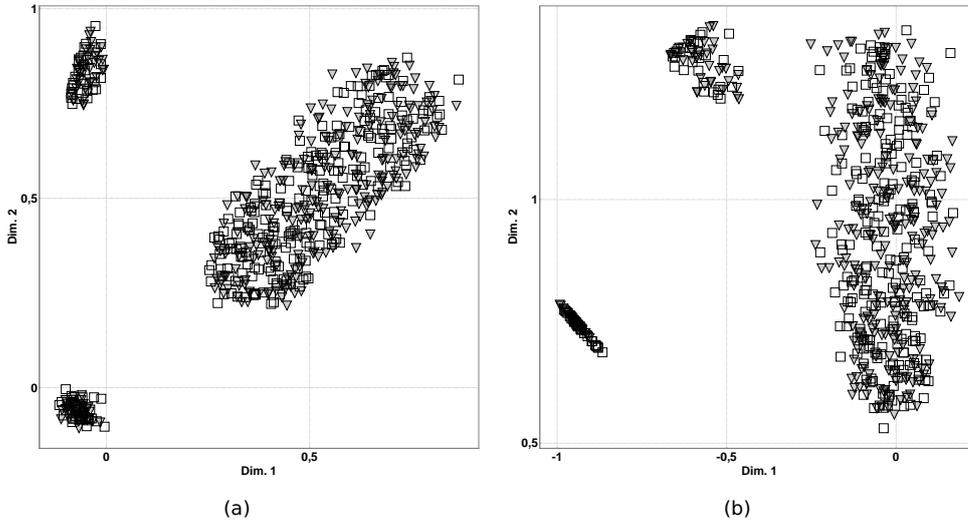


Abbildung 11.12.: Vergleich von Simulationsmodell und Dekoder für hochdimensionale Fahrpläne eines BHKW. Der Vergleich der beiden Punktwolken erfolgt hier einmal mittels Sammon-Projektion [11.12\(a\)](#) und mittels Hauptkomponentenanalyse [11.12\(b\)](#).

für  $m$  zufällige Trainingsmengen der Größe  $n$ . Hieraus kann ein modifizierter Hausdorffabstand konstruiert werden, der um den Erwartungswert des Fehlers bereinigt ist:

$$\widehat{d}_H^{(n)} = d_H - \overline{E}_{d_h}^{(n)}. \quad (11.11)$$

Für eine bessere Vergleichbarkeit wird dieser noch skaliert mit Hilfe des größten zu erwartenden Fehlers. Da alle Fahrpläne (wir betrachten hier nur skalierte) im Einheitshyperwürfel  $[0, 1]^d$  angesiedelt sind, wird als schlechteste Ausgabe des Dekoders eine gleichverteilte Punktwolke in  $[0, 1]^d$  angenommen. Dies ist die Ausgabe, wenn der Dekoder überhaupt keine Wirkung hat und somit eine Menge zufällig generierter Fahrpläne bei der Abbildung in den Lösungsraum unverändert lässt. Das hieraus konstruierte Distanzmaß für eine Fahrplanmenge hat nunmehr die folgende Form:

$$\widetilde{d}_H^{(n)} = \frac{\widehat{d}_H^{(n)}}{\max_{i,k} \{d_H(S_i, S_k) \mid S_i \in \mathcal{F}_{[0,1]}, S_k \in [0, 1]^d \sim U(0, 1)^d\}}, \quad (11.12)$$

wobei das Maximum für die Worst-Case-Abschätzung anhand von  $i$  durch das Einheits-simulationsmodell generierten und  $k$  zufällig generierten Fahrplanmengen vorab bestimmt und wiederverwendet werden kann. Ein Wert von  $\widetilde{d}_H^{(n)} = 0$  kennzeichnet eine gute Übereinstimmung.

Abbildung [11.13](#) zeigt eine beispielhafte Anwendung des Hausdorffmaßes auf den Dekoder eines Wärmepumpenmodells. Für verschiedene Bandbreiten (vgl. Abschnitt [5.3.4](#) auf Seite [70](#)) wurde das Hausdorff-Distanzmaß  $\widetilde{d}_H^{(n)}$  für verschiedene Punktzahlen ( $n \in \{100, 500, 2000\}$ ) und verschiedene Fahrplanlängen angewandt. Der Abstand wurde für Mengen von Originalfahrplänen vom Simulationsmodell und von vom Dekoder generierten Mengen an zufälligen Fahrplänen bestimmt. Wie man leicht erkennen kann (z. B. in [11.13\(c\)](#) bei einer Bandbreite von etwa 0.75) kann das Hausdorffmaß bei zu kleinen

## 11. Evaluation

Punktmengen gegebenenfalls zu Fehlinterpretationen führen, da in diesem Fall offenbar Abstände vorkommen können, die größer sind als der angenommene Worst-Case-Abstand bei wirkungslosem Dekoder. Dies liegt offenbar an der bekannten Sensitivität gegenüber Ausreißern (Agarwal et al. 2010), die bei kleinen Punktwolken stärker zur Ausprägung kommt. Zu erkennen ist dieser Umstand auch an den negativen Werten. Wird der gleiche Versuch mit einer größeren Zahl an Fahrplänen (blaue bzw. grüne Linie) durchgeführt, so treten diese Minima mit negativen Werten nicht mehr auf und das wahre Minimum wird sichtbar. Wichtig ist dieser Umstand bei der Verwendung des Hausdorff-Ansatzes zur Bestimmung guter Parameter für die Bestimmung von SVDD-Modell bzw. dem hieraus abgeleiteten Dekoder. Ein minimaler Hausdorff-Abstand zwischen Trainingsmenge und Dekoder-Fahrplänen deutet auf eine gute Dekoder-Leistung hin. Ist jedoch dieser Abstand für kleine Fahrplanmengen negativ, so ist dies auf Unzulänglichkeiten des Hausdorff-Abstands zurückzuführen. Solche Parameterwerte erzeugen keine guten Dekoder. Zur Absicherung sollte also immer ein zweiter Kennwert bei der Parameterschätzung betrachtet werden.

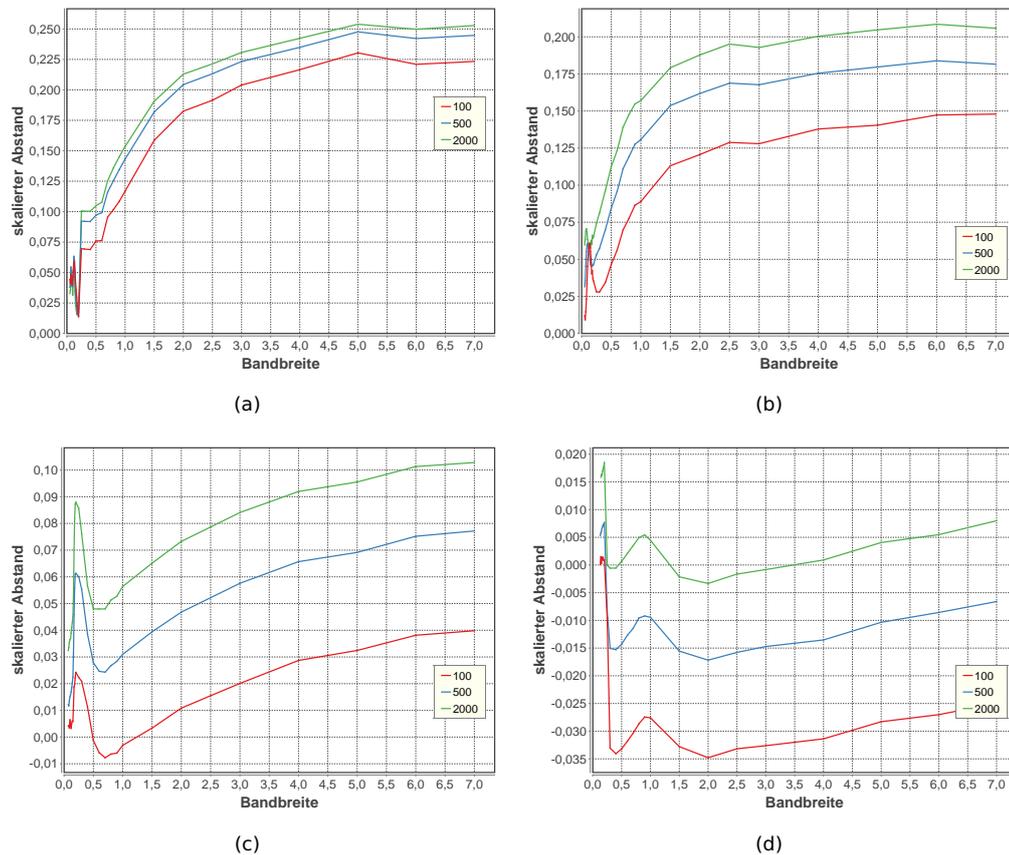


Abbildung 11.13.: Hausdorffabstand zwischen Wärmepumpenmodell und Dekoder für verschiedene Fahrplanlängen (2, 8, 32 und 96 von rechts oben nach links unten) und verschiedene Samplegrößen bei variierender Bandbreite.

Abbildung 11.14 zeigt als Ausschnittvergrößerung den 32-dimensionalen Fall dieses Versuchs zusammen mit Zeitreihendarstellungen von vom Dekoder generierten Fahr-

planmengen an ausgewählten Punkten zum Vergleich. Abbildung 11.15 zeigt den Fall eines BHKW für einen Zeitraum von 8 15-Minutenintervallen. Hier wurden zusätzlich zum Hausdorffmaß noch weitere Kennzahlen für die Korrektheit der vom Dekoder erzeugten Fahrpläne bestimmt. Hierfür wurden zwei Maße konstruiert:

- Die Korrektreproduktionsrate bezeichnet den Anteil der vom Dekoder erzeugten Fahrpläne, die vom Simulationsmodell als korrekt (also realisierbar) eingestuft werden.
- Die Korrektklassifikation der Positivklasse kennzeichnet den Anteil der vom Simulationsmodell durch Sampling erzeugten Fahrpläne die vom Klassifikator als korrekt eingestuft wurden.

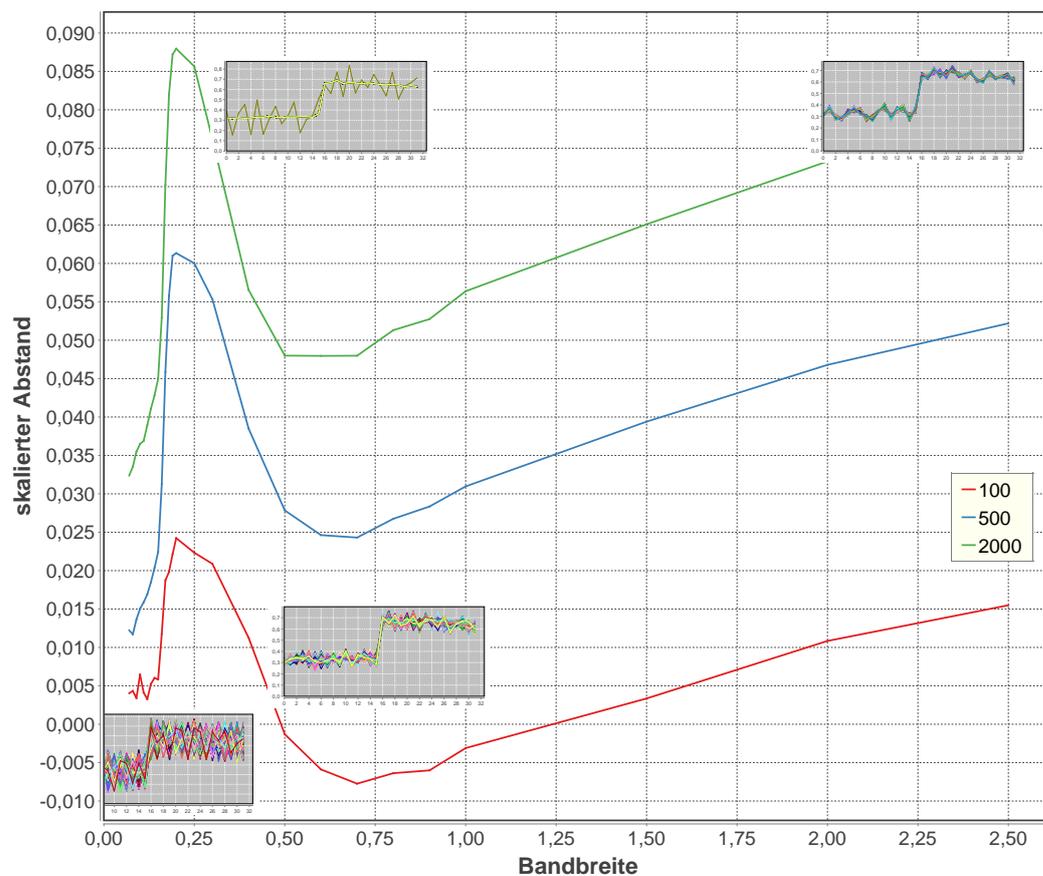


Abbildung 11.14.: Ausschnittvergrößerung mit resultierenden Fahrplanmengen (durch den jeweils resultierenden Dekoder zufällig generiert).

Diese beiden Maße sind zum Vergleich zusammen mit dem Hausdorffmaß aufgetragen. Bei der Korrektreproduktionsrate wurde eine Toleranzschwelle von 20 Watt (bei 4.7 kW Maximalleistung) vorgegeben unterhalb derer Fehler ignoriert wurden. Daher wird ab einem gewissen Punkt auch eine Rate von 1 erreicht. Man sieht in dieser Darstellung auch, dass ab einem gewissen Punkt der Dekoder besser ist als der Klassifikator. Diese Beobachtung deckt sich mit eingangs in Kapitel 9 ab Seite 117 gemachten Beobachtungen wonach die vom Dekoder erzeugten Punktmengen insbesondere an Kanten und Ecken des kodierten Unterraums schärfer mit dem ursprünglichen Sample übereinstimmen als die

## 11. Evaluation

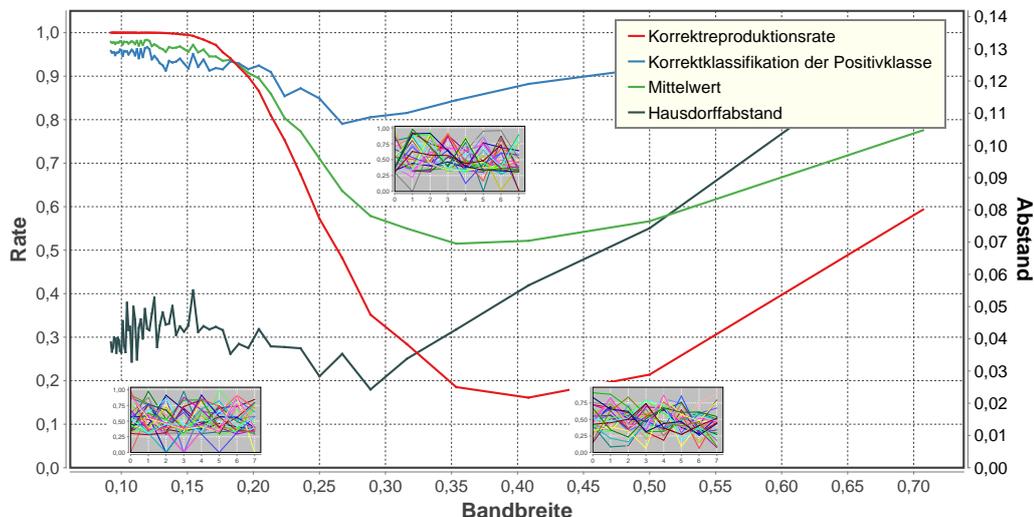


Abbildung 11.15.: Vergleich von Hausdorffabstand und Korrektklassifikationsraten für die Beurteilung der Dekoderleistung.

eher rundlich geformte Entscheidungsgrenze des Klassifikators (vgl. Abbildungen 9.5 auf Seite 128 und 9.7 auf Seite 130).

Insgesamt lässt sich an den erreichbaren Hausdorff-Abständen bei geeigneter Wahl der Bandbreite erkennen, dass – zumindest in den hier betrachteten Fällen – eine gute Überdeckung zwischen Trainingsmenge und Dekoderfahrplänen erreichbar ist. Somit bildet der Dekoder nicht nur mit akzeptabler Fehlerrate die Fahrpläne in den Bereich der gültigen Fahrpläne ab, sondern nutzt den Suchraum auch sehr gut aus. Somit wird durch den Dekoder die ursprüngliche Flexibilität der Anlage sehr gut abgebildet.

In diesem Kapitel wurden im Wesentlichen die Häufigkeiten betrachtet, mit denen mögliche Fehler auftreten können. Nun sagt die Häufigkeit noch nichts über die Größe der Fehler aus oder über die möglichen Auswirkungen. Diese aber sind abhängig vom konkreten Anwendungsfall und werden daher im nachfolgenden Kapitel im Anwendungskontext betrachtet. Prinzipiell wird bei allen Anwendungsfällen davon ausgegangen, dass eine Lösung eine Sollvorgabe für die Steuerung einer Anlage darstellt und nicht unreflektiert abgefahren sondern vielmehr von einer Steuerungseinheit umgesetzt wird, die in der Lage ist kleine Fehler zu korrigieren (z. B. eine technisch gesehen unzulässige 10 Watt Erzeugungsvorgabe für ein 4.7 kW BHKW automatisch auf 0 kW abzubilden).

### 11.4.3. Kennzahlen für die Mehrzieloptimierung

Abschließend muss noch die Integration von Kennzahlen für die differenzierte Bewertung von Fahrplänen im Rahmen von Mehrzieloptimierungsansätzen evaluiert werden. Hierfür ist insbesondere wichtig, wie gut die Kennzahlen vom Dekoder rekonstruiert werden können. Es stellt sich also die Frage nach der Abweichung von der vom Dekoder erzeugten Kennzahl (durch Abbilden eines Fahrplans mit Kennzahl in den zulässigen Raum) und der tatsächlichen Kennzahl, wie sie durch die zugehörige Funktionsvorschrift berechnet

werden würde. Interessant hierbei ist auch, wie sich die funktionale Komplexität bei der Berechnung der Kennzahl auf die erreichbare Genauigkeit auswirkt, da hierdurch Rückschlüsse möglich sind, welche Kennzahlen sich gut oder weniger gut eignen. Hierzu werden hier zunächst einige Szenarien betrachtet, bei denen Fahrpläne mit verschiedenen komplexen artifizielle Kennzahlen annotiert werden.

Ein wichtiger Aspekt bei der Integration von Kennzahlen ist also die Genauigkeit mit der sie vom Dekoder rekonstruiert werden. Hierzu wurde eine Reihe von Versuchen durchgeführt. Zusätzlich zu den in (Bremer & Sonnenschein 2013a) veröffentlichten Werten gibt Tabelle 11.6 exemplarisch für ein BHKW einige Beispiele. Verwendet wurden drei unterschiedlich komplexe artifizielle Kennzahlen:

**Linear** bezeichnet hier eine artifizielle Kennzahl, die sich über die Summennorm des Fahrplans skaliert mit der Länge  $\frac{1}{d} \|\cdot\|_1$  definiert. Da als Argument bereits mit skalierten Fahrplänen gearbeitet wird, ergibt sich hieraus bereits eine skalierte Kennzahl zur direkten Verwendung.

**Quadratisch** bezeichnet analog die  $L_2$ -Norm um eine artifizielle Kennzahl mit quadratischem Zusammenhang abzubilden.

**Rosenbrock** schließlich bezeichnet die  $d$ -dimensionale Rosenbrock-Funktion nach (vgl. z. B. Yao et al. 1999)  $\sum_{i=1}^d d - 1((1 - x_i)^2 + 100 \cdot (x_{i+1} - x_i^2)^2)$ , deren Wert für den gegebenen Fahrplanvektor berechnet und mit Hilfe des Maximalwerts der Funktion im möglichen Intervall skaliert wurde. Hiermit wird eine Kennzahl mit komplexem Zusammenhang nachgebildet.

Um den Fehler bei der Rekonstruktion durch den Dekoder zu bestimmen, wurde wie folgt vorgegangen. Als erstes wurde eine Trainingsmenge aus 800 skalierten Fahrplänen generiert. Für jeden dieser Fahrpläne wurde jede Kennzahl berechnet und jeweils als zusätzliches Element an den Fahrplan angehängt. Mit dieser Trainingsmenge wurde ein Dekoder  $\gamma$  erzeugt. Im nächsten Schritt wurde eine Menge von 10000 zufälligen Fahrplänen  $\mathbf{x}_i \in [0, 1]^{d+3} \sim U(0, 1)^{d+3}$  erzeugt und mit dem Dekoder auf gültige Fahrpläne  $\mathbf{y}_i = \gamma(\mathbf{x}_i)$  abgebildet. Hierdurch ergibt sich nunmehr eine Menge von realisierbaren Fahrplänen mit vom Dekoder automatisch zugeordneten Kennzahlen, die sich jeweils aus einem  $d$ -dimensionalen Fahrplanteil  $\mathbf{y}^{(p)} = (y_1, \dots, y_d)$  und einem Kennzahlenteil  $\mathbf{y}^{(Q)} = (y_{d+1}, \dots, y_{d+3})$  für die drei Kennzahlen zusammensetzt. Für jeden dieser Fahrpläne  $\mathbf{y}_i$  kann nun der Fehler bei der Kennzahlenrekonstruktion wie folgt berechnet werden:

$$\mathcal{E}_j = \mathcal{I}_{[0,1]}^{Q_j}(x^{(p)}, z_0) - y_j^{(Q_j)}. \quad (11.13)$$

Der Fehler stellt somit die Differenz zwischen (anhand des Simulationsmodells) berechnetem tatsächlicher Kennzahl und der rekonstruierten Kennzahl dar.

Der Fehler für Kennzahlen mit linearem Zusammenhang, wie er beispielsweise für Brennstoffkosten o. ä. zu erwarten wäre, ist erstaunlich gering und kann (zumindest in diesem Beispiel) unabhängig von der Länge der verwendeten Fahrpläne vernachlässigt werden. Aber auch für die beiden anderen Kennzahlklassen ist der mittlere Fehler durchaus akzeptabel klein. Für 32-dimensionale Fahrpläne ist der beobachtete Fehler größer. Dies dürfte daran liegen, dass in den verwendeten Szenarien gegen 7:00 Uhr jeweils ein größerer Warmwasserverbrauch für das morgendliche Duschen vorgesehen ist. Hierdurch steigt die Komplexität des Suchraums (vgl. auch Abschnitt 11.4.1 auf Seite 165 bzw.

## 11. Evaluation

---

Abbildung 11.8 auf Seite 166) sprunghaft an, womit offenbar auch ein größerer Fehler bei der Rekonstruktion anhängiger Kennzahlen einhergeht.

Tabelle 11.6.: Größe des Fehler bei der Kennzahlenrekonstruktion durch den Dekoder am Beispiel dreier artifiziieller Kennzahlen.

$d$	linear	quad.	Rosenbrock
8	$1.1319 \times 10^{-16} \pm 1.0890 \times 10^{-16}$	$1.2924 \times 10^{-4} \pm 3.5281 \times 10^{-4}$	$4.8359 \times 10^{-4} \pm 1.4051 \times 10^{-3}$
32	$1.3794 \times 10^{-16} \pm 1.1599 \times 10^{-16}$	$1.8861 \times 10^{-3} \pm 4.3726 \times 10^{-3}$	$8.3904 \times 10^{-3} \pm 1.9270 \times 10^{-2}$
96	$1.8223 \times 10^{-16} \pm 1.4216 \times 10^{-16}$	$6.1941 \times 10^{-5} \pm 5.0791 \times 10^{-5}$	$3.2065 \times 10^{-4} \pm 2.9402 \times 10^{-4}$

Wie die Versuche zeigen, erreicht der Ansatz, Kennzahlen mithilfe desselben Modells zu integrieren mit dem auch die zulässigen Leistungsbereiche einer Anlage modelliert werden, akzeptable Genauigkeiten, sodass die vom Dekoder rekonstruierten Werte im Rahmen einer Mehrzieloptimierung für den Anwendungsfall Wirkleistungsplanung verwendet werden können. Dass die Komplexität der Kennzahlen Auswirkung auf die Genauigkeit hat, mit der sie vom Model erfasst bzw. vom Dekoder rekonstruiert werden, ist unmittelbar einsichtig. Für einfache Kennzahlen mit linearem oder maximal quadratischem Zusammenhang zu den beschriebenen Fahrplänen (wie sie in der Wirkleistungsplanung zu erwarten sind) ist die erreichte Genauigkeit für den noch zu betrachtenden Anwendungsfall zunächst hinreichend.

### 11.5. Zusammenfassung

Wie eingangs des Kapitels gezeigt werden konnte, weist das Wirkleistungsplanungsproblem eine Komplexität auf, die den Einsatz von Heuristiken zumindest für nicht triviale Szenariengrößen notwendig macht. Um auf bestehende Heuristiken zurückgreifen zu können, ist eine dynamische Modellierung des Optimierungsproblems einschließlich der verteilten Information über individuelle Constraints hinsichtlich der Realisierbarkeit von Teillösungen durch die Anlagen erforderlich. Hierfür wurde ein Konzept entwickelt, das nunmehr zu evaluieren ist.

In diesem Kapitel wurden zum Zwecke der Evaluation zunächst die beiden wichtigsten Komponenten des Lösungsansatzes separat betrachtet: das Suchraummodell und der Dekoder. Hierbei konnte gezeigt werden, dass diese beiden Komponenten geeignet sind, die Modellierung individueller Fahrplanräume und die Erzeugung valider Lösungen im Rahmen einer Wirkleistungsplanung zu übernehmen.

Die Herleitung der verwendeten Szenarien sowie die verwendete Modellierung der beteiligten, dezentralen Anlagen wurden ebenfalls in diesem Kapitel eingeführt. In den nachfolgenden Kapiteln soll nunmehr der Gesamtansatz evaluiert werden; zum einen für den Fall einer zentralen Planung und zum anderen mit dezentralen Ansätzen.

# 12. Umsetzung mit zentralen Optimierungsansätzen

*The most effective way to do it, is to do it.*

*(Amelia Earhart)*

## 12.1. Einleitung

Es muss noch gezeigt werden, dass der vorgestellte Supportvektor Dekoderansatz geeignet ist, das Wirkleistungsplanungsproblem zu lösen. Hierfür muss das Konzept mit verschiedenen Optimierungsansätzen für die Planung kombiniert und mit geeigneten Szenarien für die Verbundplanung getestet werden. Als erstes soll an dieser Stelle die Integration des Supportvektor-Ansatzes in zentrale Optimierungsverfahren für die Wirkleistungsplanung demonstriert werden. Verteilte Ansätze für die Optimierung in dynamischen Verbänden werden im nächsten Kapitel gesondert betrachtet. Grundsätzlich ist ein Einsatz zentraler Ansätze im Rahmen zweier Szenarien denkbar:

**In statischen Virtuellen Kraftwerken** bestimmen alle beteiligten Anlagen ihren Lösungsraum dezentral und kommunizieren diesen an eine Steuerzentrale, welche dann die Planung für alle Anlagen übernimmt; oder

**eine zentrale Planungsautorität** in einem dynamischen Verbund löst hauptverantwortlich das Planungsproblem, um von globalem Wissen profitieren zu können.

Letzterer Fall tritt in dynamischen Verbänden ohne zentrale Optimierungsinstanz gegebenenfalls dennoch auf, wenn zusammengesetzte Einheiten (etwa eine Produktionsstraße oder mehrere BHKW in einer größeren Biogasanlage) durch einen einzelnen Controller gesteuert werden. Eine solche Einheit kann logisch als kleines statisches Virtuelles Kraftwerk innerhalb eines dynamischen Verbundes betrachtet werden. Der Verbund als Ganzes optimiert verteilt, aber der Controller der intern die zusammengesetzte Einheit im Verbund vertritt, muss intern zentral optimieren.

Es ist nicht Ziel zu evaluieren, welches Verfahren sich wie gut für die Lösung des Problems eignet. Vielmehr soll ein einheitlicher Weg aufgezeigt werden, wie das Dekoderkonzept eine einheitliche Integration in Standardoptimierungsverfahren ermöglicht. Zu diesem Zweck wird die Integration mit evolutionären bzw. populationsbasierten Heuristiken demonstriert, die von Haus aus nicht direkt für die Handhabung von Constraints konzipiert worden sind. Hierfür muss gezeigt werden, wie der Dekoder während der Optimierung verwendet wird und dass die Verfahren zu gültigen Lösungen konvergieren, die von ihrer Bewertung her besser sind als zufällig gewählte.

Die verwendeten Heuristiken werden zunächst kurz vorgestellt bevor die Integration des Suchraummodells in diese Heuristiken formal beschrieben wird. In Simulationen wird anschließend der integrierte Ansatz auf die im vorigen Kapitel beschriebenen Szenarien angewendet. Betrachtet werden ausschließlich Fahrpläne, die in Zeitintervalle mit 15 Minuten Dauer unterteilt sind. Dies stellt keine Einschränkung des Verfahrens an sich dar. Die Abstraktion als vektorielle Darstellung erlaubt prinzipiell die Darstellung von Fahrplänen mit beliebiger zeitlicher Auflösung (diese muss innerhalb eines Fahrplans nicht einmal konstant sein). Die Integration in Optimierungsansätze ist jedoch immer gleich, sodass die Betrachtung unterschiedlicher Auflösungen hier keine zusätzlichen Erkenntnisse brächte.

Anhand der dargestellten Beispiele wird auch noch einmal die generelle Wirkungsweise der Transformation des Optimierungsproblems in eine constraintfreie Formulierung aufgezeigt und visualisiert.

### 12.2. Verwendete Verfahren

In diesem Abschnitt werden zunächst fertig verfügbare Standardverfahren sowie eine speziell angepasste Eigenimplementierung des Simulated Annealing eingesetzt. Die Verfahren sollen hier einleitend kurz vorgestellt werden.

Die Computational Intelligence Library (CILib, <http://www.cilib.net>) stellt ein Open Source Framework mit einer Vielzahl an Algorithmen bereit (Cloete et al. 2008; Peer 2004). Insbesondere sind dies u. a. populationsbasierte Heuristiken. Der Vorteil bei der Verwendung eines solchen Frameworks liegt in der guten Abstraktion der Algorithmen, sodass dieselbe Implementierung eines Optimierungsproblems an verschiedene Algorithmeninstanzen zur Lösung gegeben werden kann. Im Rahmen dieser Arbeit werden exemplarisch folgende Verfahren aus dem Framework verwendet:

**Particle Swarm Optimization** (PSO) nach (Kennedy & Eberhart 1995) und (Shi & Eberhart 1998). PSO ist ein evolutionäres Optimierungsverfahren bei dem eine Population aus Lösungskandidaten iterativ verbessert wird durch Nachahmung von Schwarmverhalten. Durch Bewegungsvektoren welche sich an der besten bekannten Position einer gegebenen Nachbarschaft orientieren, werden die Lösungskandidaten in einer sich wiederholenden Prozedur durch den Lösungsraum bewegt und von guten Lösungen angezogen. Von PSO ist auch eine Variante speziell für kombinatorische Probleme bekannt (Jarboui et al. 2008).

**Artificial Bee Colony** (ABC) nach (Karaboga & Basturk 2007). Dieses Verfahren versucht ebenfalls schwarmbasiert in einer iterativen Prozedur eine Population von Lösungskandidaten sukzessive zu verbessern. Nachgebildet wird hier jedoch das Nahrungssucheverhalten eines Bienenschwarms.

**Harmony Search** (HS) nach (Lee & Geem 2005). Harmony Search ist eine Metaheuristik welche auf der Nachahmung der Improvisation von Musikern beruht und wird in der Forschung für eine Vielzahl von Anwendungsfeldern diskutiert (Abdel-raouf & Metwally 2013). HS eignet sich sowohl für diskrete als auch für kontinuierliche Suchräume

und Bedarf keiner ableitung der Zielfunktion. Die Anwendung auf höherdimensionale Problemstellungen wurde u. a. in (Omran & Mahdavi 2008) untersucht.

Von diesen Verfahren sind jeweils auch Mehrzielvarianten bekannt. Nachteil bei der Verwendung dieser Standardverfahren ist, dass sie in der vorliegenden Implementierung nicht speziell für kombinatorische Probleme sondern für numerische Funktionsoptimierung entworfen wurden. Das bedeutet, dass alle Verfahren nach dem Minimum (oder äquivalent dem Maximum) einer gegebenen Funktion  $f(x)$  suchen, wobei eine Lösung gegeben ist durch  $x \in \mathbb{R}^n$ . Die Lösungsvariation findet somit immer auch in  $\mathbb{R}^n$  statt, wobei allerdings die zulässige Domäne durch einen Boxconstraint (variabel für jede Dimension) vorgegeben werden kann. Dennoch ergibt sich im Fall der Wirkleistungsplanung als Dimension der Problemstellung  $d \cdot n$ , wenn für  $n$  Anlagen jeweils  $d$  Zeitintervalle betrachtet werden müssen. Die drei betrachteten Verfahren sind Repräsentanten für die Integration des Supportvektor-Ansatzes mit fertig verfügbaren Optimierungsansätzen.

Zusätzlich zu den gegebenen Verfahren aus der CILib-Bibliothek wurde eine Implementierung des Verfahrens Simulated Annealing umgesetzt und getestet, welches den geschilderten Nachteil nicht aufweist. Dieses Verfahren besitzt zwei Vorteile: zum einen lässt sich dieses Verfahren mit überschaubarem Aufwand implementieren, zum anderen zeigt dieses Verfahren gute Eigenschaften beim Entkommen aus lokalen Minima, was wie im vorigen Kapitel (vgl. 11.1.2 auf Seite 150) gezeigt in der Dekodervariante der Zielfunktion ein nicht unbedeutender Faktor ist.

Simulated Annealing (Kirkpatrick et al. 1983) basiert im Wesentlichen auf dem in Abbildung 12.1 dargestellten Ablauf. In Nachbildung eines Abkühlungsprozesses wird es dem Verfahren erlaubt zwischenzeitig auch schlechtere Lösungen (abhängig von Temperatur und Grad der Verschlechterung) zu akzeptieren, um aus lokalen Minima entkommen zu können. In der Umsetzung hier, besteht eine Lösung aus zwei Matrizen  $X_{ij}$  und  $M_{ij}$ , die für eine Anlage  $i$  für jedes Zeitintervall  $j$  skalierte Leistungswerte aus  $[0, 1]$  vorhalten.  $X$  wird mit zufälligen Werten initialisiert,  $M$  enthält die durch die jeweiligen Dekoder auf realisierbare Fahrpläne abgebildeten Vektoren und stellt somit die eigentliche Lösung, den skalierten Einsatzplan dar. In jeder Iteration wird nun aus  $X$  der Fahrplan genau einer Anlage zufällig gewählt und verändert (mutiert). Diese Mutation erfolgt an einem zufällig gewählten Zeitintervall  $k$ . Der Leistungswert  $p_k$  wird zu

$$p_k \leftarrow \begin{cases} p_k + r - 1 & \text{falls } p_k + r > 1 \\ p_k + r + 1 & \text{falls } p_k + r < 0 \\ p_k + r & \text{sonst,} \end{cases} \quad (12.1)$$

mit  $r \sim N(0, 1)$ . Statt nur einer können konfigurierbar auch mehrere solcher Mutationen je Iteration stattfinden, was insbesondere bei längeren Fahrplänen sinnvoll ist. Nur der mutierte Fahrplan muss im Anschluss mit dem zugehörigen Dekoder auf einen realisierbaren Fahrplan abgebildet werden, bevor eine Bewertung der neuen Lösung erfolgt.

Die so veränderte Lösung wird mit Hilfe der Zielfunktion  $f$  bewertet und mit der vorherigen Lösung verglichen. Ist die neue Lösung besser, wird sie beibehalten. Ansonsten wird sie nach dem Metropolis-Kriterium (Metropolis et al. 1953) mit einer gewissen Wahrscheinlichkeit akzeptiert und ansonsten verworfen. Der Vorteil bei dieser Umsetzung liegt darin, dass über den gesamten Zeitraum zwei Matrizen vorliegen. Alle Variationen finden in  $X$  statt, während das realisierbare Pendant  $M$  ebenfalls immer vorliegt. Die Variation einer

## 12. Umsetzung mit zentralen Optimierungsansätzen

---

```

 $\mathbf{X}_{ij} \leftarrow \mathbf{x}_i \sim U(0,1)^d, 1 \leq i \leq n$ 
 $\mathbf{M}_{ij} \leftarrow \gamma_i(\mathbf{X}_i), 1 \leq i \leq n$ 
 $\vartheta \leftarrow \vartheta_{start}$ 
while  $\vartheta < \vartheta_{min}$  do
  choose random  $k; 1 \leq k \leq n$ 
   $\mathbf{x}^* \leftarrow \mathbf{X}_k$ 
  mutate( $\mathbf{x}^*$ )
   $\mathbf{M}^* \leftarrow \mathbf{M}; \mathbf{M}_k^* \leftarrow \gamma_k(\mathbf{x}^*)$ 
  if  $e^{-\frac{f(\mathbf{M}^*)-f(\mathbf{M})}{\vartheta}} > r \sim U(0,1)$  then
     $\mathbf{M} \leftarrow \mathbf{M}^*; \mathbf{X}_k \leftarrow \mathbf{x}^*$ 
  end if
   $\vartheta \leftarrow \text{cooling} \cdot \vartheta$ 
end while

```

Abbildung 12.1.: Genereller Ablauf für Simulated Annealing mit Dekoder.

Lösung findet in jeder Iteration nur in einer Zeile in  $\mathbf{X}$  (auf einem Fahrplan) statt, sodass auf  $\mathbf{M}$  immer nur eine Update und keine komplette Neuberechnung durchgeführt werden muss. Dies ist jeweils durch nur eine Dekoderberechnung möglich.

### 12.3. Abbildungskonzept

Gegeben ist im einfachsten Fall ein nicht lineares Programm (NLP)

$$\begin{aligned}
 f(\mathbf{x}) \rightarrow \min \quad & ; \mathbf{x} \in \mathbb{R}^d \\
 \text{s.t.} & \\
 g_i(\mathbf{x}) = 0 \quad & ; 1 \leq i \leq m \\
 h_j(\mathbf{x}) \leq 0 \quad & ; 1 \leq j \leq n.
 \end{aligned} \tag{12.2}$$

In der Praxis kommen zu den linearen Constraints in (12.2) im allgemeinen Fall durchaus auch nicht-lineare Constraints  $\mathbf{x} \in \mathcal{F}$  zur Beachtung hinzu. Um gegebene Algorithmen unverändert verwenden zu können, wird die Zielfunktion mit Hilfe der Dekoder modifiziert. Statt Problem (12.2) direkt zu lösen wird das alternative Problem

$$f(\gamma(\mathbf{x})) \rightarrow \min \tag{12.3}$$

betrachtet. Hier sind keine Constraints zu beachten, da  $\gamma$  immer gültige Lösungen produziert. Zu beachten ist hierbei lediglich, dass die gefundene Lösung am Ende ebenfalls noch einmal durch den Dekoder abgebildet werden muss. Sei

$$\mathbf{x}_{opt} = \arg \min f(\gamma(\mathbf{x})), \tag{12.4}$$

dann ist  $\mathbf{x}_{min} = \gamma(\mathbf{x}_{opt})$  Lösung von (12.2). Ein evolutionäres Optimierungsverfahren dürfte also (im skalierten Fall) den vollständigen Einheitswürfel  $[0,1]^d$  durchsuchen ohne Constraints beachten zu müssen. Streng genommen stellt die Beschränkung auf Werte zwischen 0 und 1 auch einen Constraint – einen Boxconstraint – dar, derartige

Constraints können aber von praktisch allen Optimierungsverfahren einfach gehandhabt werden. Das Wirkleistungsplanungsproblem aus Gleichung (6.14) wird bei Verwendung von jeweils einem Dekoder für jede Anlage zu

$$\delta \left( \sum_{i=1}^d \gamma_i(\mathbf{p}_i), \zeta^{(P)} \right) \rightarrow \min. \quad (12.5)$$

Abbildung 12.2 zeigt ein erstes Beispiel für die Funktionsweise. Mit Hilfe eines PSO wurde für eine einzelne Anlage ein bestimmter Fahrplan gesucht. Dieser 2-dimensionale Trivialfall stellt eigentlich kein ernsthaftes Optimierungsproblem dar und dient lediglich zur Verdeutlichung. In der linken Hälfte 12.2(a) ist die Folge von Zwischenlösungen dargestellt wie sie von PSO im Datenraum aller Fahrpläne (dem Einheitswürfel  $[0, 1]^2$ ) nacheinander betrachtet wurden. Beginnend bei der rot dargestellten (zufällig generierten) Startlösung hat sich das Verfahren schrittweise zur grün dargestellten gesuchten Lösung vorgearbeitet. Evaluiert wurde jedoch der Pfad auf der rechten Seite 12.2(b), der durch Abbildung mittels des Dekoders entstand. Dieser bewegt sich vollständig im zulässigen Bereich (außerhalb der rot markierten durch Constraints unzulässigen Bereiche). Die gefundene Lösung ist somit auch der grüne Punkt auf der rechten Seite der Abbildung.

Ein zweites Beispiel mit nun vier Anlagen, die zeitgleich nach einem Summenlastgang abgesucht wurden, ist in Abbildung 12.3 dargestellt. Hier zeigt sich auch noch einmal, wie aus der Vektordarstellung die Zeitreihenlösung gewonnen wird.

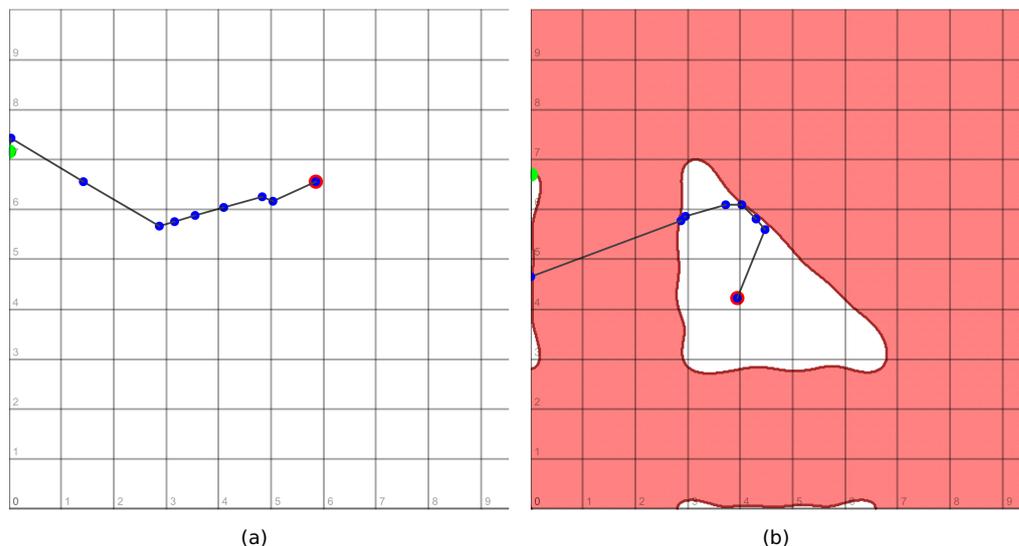


Abbildung 12.2.: Abbildung eines exemplarischen Suchpfades auf einen unbeschränkten mit Hilfe des Dekoders.

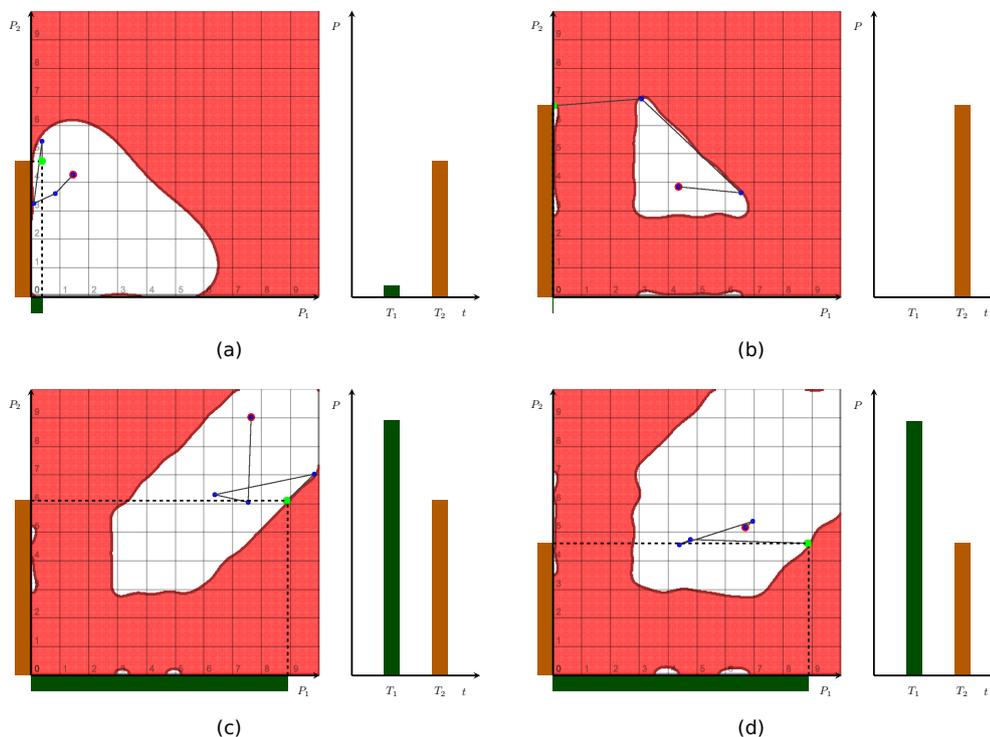


Abbildung 12.3.: Beispiel für abgebildete Suchpfade von vier Anlagen und die abschließende Lösungskonstruktion für den 2-dimensionalen Fall.

## 12.4. Simulative Evaluation

Dieser Abschnitt präsentiert eine Auswahl an Ergebnissen, die im Rahmen einer Simulationsstudie produziert worden sind. Die jeweiligen Verbünde für die die Wirkleistung zu planen ist, wurden in jeder Simulation nach den Annahmen aus Abschnitt 11.3.1 auf Seite 153 zufällig zusammengestellt, d.h. es gibt reine Erzeugerverbünde (BHKW) und gemischte Verbünde mit wenigen Verbrauchern zur Produktveredelung (Mixed). Die einzelnen Anlagen werden mit zufällig gewählter Parametrierung instanziiert. Das Produkt als Zielvorgabe ist entweder ein H0-Profil (zufällig gewählt für einen Tag des Jahres 2014), welches vom Verbund vermutlich nicht exakt gefahren werden kann, oder ein zufällig gewählter Lastverlauf (RND), der aber vom Verbund tatsächlich exakt gefahren werden kann. Letzterer wird erzeugt indem von jeder Anlage des Verbundes mit Hilfe des Simulationsmodells (durch das übliche Sampling) ein zufälliger, realisierbarer Fahrplan erzeugt wird. Dieser ist mit hoher Wahrscheinlichkeit nicht Bestandteil der Trainingsmenge für das Suchraummodell, da er separat erzeugt wird. Die Summe der so erzeugten Fahrpläne ergibt dann ein Produkt, das realisierbar ist. Von jeder Verbundtyp-Produkt-Kombination gibt es verschiedene Größen (Anzahl der Anlagen im Verbund).

Sofern nicht anders gekennzeichnet sind alle Simulationen über jeweils 100 Durchläufe und damit über 100 gelöste Wirkleistungsplanungsprobleme gemittelt, wobei für jeden

Durchlauf sowohl die Anlagen des Verbundes neu instanziiert wurden als auch das Produkt neu gewählt wurde.

### 12.4.1. Ergebnisse

Als erstes wurden die Standardverfahren aus der CILib-Bibliothek getestet. Die Tabellen 12.1 bis 12.3 zeigen die Ergebnisse. Angegeben sind verschiedene Kennzahlen zur Beurteilung des Ergebnisses:

$\delta$  gibt den Abstand – in den Beispielen hier den Euklidischen Abstand – zwischen dem Ziellastgang und dem erreichten Ergebnis an.

% **Startwert** gibt an, auf wie viel Prozent die Bewertung der Startlösung verbessert werden konnte.

**Iterationen** gibt an, über wie viele Iterationen das Verfahren durchgeführt wurde.

$n_{50\%}$  gibt an, nach wie vielen Iterationen im Mittel die Lösung auf 50% des Startwertes verbessert werden konnte.  $n_{10\%}$  und  $n_{5\%}$  geben entsprechend die Anzahl für 10 bzw. 5 Prozent an. Diese Werte geben einen Eindruck vom Konvergenzverhalten des jeweiligen Verfahrens.

**abs.  $\Delta E$**  gibt die absolute (über den gesamten Produktzeitraum gesehene) Abweichung der gelieferten von der durch das Produkt vorgegebenen Energiemenge an.

**rel.  $\Delta E$**  gibt die Energiemengenabweichung relativ zur vorgegebenen Energiemenge an, um die Abweichungen für verschiedene Produkte vergleichen zu können.

$\overline{\Delta P}$  gibt die mittlere absolute Abweichung der Leistungswerte je Zeitintervall an. Positive und negative Abweichungen werden hierbei gleich behandelt.

Im Folgenden sind zunächst die Ergebnisse für die Wirkleistungsplanung mit den CILib-Verfahren aufgeführt. Verplant wurden sowohl reine Erzeugerverbünde (BHKW) als auch gemischte Verbünde mit einigen Erzeugern zur Produktveredlung, wie in Abschnitt 11.3.1 auf Seite 153 definiert. Betrachtet wurden hier zunächst eher die kleinen Szenarien mit weniger Anlagen und kürzeren Produktzeiträumen. Die mit der CILib-Bibliothek erzielten

Tabelle 12.1.: Ergebnisse für die Optimierung mit ABC für Verbünde mit 10 Anlagen und einer Produktlänge von 8 Zeitintervallen.

Kennzahl	RND/ BHKW	H0/ BHKW	RND/ Mixed	H0/ Mixed
$\delta$	0.80212 ± 0.47278	0.95049 ± 0.83444	0.70307 ± 0.59569	0.60921 ± 0.45762
% Startwert	0.20219 ± 0.10874	0.20189 ± 0.13041	0.18465 ± 0.13242	0.18185 ± 0.13431
Iterationen	392.210 ± 209.490	374.280 ± 225.205	405.889 ± 244.201	391.192 ± 258.690
$n_{50\%}$	68.535 ± 41.133	67.732 ± 33.969	66.221 ± 34.960	57.165 ± 30.887
$n_{10\%}$	489.812 ± 185.491	400.789 ± 96.069	446.103 ± 146.737	381.667 ± 151.801
$n_{5\%}$	486.000 ± 148.142	522.000 ± 12.000	565.909 ± 144.103	548.667 ± 229.814
abs. $\Delta E$ / kWh	0.46175 ± 0.26840	0.55267 ± 0.48036	0.41160 ± 0.33838	0.35580 ± 0.26598
rel. $\Delta E$	0.00969 ± 0.00586	0.01138 ± 0.00993	0.01455 ± 0.01210	0.01260 ± 0.00965
$\overline{\Delta P}$ / kW	0.23087 ± 0.13420	0.27634 ± 0.24018	0.20580 ± 0.16919	0.17790 ± 0.13299

Ergebnisse sind bereits sehr gut, könnten aber durch Anpassung verschiedener Parameter und insbesondere durch Anpassung der Stopp-Bedingungen noch verbessert werden. Ein derartiges Feintuning wurde im Rahmen dieser Arbeit nicht mehr vorgenommen. Alle

## 12. Umsetzung mit zentralen Optimierungsansätzen

Tabelle 12.2.: Ergebnisse für die Optimierung mit PSO für Verbünde mit 10 Anlagen und einer Produktlänge von 8 Zeitintervallen.

Kennzahl	RND/ BHKW	H0/ BHKW	RND/ Mixed	H0/ Mixed
$\delta$	0.59857 ± 0.26386	0.69636 ± 0.32449	0.43515 ± 0.28886	0.34389 ± 0.26606
% Startwert	0.13504 ± 0.05968	0.14549 ± 0.06228	0.11355 ± 0.07361	0.09606 ± 0.06495
Iterationen	1556.630 ± 619.377	1297.180 ± 687.871	1378.200 ± 692.139	1548.220 ± 599.163
$n_{50\%}$	18.250 ± 16.403	19.820 ± 20.751	19.530 ± 16.444	19.740 ± 26.561
$n_{10\%}$	757.033 ± 458.303	674.160 ± 512.228	561.735 ± 456.268	644.754 ± 413.980
$n_{5\%}$	1240.600 ± 519.284	539.000 ± 350.047	800.333 ± 433.124	1039.458 ± 501.942
abs. $\Delta E$ / kWh	0.34470 ± 0.15546	0.40566 ± 0.19252	0.25204 ± 0.16669	0.19991 ± 0.15618
rel. $\Delta E$	0.00720 ± 0.00332	0.00854 ± 0.00410	0.00893 ± 0.00603	0.00714 ± 0.00528
$\Delta \bar{P}$ / kW	0.17235 ± 0.07773	0.20283 ± 0.09626	0.12602 ± 0.08335	0.09995 ± 0.07809

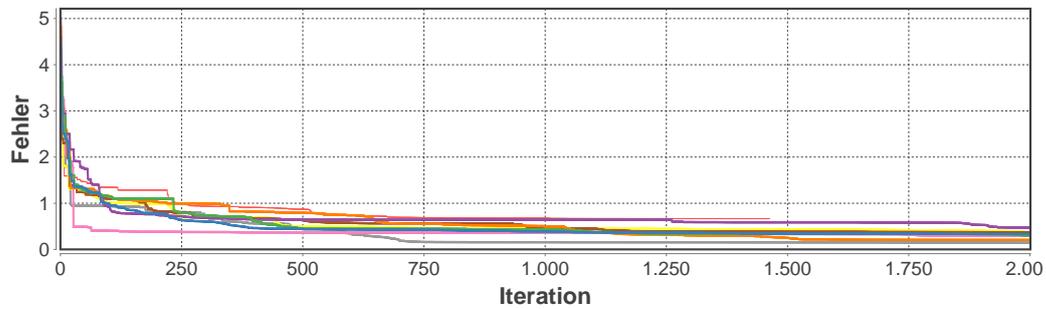
Tabelle 12.3.: Ergebnisse für die Optimierung mit HS für Verbünde mit 10 Anlagen und einer Produktlänge von 8 Zeitintervallen.

Kennzahl	RND/ BHKW	H0/ BHKW	RND/ Mixed	H0/ Mixed
$\delta$	0.77828 ± 0.53589	0.91631 ± 0.56203	0.41295 ± 0.37835	0.43355 ± 0.37802
% Startwert	0.05185 ± 0.03609	0.05832 ± 0.04076	0.02082 ± 0.02575	0.02703 ± 0.03522
Iterationen	250000 ± 0	250000 ± 0	250000 ± 0	250000 ± 0
$n_{50\%}$	339.260 ± 100.309	309.440 ± 90.865	280.869 ± 139.027	257.080 ± 107.527
$n_{10\%}$	59096.4 ± 49556.5	59068.1 ± 60829.3	27520.9 ± 37770.7	35228.6 ± 46579.6
$n_{5\%}$	60274.4 ± 43367.4	59793.0 ± 56419.2	32121.8 ± 40257.9	40002.5 ± 43938.6
abs. $\Delta E$ / kWh	0.45528 ± 0.31613	0.53317 ± 0.33364	0.23139 ± 0.20115	0.25333 ± 0.22146
rel. $\Delta E$	0.00933 ± 0.00635	0.01104 ± 0.00677	0.00801 ± 0.00660	0.00896 ± 0.00773
$\Delta \bar{P}$ / kW	0.22764 ± 0.15807	0.26659 ± 0.16682	0.11570 ± 0.10057	0.12666 ± 0.11073

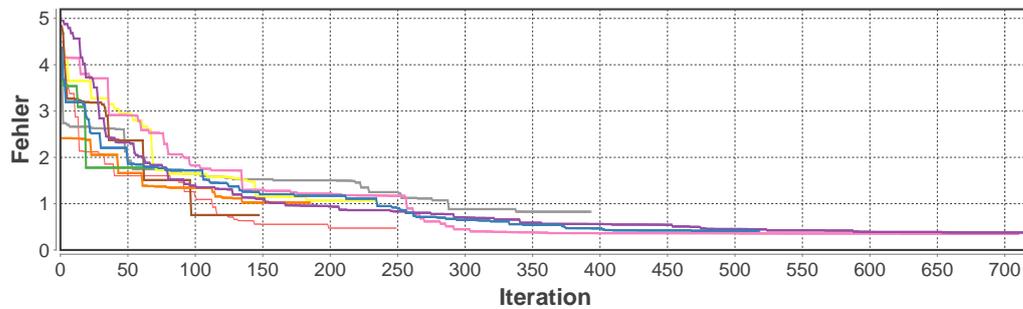
Verfahren wurden abgebrochen bei einem Stocken des Fortschritts; detektiert dadurch dass sich die Fitnesswerte über mehr als 10 Iterationen bei 8 Nachkommastellen nicht mehr verbessern. Für Harmony Search wurde zusätzlich eine obere Schranke von 250000 Iterationen eingeführt.

Absolut betrachtet, sind die Ergebnisse der verschiedenen Verfahren bezüglich der erreichten Lösungsgüte miteinander vergleichbar. Anzumerken ist hier noch, dass die Startlösung der beiden populationsbasierten Verfahren bereits die jeweils beste der 10 zufällig gewählten Individuen ist. Hierdurch starten sowohl PSO als auch ABC bereits mit einer relativ guten Startlösung, wodurch die erzielten Verbesserungen prozentual kleiner sind als bei HS. Zudem benötigen diese beiden Verfahren deutlich weniger Iterationen, sind aber in jeder Iteration langsamer, da für jedes Individuum eine Lösungsbewertung mit Dekoderabbildung vorgenommen wird. Bei HS ist es in jeder Iteration nur eine Bewertung. [Abbildung 12.4 auf der nächsten Seite](#) zeigt das Konvergenzverhalten der drei Algorithmen.

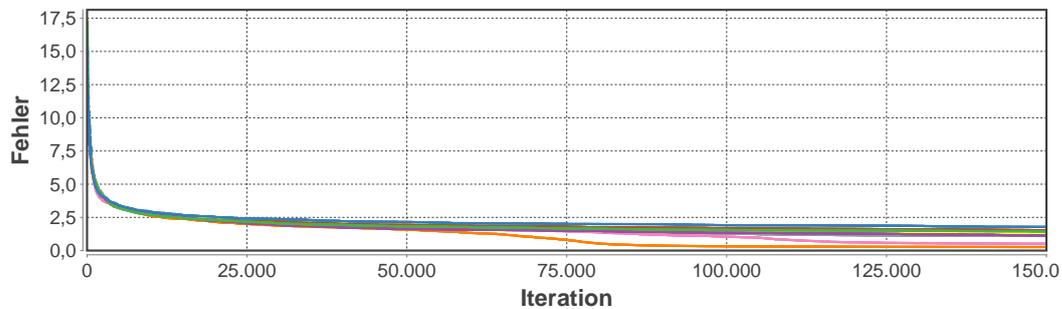
Ein großer Nachteil bei der Verwendung der CILib-Verfahren liegt in dem Umstand, dass hier die Fahrpläne der einzelnen Anlagen zu einer entsprechend hochstelligen Funktion kombiniert werden müssen, wodurch diese Verfahren aufgrund der sehr schnell wachsenden Dimensionalität des Problems an ihre Grenzen kommen. Dies liegt weniger in den Verfahren selbst als vielmehr in der Implementierung der Verfahren begründet. Es können somit leider nur Verbünde mit wenigen Anlagen oder über wenige Zeitintervall optimiert werden. Diese Einschränkung war vorab nicht bekannt oder dokumentiert. Ähnliche Implementierungen anderer Bibliotheken hätten diese Einschränkung eventuell nicht.



(a)



(b)



(c)

Abbildung 12.4.: Konvergenzverhalten der CILib Standardverfahren mit Dekoder für einen Erzeugerverbund aus 10 BHKW mit erreichbarem Ziellastgang; v.o.n.u.: PSO, ABC und HS. Die beiden populationsbasierten Verfahren 12.4(a) und 12.4(b) waren mit einem zusätzlichen Abbruchkriterium (bei stockendem Fortschritt) ausgestattet, sodass nicht bei allen Läufen die maximale Anzahl an Iterationen ausgenutzt wurde.

## 12. Umsetzung mit zentralen Optimierungsansätzen

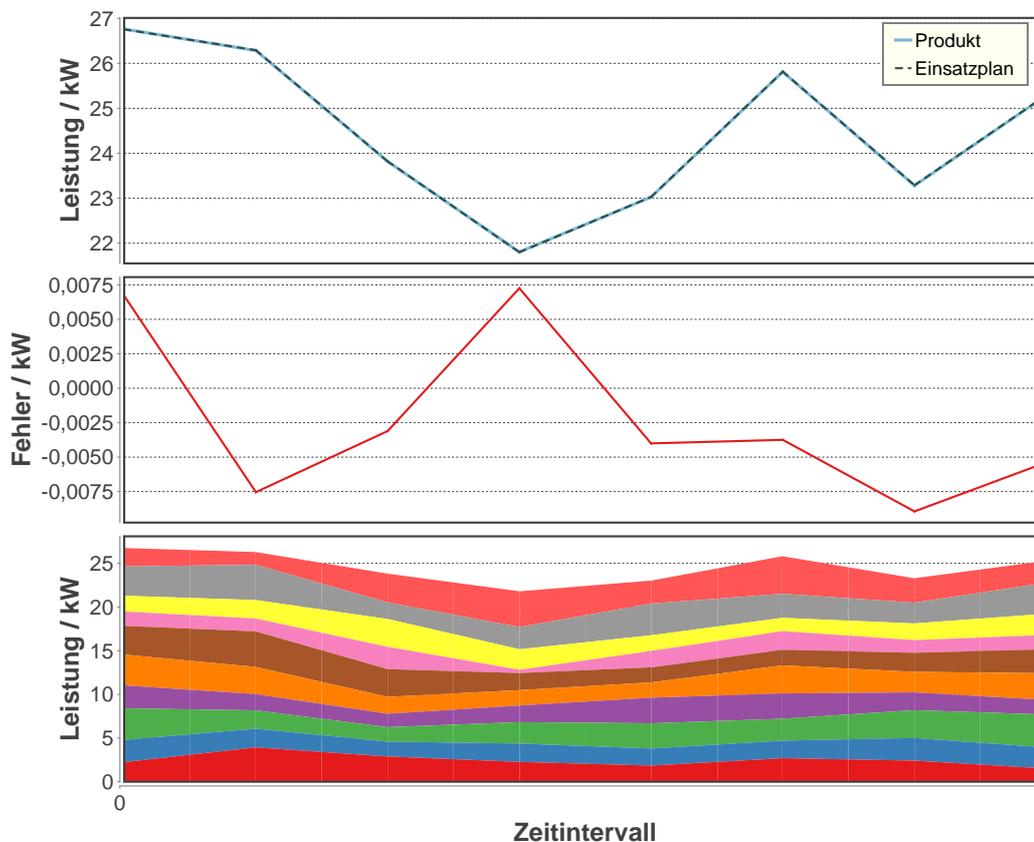


Abbildung 12.5.: Beispielergebnis für die Einsatzplanerstellung mit Simulated Annealing für 10 BHKW und einem garantiert erreichbaren Produkt. Das Verfahren wurde für schnelle Ausführung (durch schnelle Abkühlung) parametrisiert.

Für das selbst implementierte Simulated Annealing gilt dies nicht. Die beiden Abbildungen 12.5 und 12.6 zeigen zwei Beispielergebnisse, die mit Hilfe des SA erzielt wurden. Das Verfahren wurde für den zweiten Fall (ein Szenario mit Ganztagesfahrplänen) entsprechend mit einer höheren Mutationsrate und einer langsameren Abkühlung parametrisiert, sodass in Szenarien mit einem tatsächlich erreichbaren Zielprodukt, auch beinahe immer eine passende Kombination von Fahrplänen gefunden wurde (vgl. Fehler in Abbildung 12.6). In dem einfacheren Fall 12.5 wurde eine 50-fach schnellere Abkühlung verwendet. In beiden Fällen wurden reine Erzeugerverbünde mit jeweils 10 zufällig initialisierten BHKW verplant.

Eine Auswahl an umfangreicheren Ergebnissen findet sich in den Tabellen 12.4 bis 12.9, wobei wieder die vorab bereits erläuterten Kennzahlen zu den Ergebnissen bestimmt wurden. Für diese Ergebnisse wurden die verschiedensten Kombinationen von Verbundzenarien und Produkten (erreichbaren sowie nicht erreichbaren) jeweils in einer Serie von Simulationen betrachtet. Generell ist zu diesen Ergebnissen zunächst anzumerken, dass eine einzelne Iteration (wie auch bei HS) sehr viel schneller zu berechnen ist als bei den populationsbasierten Verfahren, so dass ein übergreifender Vergleich der benötigten Iterationen bei verschiedenen Algorithmen nicht sinnvoll ist. Zudem ist im Falle des SA die Zahl der durchgeführten Iterationen für eine gegebene Abkühlrate konstant. Da kein

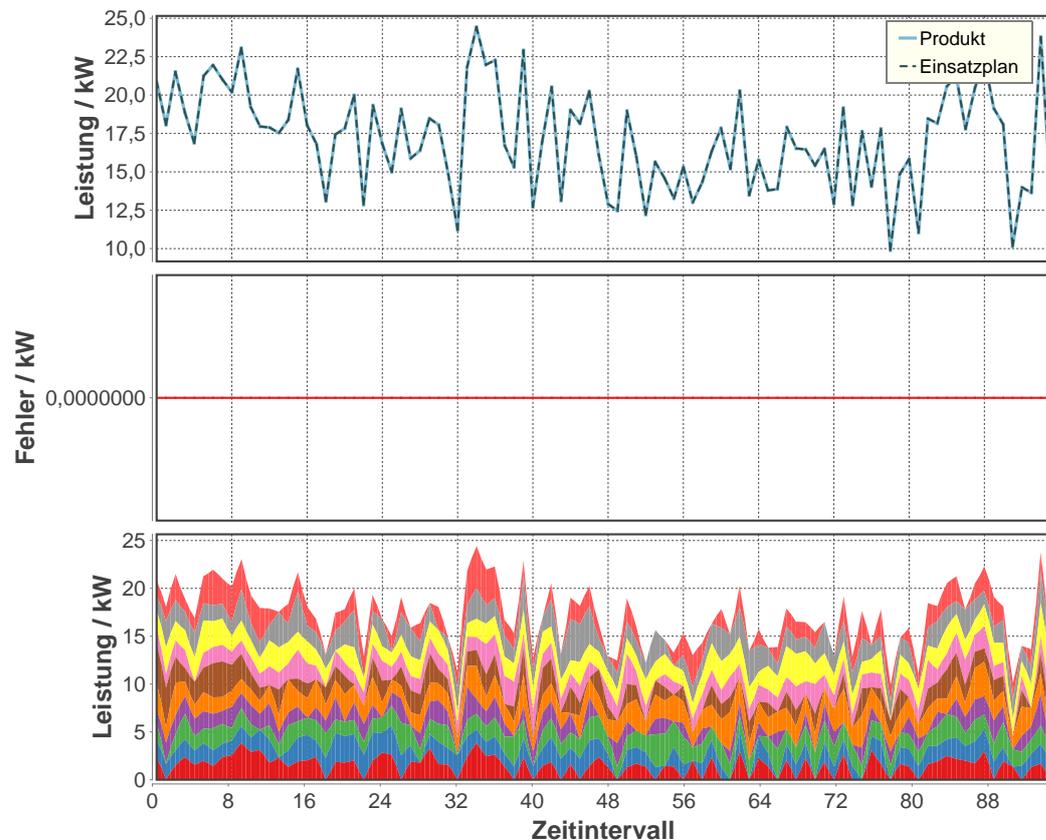


Abbildung 12.6.: Beispielergebnis für die Einsatzplanerstellung mit Simulated Annealing für 10 BHKW und einem garantiert erreichbaren Ganztages-Produkt. Das Verfahren wurde mit langsamerer Abkühlung parametrisiert, wodurch eine optimale Produktannäherung auch erreicht wurde.

zusätzliches Abbruchkriterium verwendet wurde, ist die Zahl der Iterationen innerhalb eines Versuches somit immer gleich.

Zudem ist in diesen Versuchen zu beobachten, dass eine Verbesserung des Zielwertes auf 5% der Startlösung oder weniger in einigen Fällen nicht erreicht wurde. Dies ist häufig bei einer Produktzielvorgabe der Fall, die von den Anlagen nicht erreicht werden kann. Hierfür gibt es prinzipiell zwei Ursachen. Das Verfahren startet bereits mit einem zufällig gewählten aber gültigen Fahrplan für jede Anlage. Daher ist die Startlösung meist bereits einigermaßen gut (zumindest bei höherdimensionalen Produkten), da ja alle Zielvorgaben zumindest größenordnungsmäßig so gewählt sind, dass der Verbund zum Produkt passt. Es kommt also nicht vor, dass beispielsweise ein Verbund aus nur wenigen häuslichen BHKW eine ganze Siedlung versorgen soll. Zum anderen ist naturgemäß bei nicht erreichbaren Produkten die erreichte Lösungsgüte nicht so hoch.

Insgesamt lässt sich aus diesen Ergebnissen ersehen, dass Simulated Annealing als zentraler Ansatz sehr gut für die Aufgabenstellung der Wirkleistungsplanung mit Supportvektor-Dekoder geeignet ist. Auch große Verbünde über längere Zeiträume (z. B. Tabelle 12.9) lassen sich damit bewältigen, ohne Performanzprobleme zu bekommen. Die erreichte Lösungsgüte scheint hierbei akzeptabel: mit einer durchschnittlichen Abweichung von der

## 12. Umsetzung mit zentralen Optimierungsansätzen

vorgegebenen Leistung von weniger als 1 kW bei 4700 kW Nennleistung in dem Verbund mit 1000 Anlagen. Ein Tradeoff zwischen Lösungsgüte und Performanz lässt sich jederzeit parametrieren. Die Tabellen 12.6 und 12.7 geben einen Vergleich der Ergebnisse für jeweils identische Szenarien, die aber mit unterschiedlichen Abkühlraten (etwa 20-facher Geschwindigkeitsvorteil) als Parametrierung geplant wurden.

Tabelle 12.4.: Ergebnisse für Simulated Annealing für Verbünde mit 10 Anlagen und einer Zielvorgabe für 8 Zeitintervalle.

Kennzahl	RND/ BHKW	H0/ BHKW	RND/ Mixed	H0/ Mixed
$\delta$	0.03773 ± 0.05467	0.02339 ± 0.03689	0.12797 ± 0.16311	0.10695 ± 0.12269
% Startwert	0.00364 ± 0.00523	0.00236 ± 0.00387	0.01546 ± 0.02045	0.01633 ± 0.01690
Iterationen	253272 ± 0	253272 ± 0	253272 ± 0	253272 ± 0
$n_{50\%}$	1330.833 ± 2635.333	633.333 ± 1123.236	2645.450 ± 5826.525	1239.667 ± 3792.743
$n_{10\%}$	102138.23 ± 9153.32	101143.70 ± 6656.29	104913.30 ± 4605.71	104514.63 ± 12354.02
$n_{5\%}$	116141.06 ± 8044.03	112142.07 ± 6272.31	114400.25 ± 6134.99	117132.70 ± 11683.98
abs. $\Delta E$ / kWh	0.02296 ± 0.03479	0.01321 ± 0.02029	0.02990 ± 0.03216	0.02633 ± 0.03094
rel. $\Delta E$	0.00048 ± 0.00072	0.00028 ± 0.00043	0.00104 ± 0.00110	0.00096 ± 0.00110
$\overline{\Delta P}$ / kW	0.01148 ± 0.01740	0.00661 ± 0.01015	0.01495 ± 0.01608	0.01316 ± 0.01547

Tabelle 12.5.: Ergebnisse für Simulated Annealing für Verbünde mit 10 Anlagen und einer Zielvorgabe für 32 Zeitintervalle. Einige prozentuale Verbesserungen der Startlösung konnten nicht erreicht werden (n/a).

Kennzahl	RND/ BHKW	H0/ BHKW	RND/ Mixed	H0/ Mixed
$\delta$	4.39121 ± 1.17663	31.76620 ± 5.96226	3.88776 ± 1.57257	17.25568 ± 3.44725
% Startwert	0.23128 ± 0.06241	0.56846 ± 0.05249	0.20841 ± 0.08708	0.50477 ± 0.07268
Iterationen	253272 ± 0	253272 ± 0	253272 ± 0	253272 ± 0
$n_{50\%}$	75052.650 ± 30095.615	105503.500 ± 6177.992	60620.500 ± 37002.046	106190.750 ± 6692.115
$n_{10\%}$	135132 ± 0	n/a	107804 ± 0	n/a
$n_{5\%}$	135132 ± 0	n/a	107804 ± 0	n/a
abs. $\Delta E$ / kWh	4.99476 ± 1.33255	37.12361 ± 8.29832	4.30753 ± 1.73967	20.11253 ± 4.76048
rel. $\Delta E$	0.03121 ± 0.00860	0.23763 ± 0.05223	0.05398 ± 0.02251	0.25629 ± 0.06083
$\overline{\Delta P}$ / kW	0.62434 ± 0.16657	4.64045 ± 1.03729	0.53844 ± 0.21746	2.51407 ± 0.59506

Tabelle 12.6.: Ergebnisse für Simulated Annealing für Verbünde mit 100 Anlagen und einer Zielvorgabe für 96 Zeitintervalle. Einige prozentuale Verbesserungen der Startlösung konnten nicht erreicht werden (n/a).

Kennzahl	RND/ BHKW	H0/ BHKW	RND/ Mixed	H0/ Mixed
$\delta$	13.8557 ± 0.8370	582.4645 ± 60.5922	11.59533 ± 0.9010	336.63273 ± 40.5767
% Startwert	0.12963 ± 0.01045	0.74167 ± 0.02917	0.12106 ± 0.00888	0.71449 ± 0.03145
Iterationen	506557 ± 0	506557 ± 0	506557 ± 0	506557 ± 0
$n_{50\%}$	189948.050 ± 3831.857	n/a	192140.800 ± 3643.218	n/a
$n_{10\%}$	n/a	n/a	n/a	n/a
$n_{5\%}$	n/a	n/a	n/a	n/a
abs. $\Delta E$ / kWh	27.2021 ± 1.8394	1178.1589 ± 143.1157	22.6151 ± 1.9935	634.8285 ± 69.7540
rel. $\Delta E$	0.00669 ± 0.00049	0.28806 ± 0.03330	0.01148 ± 0.00112	0.31771 ± 0.03650
$\overline{\Delta P}$ / kW	1.13342 ± 0.07664	49.08996 ± 5.96315	0.94230 ± 0.08307	26.45119 ± 2.90642

Einen ersten exemplarischen Eindruck von der Konvergenz des Verfahrens bei Einsatz mit Dekoder gibt Abbildung 12.7. Hieraus ist auch noch einmal anhand der verschiedenen Konvergenzgeschwindigkeiten – es wurden jeweils verschiedene Durchläufe auf demselben exemplarischen Szenario durchgeführt – zu ersehen, dass es sich lohnt, das Verfahren (wie

Tabelle 12.7.: Ergebnisse für Simulated Annealing für Verbünde mit 100 Anlagen und einer Zielvorgabe für 96 Zeitintervalle. Diesmal wurde der SA allerdings für schnelle Optimierung parametrisiert (schnellere Abkühlung). Einige prozentuale Verbesserungen der Startlösung konnten nicht erreicht werden (n/a).

Kennzahl	RND/ BHKW	H0/ BHKW	RND/ Mixed	H0/ Mixed
$\delta$	19.58500 $\pm$ 0.62450	632.46024 $\pm$ 64.08755	17.44437 $\pm$ 1.15366	367.03011 $\pm$ 34.11098
% Startwert	0.18444 $\pm$ 0.01756	0.79081 $\pm$ 0.02222	0.18003 $\pm$ 0.01480	0.76661 $\pm$ 0.02899
Iterationen	25316 $\pm$ 0	25316 $\pm$ 0	25316 $\pm$ 0	25316 $\pm$ 0
$n_{50\%}$	10533.500 $\pm$ 383.000	n/a	10660.600 $\pm$ 436.291	n/a
$n_{10\%}$	n/a	n/a	n/a	n/a
$n_{5\%}$	n/a	n/a	n/a	n/a
abs. $\Delta E$ / kWh	38.97183 $\pm$ 1.18493	1291.25824 $\pm$ 155.12134	33.92714 $\pm$ 2.57344	689.77319 $\pm$ 67.47972
rel. $\Delta E$	0.00951 $\pm$ 0.00037	0.31378 $\pm$ 0.03479	0.01733 $\pm$ 0.00164	0.34870 $\pm$ 0.03969
$\overline{\Delta P}$ / kW	1.62383 $\pm$ 0.04937	53.80243 $\pm$ 6.46339	1.41363 $\pm$ 0.10723	28.74055 $\pm$ 2.81165

Tabelle 12.8.: Ergebnisse für Simulated Annealing für Verbünde mit 1000 Anlagen und einer Zielvorgabe für 32 Zeitintervalle.

Kennzahl	RND/ BHKW	H0/ BHKW	RND/ Mixed	H0/ Mixed
$\delta$	1.7919 $\pm$ 0.1598	2503.0716 $\pm$ 1221.0867	0.6011 $\pm$ 0.1062	1311.1188 $\pm$ 629.8957
% Startwert	0.00965 $\pm$ 0.00162	0.45971 $\pm$ 0.19617	0.00352 $\pm$ 0.00063	0.38407 $\pm$ 0.1645
Iterationen	921021 $\pm$ 0	921021 $\pm$ 0	921021 $\pm$ 0	921021 $\pm$ 0
$n_{50\%}$	173071.33 $\pm$ 13878.57	377610.00 $\pm$ 141135.66	163889.73 $\pm$ 37328.39	471710.55 $\pm$ 173707.82
$n_{10\%}$	232760.30 $\pm$ 4714.14	343268.75 $\pm$ 5385.54	237678.87 $\pm$ 5316.43	495528.80 $\pm$ 209069.47
$n_{5\%}$	258613.60 $\pm$ 5912.11	389839.75 $\pm$ 16594.54	259556.97 $\pm$ 6416.39	276001.00 $\pm$ 4064.45
abs. $\Delta E$ / kWh	2.0389 $\pm$ 0.2242	2922.7371 $\pm$ 1503.7406	0.6750 $\pm$ 0.1025	1606.1906 $\pm$ 805.4022
rel. $\Delta E$	0.00013 $\pm$ 0.00001	0.18214 $\pm$ 0.09376	0.00008 $\pm$ 0.00001	0.19940 $\pm$ 0.10022
$\overline{\Delta P}$ / kW	0.2548 $\pm$ 0.0280	365.3421 $\pm$ 187.9675	0.0843 $\pm$ 0.0128	200.7738 $\pm$ 100.6752

bei Heuristiken durchaus üblich) mehrfach mit einer jeweils anderen zufälligen Startlösung zu starten.

Ein wichtiger Aspekt bei der Bewertung der Ergebnisse ist die Realisierbarkeit der Fahrpläne und damit die Gültigkeit der gefundenen Lösungen. Ziel des in dieser Arbeit entwickelten Ansatzes ist es, Anlagen-Constraints so zu kapseln, dass eine Einbindung der beschränkten Suchräume ohne Kenntnis der Anlagenmodellierung möglich wird und hierbei dennoch nur gültige Lösungen bei der Planung betrachtet werden. Bisher wurde nur die Güte der gefundenen Lösungen bezüglich der Zielfunktion betrachtet. Nun soll noch die Realisierbarkeit der gefundenen Einsatzpläne betrachtet werden. Hierfür werden

Tabelle 12.9.: Ergebnisse für Simulated Annealing für Verbünde mit 1000 Anlagen und einer Zielvorgabe für 96 Zeitintervalle. Für einige Szenarien wurde das Verfahren nicht hinreichend lange durchgeführt (kenntlich durch n/a) um innerhalb der begrenzten Zahl von Iterationen eine ausreichende Verbesserung der Startlösung zu bekommen.

Kennzahl	RND/ BHKW	H0/ BHKW	RND/ Mixed	H0/ Mixed
$\delta$	9.0616 $\pm$ 0.3661	6142.7702 $\pm$ 733.8334	4.5353 $\pm$ 0.3236	3214.6265 $\pm$ 227.6195
% Startwert	0.0264 $\pm$ 0.0017	0.76075 $\pm$ 0.0234	0.01493 $\pm$ 0.0020	0.70689 $\pm$ 0.0184
Iterationen	921021 $\pm$ 0	921021 $\pm$ 0	921021 $\pm$ 0	921021 $\pm$ 0
$n_{50\%}$	204162.850 $\pm$ 4720.919	n/a	207297.050 $\pm$ 5430.509	n/a
$n_{10\%}$	262013.800 $\pm$ 3138.201	n/a	266740.000 $\pm$ 4496.587	n/a
$n_{5\%}$	296886.650 $\pm$ 6112.611	n/a	296888.450 $\pm$ 6556.461	n/a
abs. $\Delta E$ / kWh	17.8227 $\pm$ 0.8633	12583.4960 $\pm$ 1920.5662	8.8627 $\pm$ 0.6306	6041.9604 $\pm$ 518.1090
rel. $\Delta E$	0.00044 $\pm$ 0.00002	0.30847 $\pm$ 0.04700	0.00044 $\pm$ 0.00003	0.30327 $\pm$ 0.02564
$\overline{\Delta P}$ / kW	0.74262 $\pm$ 0.03597	524.31234 $\pm$ 80.02359	0.36928 $\pm$ 0.02628	251.74835 $\pm$ 21.58788

## 12. Umsetzung mit zentralen Optimierungsansätzen

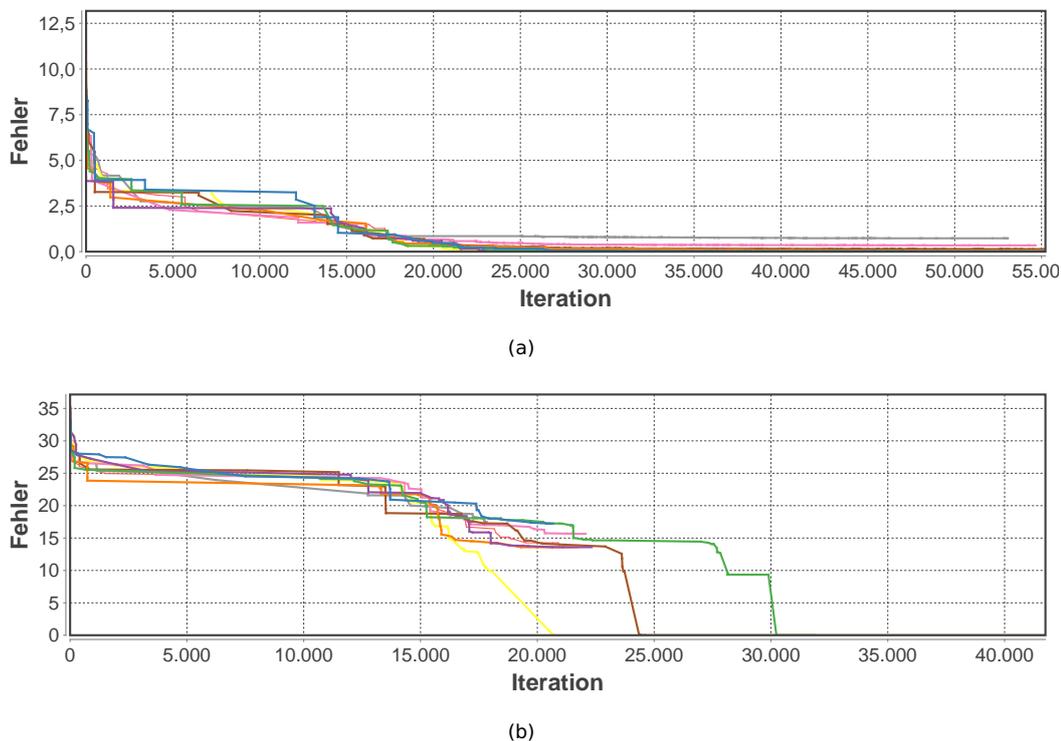


Abbildung 12.7.: Konvergenzverhalten des Simulated Annealing mit Dekoder an zwei Beispielen: Erzeugerverbünde mit 10 Anlagen für 8 bzw. 96 Zeitintervalle (das Verhältnis aus Mutationsrate und Fahrplandimension wurde beibehalten).

die einzelnen Fahrpläne des Einsatzplanes nach der Optimierung mit Hilfe der Anlagensimulationsmodelle auf ihre Durchführbarkeit geprüft. Tabelle 12.10 zeigt die Ergebnisse für zwei Szenarien mit 32-dimensionalen Fahrplänen.

Betrachtet wurden zwei Szenarien: reine Erzeugerverbünde und gemischte Verbünde. Betrachtet wurden jeweils Verbünde mit 100 Anlagen (bzw. 25 für PSO wg. der Dimensionsbeschränkung für die Zielfunktion) für die Einsatzpläne gesucht wurden. Als Produktvorgabe wurde ein H0-Profil gewählt, welches durch die Verbünde nicht erreichbar ist, um Lösungen jeweils am Rand der zulässigen Regionen zu provozieren.

Für die BHKW wurde in beiden Szenarien eine Toleranzschwelle von 20 Watt (bei 4.7 kW Nennleistung) vorgegeben unterhalb derer Fehler ignoriert wurden. Berechnet wurde für jedes Szenario die Realisierbarkeitsrate (gemittelt über 100 Durchläufe mit jeweils anders, zufällig konfigurierten Anlagen). Die Realisierbarkeitsrate  $r_{\mathcal{F}}$  gibt an, welcher Anteil der Fahrpläne im Einsatzplan realisierbar ist ohne eines der gegebenen Constraints zu verletzen.

Tabelle 12.11 zeigt analog die Realisierbarkeitsraten für Szenarien mit 96 Dimensionen und Verbünde mit 1000 Anlagen. Aufgrund der Beschränkungen für die Standard-CILib Algorithmen konnten hier bereits nur noch der SA verwendet werden.

Tabelle 12.10.: Exemplarische Realisierbarkeitsraten für zwei Szenarien und zwei Optimierungsverfahren über 32 Zeitintervalle: 100 BHKW (Szenario 1) und gemischter Verbund mit 100 Anlagen (Szenario 2). Vorgabe ist jeweils ein nicht erreichbares H0-Profil.

Algorithmus	Szenario 1	Szenario 2
PSO	0.99135 ± 0.01917	0.99387 ± 0.01498
SA	0.99510 ± 0.00767	0.99769 ± 0.00438

Tabelle 12.11.: Exemplarische Realisierbarkeitsraten für 96 Zeitintervalle und gemischte Verbünde: 10 Anlagen für ein H0-Profil (Szenario 1), 1000 Anlagen (Szenario 2) und 1000 Anlagen für erreichbare Produkte (Szenario 3).

Algorithmus	Szenario 1	Szenario 2	Szenario 3
SA	1.0 ± 0.0	0.9624 ± 0.00688	0.99998 ± 0.000141

Abbildung 12.8 visualisiert noch einmal die Einhaltung der Anlagenconstraints für einen reinen BHKW-Erzeugerverbund. Dargestellt sind neben dem Optimierungsergebnis und dem verbleibenden Restfehler auch die Fahrpläne der einzelnen Anlagen und die Temperaturverläufe in den jeweiligen Pufferspeichern (von oben nach unten). Die grauen Flächen kennzeichnen jeweils den zulässigen Bereich in dem sich die Verläufe bewegen dürfen. Weitere Constraints (z. B. maximale Leistungsänderungen) sind zwar im Szenario vorhanden, aber in der Grafik nicht sinnvoll darstellbar.

Die Ergebnisse bis hierher zeigen bereits, dass das Supportvektor-Dekoder-Verfahren geeignet ist, gängige Optimierungsalgorithmen derart zu erweitern, dass eine Behandlung von Constraints möglich wird. Hierbei können bestehende Implementierungen erweitert werden, durch Ergänzung der Zielfunktion um die Abbildung auf gültige Lösungen durch den Dekoder. Die Implementierung des eigentlichen Algorithmus braucht nicht geändert zu werden. Somit ist auch der Einsatz von etablierten und getesteten Algorithmen-Bibliotheken möglich. Durch Integration des Dekoders bereits in die Implementierung eines Optimierungsverfahrens – wie am Simulated Annealing Beispiel demonstriert – lässt sich der Dekoder zudem auch problemspezifisch für eine verbesserte Effizienz einsetzen.

Die Ergebnisse zeigen vor allen zwei Aspekte: Trotz des zusätzlichen Rechenaufwands durch den Einsatz des Dekoders lassen sich gute Lösungsgüten dennoch performant erreichen, aber durch Verwendung des Dekoder sind die gefundenen Lösungen durch die Anlagen auch realisierbar.

### 12.4.2. Vergleich mit Penalty-Ansatz

Abschließend soll noch evaluiert werden, wie der Supportvektor-Dekoder im Vergleich zu einer anderen Technik für die Constrainthandhabung abschneidet. Hierfür wurde exemplarisch der Penalty-Ansatz gewählt. Verglichen werden sollen allerdings nicht die jeweils erreichten Lösungsgüten, sondern lediglich der Anteil der gültigen Lösungen.

Hierfür wurde ein Penaltyverfahren wie folgt umgesetzt: Für jede Anlage wird das SVDD-Modell von Suchraum und Constraints – allerdings ohne den Dekoder – verwendet. Mit Hilfe dieser Modelle kann für die individuellen Lösungsbeiträge der einzelnen Anlagen ein Maß für die jeweilige Gültigkeit der Anlagenbeiträge bestimmt werden. Für den Fahrplan einer

## 12. Umsetzung mit zentralen Optimierungsansätzen

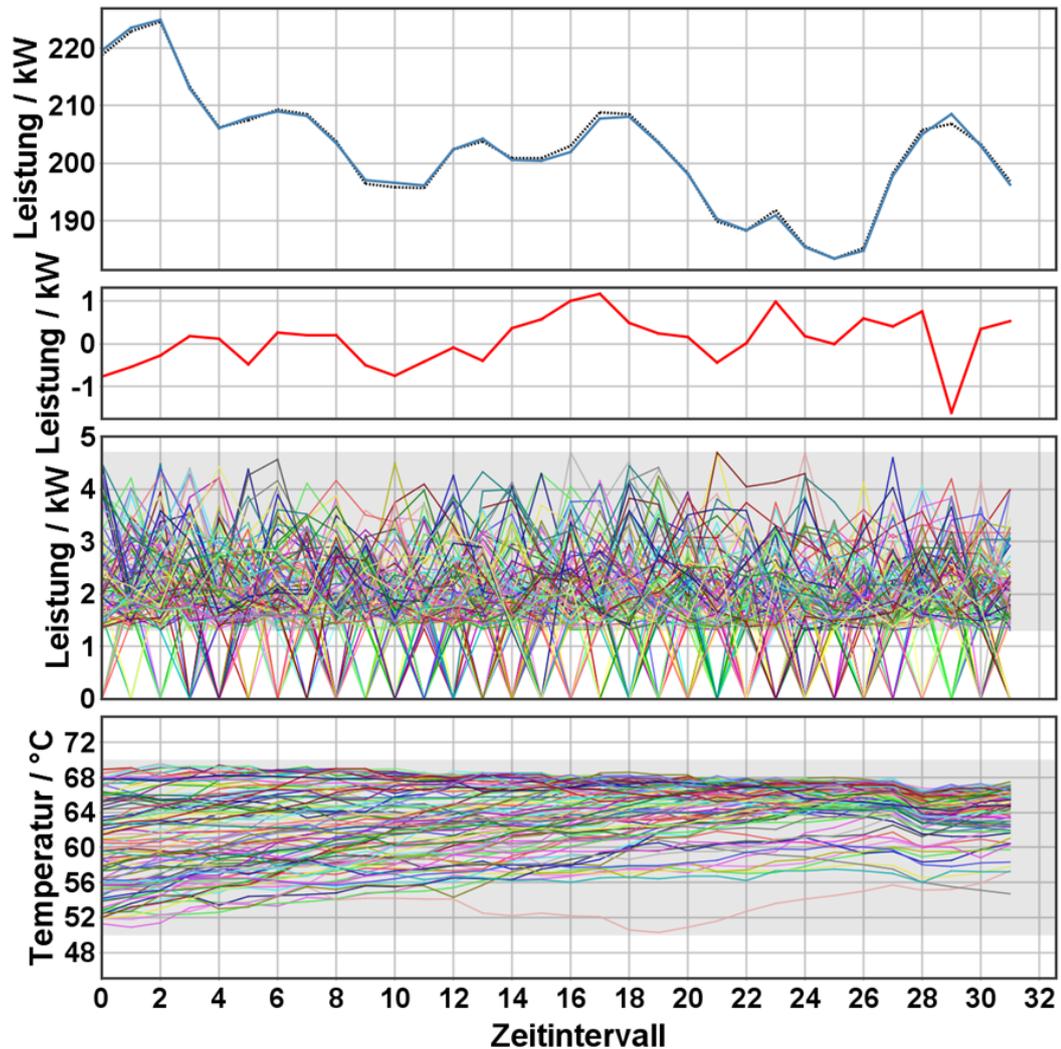


Abbildung 12.8.: Beispielergebnis mit Anlagenconstraints. Die grauen Flächen geben jeweils den zulässigen Bereich an, in dem sich Leistungswerte (mittlerer Graph) und Temperaturverlauf im Speicher (unten) bewegen dürfen. Bei den Leistungswerten ist zu beachten, dass laut Parametrierung ein jederzeitiges Abschalten ebenfalls zulässig war. Die jeweilige Anlage war dann während des gesamten Zeitintervalls inaktiv, nicht nur momentan, wie die Darstellung vermuten lassen könnte. Ein sofortiges Wiederabschalten war ohne hinreichende Betriebstemperatur jedoch nicht zulässig.

Anlage als Lösungsbeitrag kann mit dem SVDD-Modell der Abstand  $R_i$  vom Mittelpunkt der Hypersphäre (deren Urbild den zulässigen Lösungsraum modelliert) bestimmt werden. Dieser ist umso kleiner je näher eine Lösung am zulässigen Bereich ist. Hiermit wird nun ein externer (da keine zulässige Startlösung bekannt ist) Penalty-Ansatz für die Wirkleistungsplanung nach Abschnitt 6.5 auf Seite 86 wie folgt umgesetzt (vgl. auch Gleichung 6.14):

$$\omega_\delta \cdot \delta \left( \sum_{i=1}^n \mathbf{p}_i, \zeta^{(P)} \right) + \omega_{r_{\mathcal{F}}} \cdot \sum_{i=1}^d R_i(\mathbf{p}_i) \rightarrow \min \quad (12.6)$$

Der erste Term bewertet auch hier den erreichten Abstand  $\delta(\cdot, \cdot)$  zwischen dem Summenlastgang des Einsatzplans (durch individuelle Fahrpläne  $\mathbf{p}_i$  definiert) und dem angestrebten Ziellastgang  $\zeta^{(P)}$ . Der zweite Term hingegen bestraft ungültige Lösungen in Abhängigkeit vom individuellen anlagenspezifischen Abstand  $R_i$  vom Zentrum der Hyperkugel, so wie er durch die SVDD definiert ist. Ungültige Lösungen werden somit durch den zweiten Term in (12.6) quasi in Richtung des zulässigen Bereichs gezogen. Die beiden Parameter  $\omega_\delta$  und  $\omega_{r_{\mathcal{F}}}$  erlauben eine unterschiedliche Gewichtung des Einflusses von Zielfunktion und Gültigkeit der Lösung.

Ein exemplarisches Ergebnis für verhältnismäßig kleine Szenarien mit 10 Anlagen (nur BHKW) und Fahrplänen über 4, 8 bzw. 16 Zeitintervallen zeigt Tabelle 12.12. Bereits bei diesen kleinen Szenarien bleibt der Penalty-Ansatz offensichtlich zu oft in Minima mit ungültigen Lösungen stecken. Verwendet wurde das PSO-Verfahren aus der CILib-Bibliothek. Zwar kann durch hinreichend starke Gewichtung des Penaltywertes (das Gewicht für die Zielfunktion in der obersten Zeile ist nicht wie angegeben exakt 0, sondern 0.001) eine im Vergleich zu den Dekoderergebnissen konkurrenzfähige Realisierbarkeitsrate erreicht werden, jedoch ist die Lösungsgüte bei diesen Lösungen dann im Vergleich inakzeptabel schlecht.

Tabelle 12.12.: Ergebnisse für die Optimierung mittels Penalty-Ansatz für verschiedene Gewichtungsverhältnisse von  $\omega_\delta$  und  $\omega_{r_{\mathcal{F}}}$ . Angegeben ist jeweils die erreichte Lösungsgüte  $\delta$  und die Realisierbarkeitsrate der Lösung  $r_{\mathcal{F}}$ .

$\omega_\delta : \omega_{r_{\mathcal{F}}}$	Ziel	4	8	16
0.0 : 1.0	$\delta$	7.451 ± 3.137	57.843 ± 155.481	104.111 ± 136.893
	$r_{\mathcal{F}}$	0.940 ± 0.071	0.596 ± 0.146	0.116 ± 0.037
0.1 : 0.9	$\delta$	0.015 ± 0.007	0.000 ± 0.000	0.000 ± 0.000
	$r_{\mathcal{F}}$	0.616 ± 0.121	0.168 ± 0.122	0.008 ± 0.028
0.5 : 0.5	$\delta$	0.002 ± 0.007	0.000 ± 0.000	0.000 ± 0.000
	$r_{\mathcal{F}}$	0.460 ± 0.131	0.112 ± 0.083	0.008 ± 0.0029
0.9 : 0.1	$\delta$	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
	$r_{\mathcal{F}}$	0.376 ± 0.139	0.104 ± 0.084	0.016 ± 0.037

Dieser Vergleich zeigt, dass die alleinige Nutzung des Wissens um die Realisierbarkeit eines Lösungskandidaten mit Hilfe des Modells ohne systematische Lösungsgenerierung durch den Dekoder zumindest bei Verwendung eines Penaltyansatzes noch nicht zu akzeptablen Lösungen führt. Dies liegt weniger an den erreichten Lösungsgüten als vielmehr an der mangelnden Realisierbarkeit der so gefundenen Lösungen.

Zwar wurde im Rahmen dieser Arbeit nur der Penaltyansatz als weitere Constraint-Handhabungstechnik betrachtet, dennoch lässt sich bereits vermuten, dass die systematische Konstruktion von gültigen Lösungen solchen Verfahren überlegen ist, bei denen

auch ungültige Lösungen betrachtet und bewertet werden müssen. Ein Barriereansatz bei dem ausgehend von einer bereits gültigen Lösung nur Nachbarlösungen betrachtet werden, die ebenfalls gültig sind (der SVDD-Abstand, der den Rand des gültigen Bereichs markiert würde hier als Barriere dienen) wurde nicht betrachtet, da hierdurch nur solche Spezialfälle abgedeckt sind in denen der Suchraum aller Anlagen kompakt ist, also nicht in separierte Teilbereiche zerfällt (vgl. Abschnitt 6.3.2 auf Seite 80). Ansätze, die statt des SVDD-Modells, das Simulationsmodell direkt zur Entscheidung über die Realisierbarkeit eines Lösungskandidaten entscheiden, wurden ebenfalls nicht betrachtet, da die Simulationsmodelle so wie sie in dieser Arbeit Verwendung finden bzgl. möglicher Aussagen über die Realisierbarkeit weniger mächtig (nur ja oder nein) sind als die SVDD-Modelle. Spezielle Simulationsmodelle, die direkt in die Optimierung integriert werden, müssten anlagenspezifisch implementiert werden und würden dem Konzept der ad hoc Integration bei der automatisierten Ableitung des Optimierungsmodells widersprechen. Hierfür gibt es zahlreiche Beispiele in der Literatur. Sie erreichen zwar gute Ergebnisse für die Spezialfälle für die sie implementiert sind (manchmal auch ohne Einbezug von Constraints), weisen aber nicht die notwendige Flexibilität für eine dynamische Neuzusammenstellung auf (vgl. Kapitel 3 ab Seite 23). Ein direkter Vergleich ist hier daher auch nicht zulässig.

### 12.4.3. Die Rolle von Batteriespeichern

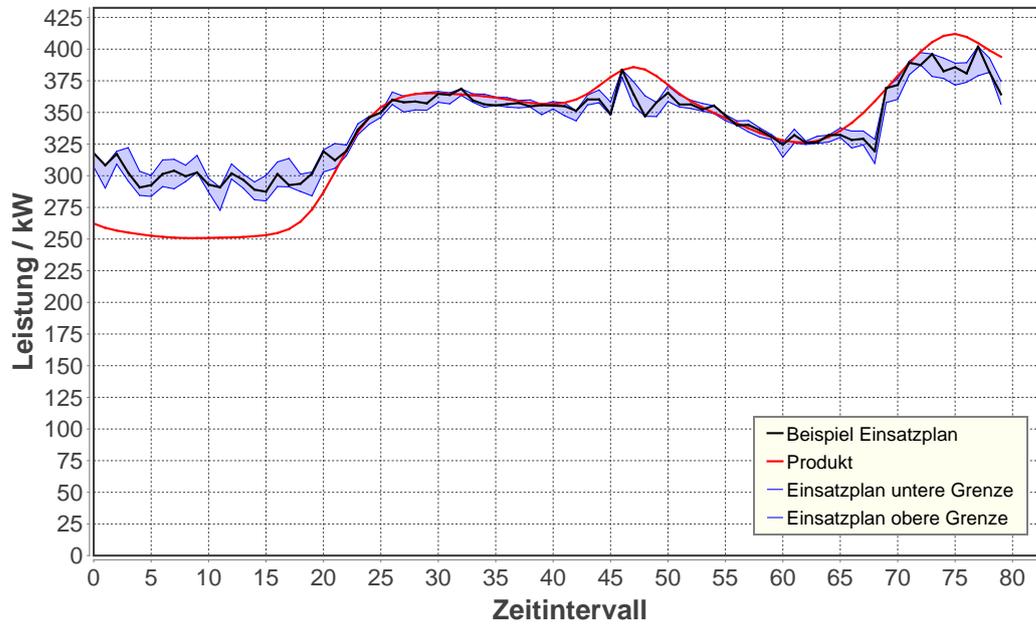
Ein interessanter Effekt konnte im Zusammenhang mit dem Einsatz von Batteriespeichern, die eigentlich zur Lösungsverbesserung gedacht waren, im Verbund beobachtet werden.

Abbildung 12.9 zeigt ein Beispiel. Zunächst wurde für einen ausschließlich aus BHKW bestehenden Verbund ein Einsatzplan erstellt, um ein aus einem Standardlastprofil abgeleitetes Produkt zu erzeugen. Der Verbund kann das Produkt in dieser Form nur teilweise erreichen. Die blaue Fläche in Abbildung 12.9(a) kennzeichnet den Bereich, in dem die Einsatzpläne mehrerer Planungsläufe liegen; ein Einsatzplan als Ergebnis ist exemplarisch hervorgehoben.

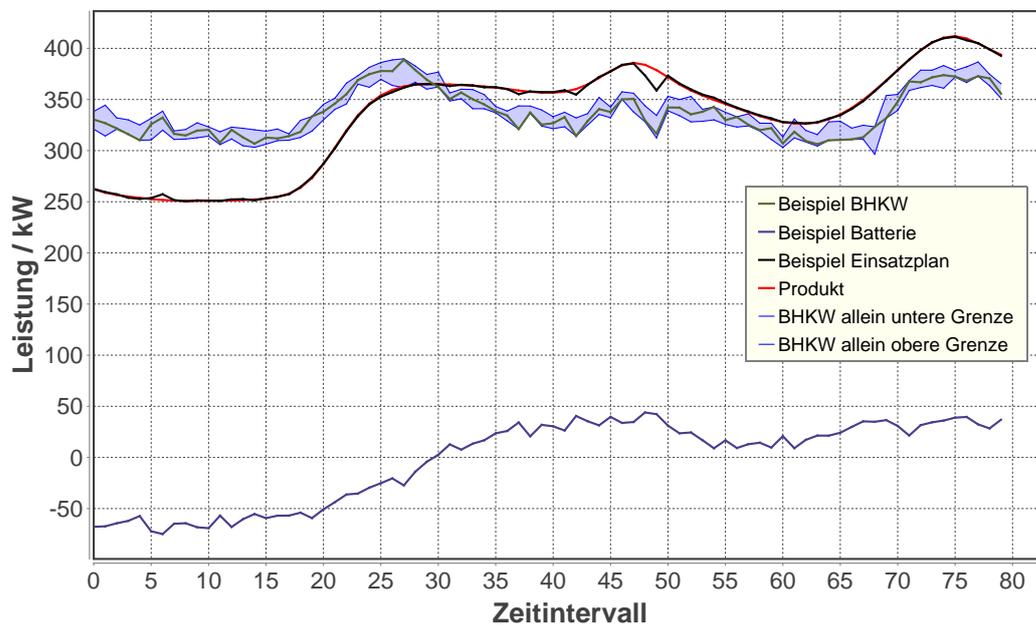
Kommt nun ein Batteriespeicher als zusätzliche Anlage zu dem Verbund hinzu, so kann das Ergebnis erwartungsgemäß dramatisch verbessert werden (schwarze Linie in 12.9(b)). Betrachtet man jedoch den Beitrag der BHKW zu diesem verbesserten Ergebnis in Abbildung 12.9(b), so sieht man deutlich, dass diese nun weiter von dem Produkt entfernt sind als vorher. Offenbar übernimmt der Batteriespeicher mehr als notwendig und kommt damit übermäßig stark zum Einsatz.

Der Fehler der BHKW (Abstand zwischen der Erzeugung nur der BHKW und dem gewünschten Produktfahrplan) allein betrachtet ist ohne Batterie kleiner. In einem Szenario mit 10 BHKW und 32 Zeitintervallen beispielsweise beträgt der Fehler durchschnittlich nur 52.07% im Vergleich zu den BHKW in einem Verbund mit Batteriespeicher. Natürlich sinkt der Gesamtfehler (BHKW und Batterie gemeinsam betrachtet) deutlich (im Beispiel auf etwa 2.25% des ursprünglichen Fehlers) ab. In einem Ganztageszenario mit 10 BHKW waren es 66.61%, bei 100 BHKW für 64 Zeitintervalle konnte eine Verschlechterung auf 71.55% festgestellt werden.

Das Hinzufügen eines Batteriespeichers zu dem Verbund verbessert zwar das Gesamtergebnis des Verbundes ggf. dramatisch, da der Batteriespeicher bei hinreichender



(a)



(b)

Abbildung 12.9.: Rolle der Batteriespeicher für die Einsatzplanoptimierung. Vergleich der Optimierungsergebnisse mit und ohne Speicher.

Dimensionierung quasi alle Unzulänglichkeiten des Verbundes ausgleichen kann, betrachtet man jedoch die aktiven Komponenten des Verbundes allein – die Erzeuger – so lässt sich hier eine Verschlechterung der Anpassung an den gewünschten Produktlastgang feststellen. Dieser Effekt lässt sich dadurch erklären, dass bei Vorhandensein eines hinreichend dimensionierten Energiespeichers dieser quasi alles ausgleichen kann, was die aktiven Anlagen falsch machen und somit (zumindest, wenn man nur die Annäherung an einen Ziellastgang betrachtet) jede Lösung gleich gut ist. Dies ist zunächst auch nicht ungewöhnlich. Wenn die Optimierung lediglich versucht, den Ziellastgang möglichst genau zu erreichen, so unterscheidet das Verfahren nicht die Sinnhaftigkeit des Einsatzes von BHKW oder Batteriespeicher. Hierfür fehlen in den hier betrachteten Szenarien geeignete, weitere Bewertungskriterien. Zudem vergrößert die Integration des Batteriespeichers in die Optimierung die Komplexität des Planungsproblems, da die Batterie aktiv als weitere Anlage mit verplant wird.

Insgesamt konnte beobachtet werden, dass sich bessere Resultate (reine Angleichung des Einsatzplans an den Ziellastgang) erzielen lassen, wenn ein Batteriespeicher separat betrachtet wird. In den Szenarien in dieser Arbeit werden daher zunächst keine Batteriespeicher in die Optimierung mit eingebunden. Optimiert wird immer nur der Pool an Anlagen (ohne Speicher) in einem Verbund. Der verbleibende Restlastgang kann üblicherweise von einem Batteriespeicher übernommen werden. Ergebnis der Optimierung ist ggf. die notwendige Mindestdimensionierung einer Batterie, die in der Lage ist den Restlastgang zu übernehmen. Sobald geeignete Kostenkennzahlen für die Mehrzieloptimierung bereitstehen, welche die Abnahme von BHKW-Strom gegenüber einer nicht notwendigen Einspeicherung bevorzugen bzw. in der Lage sind die Umwandlungsverluste durch die Ein- und Ausspeicherung einzubeziehen, sollte dieses Problem nicht mehr entstehen.

### 12.5. Zusammenfassung

In diesem ersten Teil der Evaluation des Gesamtkonzeptes, bei dem zunächst zentral arbeitende Standardverfahren für den Optimierungsteil bei der Wirkleistungsplanung zum Einsatz kamen, konnten die Umsetzbarkeit und die Wirksamkeit des Konzeptes gezeigt werden. Durch die Verwendung des Dekoders wird es tatsächlich möglich, das eigentlich constraintbehaftete Optimierungsproblem so in ein anderes Optimierungsproblem ohne (bis auf einfach zu handhabende Boxconstraints) Constraints zu überführen, dass dessen Lösung eine gültige Lösung des Ursprungsproblems darstellt. Erreicht wird dies durch Abbildung in einen anderen Suchraum.

Auf diese Weise wird es möglich, zur Lösung auch solche Heuristiken einzusetzen, die nicht explizit für den Einsatz in Constraint-basierten Problemen konzipiert sind. Als Vorteil kann auf erprobte Standardansätze zurückgegriffen werden.

Getestet wurde der Ansatz sowohl mit verschiedenen fertig verfügbaren Implementierungen von Standard-Heuristiken als auch mit einer eigenen Simulated Annealing Implementierung für den Einsatz mit Dekodern.

Ziel dieser Arbeit ist es nicht, ein spezielles Verfahren für die Optimierung zu entwickeln oder zu finden. Vielmehr soll die Voraussetzung geschaffen werden, verschiedene

Optimierungsverfahren einbinden zu können, ohne diese Verfahren speziell für das hier betrachtete Problem bezüglich der dynamischen Problemmodellierung oder der Integration und Behandlung individueller Constraints anpassen zu müssen. Auf diese Weise kann für die Planung auf anerkannte und erprobte Verfahren mit definierten Eigenschaften zurückgegriffen werden.

Dies konnte für zentrale Ansätze in diesem Kapitel bereits gezeigt werden. Es ist jedoch zu erwarten, dass in einem zukünftigen, verteilten System auch verteilte Verfahren zum Einsatz kommen werden. Mögliche Umsetzungsansätze für solche Szenarien werden im nächsten Kapitel behandelt.



# 13. Umsetzung mit dezentralen Optimierungsansätzen

*A basketball team is like the five fingers on your hand.  
If you can get them all together, you have a fist.  
That's how I want you to play.*

---

*(Michael William Krzyzewski)*

## 13.1. Einleitung

Im vorigen Kapitel konnte bereits gezeigt werden, dass die Integration des Suchraummodells mit zentralen Verfahren zur heuristischen Optimierung ein Verfahren liefert, welches in der Lage ist das Wirkleistungsplanungsproblem mit seinen Randbedingungen zu lösen.

Diesem Vorgehen sind allerdings Grenzen durch die Problemgröße gesetzt. Darüber hinaus lässt sich vermuten, dass ein Selbstorganisationsansatz die besten Voraussetzungen mitbringt, um dem dynamischen Charakter der Problemstellung am besten gerecht zu werden, insbesondere bei Einbezug dynamischer Adaption einer großen Anzahl dezentraler Anlagen an jeweils aktuelle Netzerfordernisse oder der selbsttätigen Adaption an Planänderungen durch Anlagenausfall o. ä. (Nieße et al. [2012b](#)). Die Wirkleistungsplanung ist in diesem Kontext nur ein Baustein von vielen; im Falle eines selbstorganisierten Gesamtsystems muss jedoch auch diese dezentral funktionieren können. Daher wird in diesem Kapitel noch die Einbettung in dezentrale Ansätze betrachtet und evaluiert. Dass eine solche Integration möglich und insbesondere auch sinnvoll ist wurde bereits in (Hinrichs et al. [2013b](#)) gezeigt. Neben der Reflektion dieses Integrationsansatzes befasst sich dieses Kapitel im Wesentlichen mit einem Greedy-basierten Ansatz aus (Bremer & Sonnenschein [2012a](#)) und dessen Mehrzielerweiterung sowie der Integration in die vollständig dezentral arbeitende Heuristik COHDA (Hinrichs et al. [2013b](#)).

## 13.2. Verfahren

Im Rahmen dieser Arbeit sollen zwei dezentrale Optimierungsverfahren zur Lösung des Wirkleistungsplanungsproblems betrachtet werden. Der erste Ansatz stellt eine relativ einfache Heuristik (Bremer & Sonnenschein [2012a](#)) dar, die allerdings speziell für den Einsatz mit dem Supportvektor Dekoder entwickelt wurde. Der zweite Ansatz (Hinrichs et al. [2014](#)) wurde zunächst mit diskreten Suchräumen für die lokalen Entscheidungen

betrieben und später auch mit dem hier entwickelten Supportvektor Dekoder kombiniert (Hinrichs et al. 2013b).

#### 13.2.1. Ein verteilter Mehrziel-Greedy-Ansatz

Das verteilte Greedy-Verfahren so wie es in (Bremer & Sonnenschein 2012a) vorgestellt wurde, basiert in erster Linie auf der Idee, dass jeder Agent in einem verteilten System lokal entscheidet, was er im besten Fall beitragen kann, um die Gruppe als Ganzes möglichst weit voranzubringen. Im Rahmen dieses Algorithmus gibt es einen Typ von Agent, welcher genau eine Einheit vertritt. D.h. der Agent kennt ein Simulationsmodell der zugehörigen Anlage und ist in der Lage ein Suchraummodell mitsamt Dekoder zu generieren. Dieses bleibt als lokales Wissen beim jeweiligen Agenten und braucht nicht nach außen gegeben zu werden. Es dient lediglich zur systematischen Lösungsreparatur in dem folgenden Prozess, der wiederholt durchgeführt wird: Jeder Agent kennt die aktuelle Lösung des Verbundes. Er kennt allerdings nicht den Einsatzplan, sondern lediglich den Summenfahrplan aller Anlagen. Er kennt zudem seinen eigenen Fahrplan als aktuellen Beitrag zur Lösung; initial startet jeder Agent mit einem zufällig gewählten, jedoch realisierbaren Fahrplan. Es handelt sich somit um einen so genannten Blackboard-Ansatz (Corkill 1991). Die Agenten haben also die Möglichkeit auf einen zentralen Wissensspeicher zuzugreifen. Aus dem Summenfahrplan aller Anlagen und dem eigenen Fahrplan kann jeder Agent durch Differenzbildung einfach ermitteln, was alle anderen Agenten (gemeinsam) gerade planen. Da der Ziellastgang bekannt ist, kann nun einfach derjenige Fahrplan bestimmt werden der (wenn von der eigenen Anlage so umgesetzt) genau den Ziellastgang ergibt. Vermutlich ist dieser Fahrplan von der eigenen Anlage jedoch nicht exakt umsetzbar. Daher wird dieser vom eigenen Dekoder auf einen möglichst ähnlichen, aber realisierbaren Fahrplan abgebildet. Dieser Fahrplan stellt den aktuell bestmöglichen Beitrag eines Agenten zur Gesamtlösung dar.

Reihum in einem wiederholten Prozess bestimmen nun die Agenten jeweils den Betrag der anderen und die bestmögliche eigene Ergänzung zur Erlangung des Ziellastgangs. nun kann hieraus eine aktualisierte Gesamtlösung berechnet werden, die den anderen Agenten zugänglich gemacht wird. Im Anschluss ist ein anderer Agent an der Reihe, den gleichen Aktualisierungsschritt lokal bei sich auszuführen. Dies kann in festgelegter Reihenfolge geschehen oder auch in zufälliger Reihenfolge. Zudem wurden Tests durchgeführt, diesen Prozess asynchron auszuführen. Eine asynchrone Ausführung führt ebenfalls zur Konvergenz der Lösung, wenn auch etwas langsamer da Agenten ggf. auch widersprüchliche Updates nebenläufig vornehmen.

Der Pseudocode in Abbildung 13.1 auf der nächsten Seite aus (Bremer & Sonnenschein 2012a) zeigt den Ablauf in formalisierter Form. Eine graphische Darstellung der Situation in zwei aufeinander folgenden Schritten ist in Abbildung 13.2 auf der nächsten Seite gegeben. Im ersten Schritt 13.2(a) ist aktuell Agent 3 an der Reihe.  $x_{Rest}$  kennzeichnet die (vektorielle) Summe der Fahrpläne der anderen Agenten und Agent 3 bestimmt den Differenzvektor  $\Delta x$  welcher notwendig wäre als eigener Beitrag, um den Ziellastgang exakt zu erreichen. Um einen für die Einheit 3 realisierbaren Fahrplan zu erhalten wird  $\Delta x$  mit Hilfe des Dekoders  $\gamma_3$  abgebildet auf einen Fahrplan aus dem Suchraum  $\mathcal{F}_3$  von

```

 $\mathcal{A} \leftarrow$  List of all agents
 $s_{target} \leftarrow$  target schedule
 $S_a^t \leftarrow$  This agents contribution (initially zero)
if is initiator then
   $S \leftarrow zeros(n, d)$ 
else
   $S \leftarrow$  aggregated schedule
   $S_a^{t+1} \leftarrow \gamma(s_{target} - (S - S_a))$ 
   $S \leftarrow S - S_a^t + S_a^{t+1}$ 
  if no stop criterion met then
    choose random agent  $A \in \mathcal{A}$ 
    send message with  $S$  to  $A$ 
  else
    publish solution  $S$ 
  end if
end if

```

Abbildung 13.1.: Aktualisierungsschritt, den jeder Agent im Laufe des Greedy-Verfahrens wiederholt durchführt, um die Lösung zu verbessern (aus Bremer & Sonnenschein 2012a).

Einheit 3. Wenn im nächsten Schritt nun Agent 2 am Zug ist, so stellt sich der Schritt aus Sicht von Agent 2 analog wie in Abbildung 13.2(b) dar.

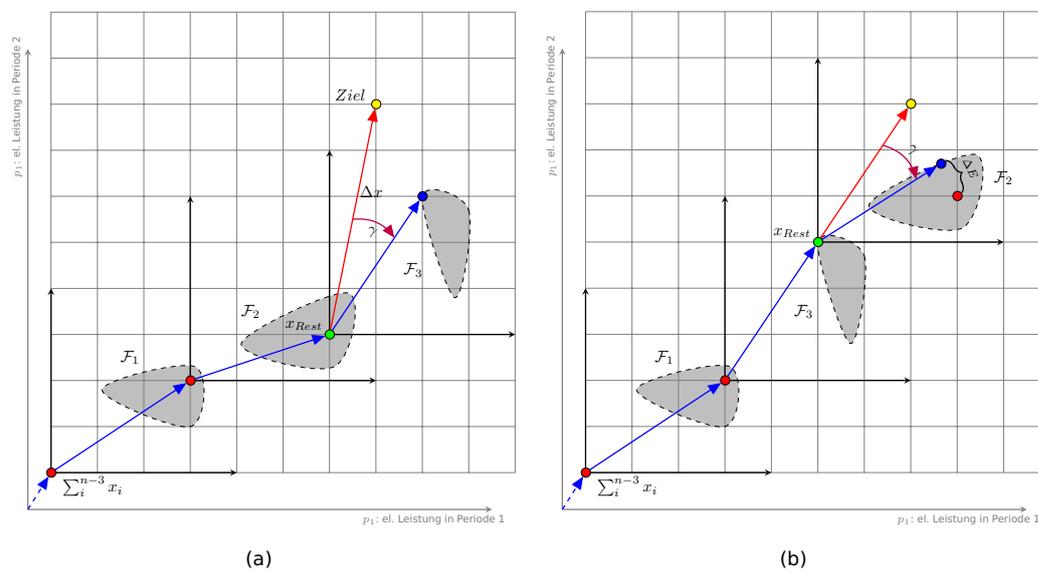


Abbildung 13.2.: Grundprinzip des Greedyverfahrens anhand zweier Iterationsschritte nach (Bremer & Sonnenschein 2012a).

## Ergebnisse

Tabellen 13.1 bis 13.3 zeigen exemplarisch einige Ergebnisse, die mit diesem Verfahren erzielt wurden. Insgesamt wurden auch bei den verteilten Verfahren die in Abschnitt 11.3.1 auf Seite 153 diskutierten Szenarien zugrunde gelegt. Für die Beurteilung der Ergebnisse wurden auch hier wieder die schon aus dem vorigen Kapitel bekannten Kennzahlen bestimmt. Das Verfahren selbst wurde bezüglich Konvergenzverhalten, Kommunikationsaufwand und ähnlichem nicht weiter untersucht als in (Bremer & Sonnenschein 2012a). Die entscheidende Erkenntnis ist zunächst, dass sich der Supportvektor Dekoder auf einfache Art und Weise auch für verteilte Verfahren verwenden lässt. Private Daten über Suchraum, technische Einschränkungen oder Kostenfaktoren brauchen in diesem Fall nicht zwingend (auch nicht als Dekoderabstraktion) aus der Hand gegeben werden. Dennoch ist der Dekoder während der Optimierung für lokale Entscheidungen ein unverzichtbares Mittel zur systematischen Lösungsreparatur. Eine solche systematische Konstruktion gültiger Lösungen als benachbarte Näherung zur optimalen (im Sinne von Annäherung an die Zielvorgabe) Lösung wäre mit einem mathematischen Modell der Anlage allein in dieser Form schwerlich zu bewältigen. Der Dekoder bietet so auch in der verteilten Optimierung den einzelnen Agenten eine Unterstützung bei der lokalen Entscheidungsfindung und somit einen entscheidenden Mehrwert.

Tabelle 13.1.: Ergebnisse für den verteilten Greedyalgorithmus für Verbünde mit 10 Anlagen.

Kennzahl	RND/ BHKW	H0/ BHKW	RND/ Mixed	H0/ Mixed
$\delta$	0.59366 ± 0.26855	0.57954 ± 0.25748	0.38467 ± 0.22102	0.37184 ± 0.18393
Iterationen	363.162 ± 218.423	338.834 ± 174.596	314.816 ± 148.714	313.529 ± 195.941
$n_{50\%}$	4.380 ± 2.238	4.368 ± 2.133	4.822 ± 2.389	4.179 ± 2.368
$n_{10\%}$	18.853 ± 12.845	16.955 ± 9.873	15.952 ± 8.702	15.799 ± 9.497
$n_{5\%}$	30.441 ± 19.176	28.154 ± 18.516	26.371 ± 15.668	24.472 ± 14.601
abs. $\Delta E$ / kWh	0.34543 ± 0.15659	0.33871 ± 0.15484	0.22258 ± 0.12912	0.21712 ± 0.11139
rel. $\Delta E$	0.00667 ± 0.00307	0.00676 ± 0.00313	0.00807 ± 0.00487	0.00731 ± 0.00385
$\overline{\Delta P}$ / kW	0.17272 ± 0.07830	0.16936 ± 0.07742	0.11129 ± 0.06456	0.10856 ± 0.05570

Tabelle 13.2.: Ergebnisse für den verteilten Greedyalgorithmus für Verbünde mit 100 Anlagen.

Kennzahl	RND/ BHKW	H0/ BHKW	RND/ Mixed	H0/ Mixed
$\delta$	0.05726 ± 0.02339	0.06586 ± 0.02821	0.03832 ± 0.01502	0.04041 ± 0.01980
Iterationen	3151.328 ± 741.275	3172.012 ± 622.753	3038.088 ± 570.636	3186.694 ± 1047.498
$n_{50\%}$	13.000 ± 4.577	30.088 ± 14.375	14.781 ± 5.799	21.000 ± 9.785
$n_{10\%}$	25.306 ± 7.530	66.668 ± 33.472	27.860 ± 8.974	42.265 ± 19.788
$n_{5\%}$	29.484 ± 7.580	73.738 ± 37.069	31.762 ± 9.207	46.327 ± 20.762
abs. $\Delta E$ / kWh	0.03326 ± 0.01390	0.03847 ± 0.01700	0.02246 ± 0.00925	0.02413 ± 0.01212
rel. $\Delta E$	0.00007 ± 0.00003	0.00008 ± 0.00003	0.00007 ± 0.00003	0.00009 ± 0.00004
$\overline{\Delta P}$ / kW	0.01663 ± 0.00695	0.01924 ± 0.00850	0.01123 ± 0.00462	0.01206 ± 0.00606

Getestet wurde mit Verbänden bis 1000 Anlagen, wobei die Simulation allerdings synchron auf einem einzelnen Prozessor ausgeführt wurde. Dies stellt aber keine Einschränkung für die Umsetzbarkeit als wirklich verteilt laufendes System dar. Eine verteilte Umsetzung mit asynchroner Ausführung wurde bereits in (Bremer & Sonnenschein 2013g) evaluiert.

Tabelle 13.3.: Ergebnisse für den verteilten Greedyalgorithmus für Verbünde mit 1000 Anlagen.

Kennzahl	RND/ BHKW	H0/ BHKW	RND/ Mixed	H0/ Mixed
$\delta$	0.00382 ± 0.00181	0.00565 ± 0.00312	0.00416 ± 0.00207	0.00580 ± 0.00308
Iterationen	30746.680 ± 3102.523	32883.720 ± 3917.434	30795.588 ± 3494.363	31806.452 ± 3522.818
$n_{50\%}$	47.940 ± 15.777	435.700 ± 205.244	52.845 ± 16.358	294.731 ± 149.111
$n_{10\%}$	98.650 ± 30.696	1144.870 ± 644.263	98.371 ± 28.258	678.140 ± 377.007
$n_{5\%}$	106.200 ± 32.297	1273.640 ± 731.976	104.351 ± 29.459	734.269 ± 412.488
abs. $\Delta E$ / kWh	0.00219 ± 0.00108	0.00327 ± 0.00189	0.00242 ± 0.00123	0.00345 ± 0.00192
rel. $\Delta E$	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000
$\overline{\Delta P}$ / kW	0.00110 ± 0.00054	0.00164 ± 0.00095	0.00121 ± 0.00062	0.00172 ± 0.00096

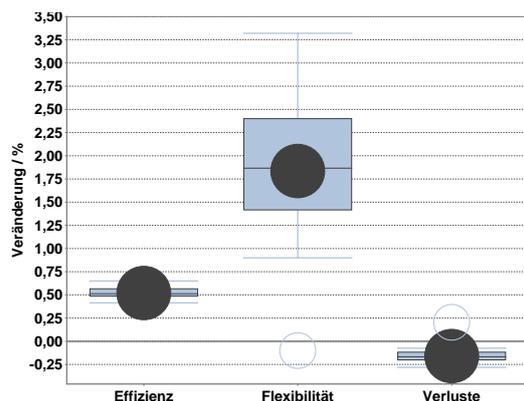


Abbildung 13.3: Beispielresultat für Mehrzieloptimierung mit drei zusätzlichen Optimierungskriterien. Optimierte wurde ein Erzeugerverbund mit 25 Anlagen für einen Zeitraum von 8 15-minütigen Intervallen. Die gleichzeitig einhergehende Verschlechterung der erzielten Wirkleistungsannäherung betrug durchschnittlich  $0.359 \pm 0.197$  kW, also im Mittel etwa 14 Watt je Anlage und Zeitintervall (in der Grafik nicht mit abgebildet).

### Erweiterung für Mehrzieloptimierung

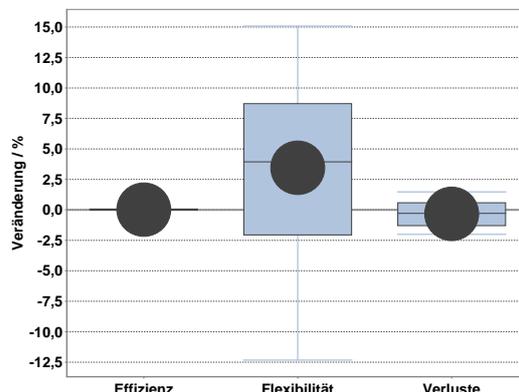
An dieser Stelle soll noch die Umsetzung eines Mehrzieloptimierungsansatzes mit Hilfe des Supportvektor Dekoders betrachtet werden. Das Grundprinzip ist auf jeden der hier betrachteten Algorithmen anwendbar und wird hier lediglich exemplarisch am Beispiel des Greedy-Verfahrens betrachtet.

Wie bereits in Kapitel 10 ab Seite 133 erläutert, werden Kennzahlen für die differenzierte Bewertung im Rahmen einer Mehrzieloptimierung integriert durch Erzeugen von zusammengesetzten Featurevektoren bestehend aus Leistungswerten  $p_i$  und Kennzahlenwerten  $I_j$   $\mathbf{x} = (p_1, \dots, p_d, I_1, \dots, I_\ell)$ . Analog kann nun auch der Vektor für das Optimierungsziel als solcher Featurevektor aufgebaut werden, sodass dieser ebenfalls Werte für die in jedem Zeitintervall zu erzielenden Wirkleistungswerte und gewünschte Kennzahlenwerte enthält. Da auch die Kennzahlenwerte auf das Intervall  $[0, 1]$  skaliert sind, ist die Wahl für die entsprechenden Zielvorgaben recht einfach: für alle zu minimierenden Ziele wird eine 0 vorgegebenen, für alle zu maximierenden Ziel  $n$ . Da  $n$  die Anzahl der Anlagen ist, ist dies die größtmögliche Summe die die Werte aller Anlagen haben können. Eine individuelle Gewichtung einzelner Optimierungsziele erfolgt hier nicht.

Mit einem derartigen Zielvektor können nun alle Verfahren welche in der Einzelvariante als Zielfunktion lediglich den Abstand  $\delta$  zwischen Einsatzplan und Ziellastgang minimiert haben unverändert verwendet werden. Minimiert wird dann gleichzeitig der Abstand zwischen den Leistungswerten und ihrer Zielvorgaben und den jeweiligen Kennzahlen und ihrer Zielvorgabe. Jedes herkömmliche Mehrzielverfahren könnte analog durch separate Weiterverarbeitung und Auswertung der Kennzahlenwerte verwendet werden. Hier soll jedoch die direkte Verwendung getestet werden.

### 13. Umsetzung mit dezentralen Optimierungsansätzen

Abbildung 13.4: Beispielergebnis für Mehrzieloptimierung mit drei zusätzlichen Optimierungskriterien, diesmal für längere Zeiträume und weniger Anlagen (10 BHKW über 96 Zeitintervalle für einen ganzen Tag). Die gleichzeitig einhergehende Verschlechterung der erzielten Wirkleistungsannäherung (nicht mit dargestellt) betrug durchschnittlich  $1.491 \pm 0.339$  kW (bei 47 kW Nennleistung).



Die Abbildungen 13.3 und 13.4 zeigen exemplarisch zwei Ergebnisse, die mit den in Abschnitt 10.3 auf Seite 135 beschriebenen Kennzahlen für die zu maximierende Anlageneffizienz, die zu maximierende verbleibende Flexibilität und die zu minimierenden Speicherverluste erzielt wurden. Hierzu wurde dieselbe Wirkleistungsplanung einmal als Einzelvariante (nur Beachtung der Leistungswerte) und einmal als Mehrzielvariante (mit kombinierten Wirkleistungs-Kennzahlen-Vektoren) gelöst. Dargestellt ist nur der Vergleich der Werte, die für die drei zusätzlichen Optimierungsziele erreicht wurden. Bei reiner Beachtung der Wirkleistung, erfolgt die Auswahl der Fahrpläne in Bezug auf die Bewertung hinsichtlich der zusätzlichen Optimierungsziele zufällig, im Mehrziel Fall werden die Fahrpläne diesbezüglich gezielt gewählt. Die beiden Abbildungen zeigen die Veränderungen in Prozent.

Betrachtet wurden reine Erzeugerverbünde mit erreichbarer, zufälliger Wirkleistungsvorgabe. Insgesamt wurden jeweils über 100 Durchläufe gemittelt. Die BHKW wurden jeweils gleichverteilt zufällig Parametrisiert mit verschiedenen Startspeichertemperaturen, unterschiedlichen Außentemperaturen, individuellem thermischen Bedarf und unterschiedlichen zulässigen Leistungsgradienten (vgl. Abschnitt 11.3.2 auf Seite 154).

Dargestellt sind für die drei zusätzlichen Kennzahlen jeweils die prozentuale Änderung gegenüber den zufälligen Werten (bei Nichtbeachtung), wenn der Mehrzielansatz verwendet wurde. Man sieht deutlich, dass die beiden zu maximierenden Werte eine positive Veränderung gegenüber dem Einzelfall haben. Die zu minimierenden Verluste fallen bei Beachtung im Mehrzielansatz geringer aus. Die absolut erreichten Verbesserungen fallen eher gering aus. Dies liegt vornehmlich aber an den verwendeten Szenarien und Kennzahlen. So ist beispielsweise der betrachtete Zeitraum in den Szenarien in Ergebnis 13.3 zu gering als dass die BHKWs eine Chance hätten große Abweichungen im Speicherfüllstand (als Maß für die verbleibende Anlagenflexibilität) zu erzielen. Bei den betrachteten Ganztagesfahrplänen in 13.4 fällt die Veränderung entsprechend größer aus. Die Unterschiede in der Effizienz verschiedener Leistungsstufen der BHKW sind ebenfalls sehr gering (bei den betrachteten Typen etwa 2 Prozentpunkte zwischen Minimum und Maximum), bei längeren Szenarien mitteln sich die Unterschiede über den Tag offenbar fast vollständig heraus, sodass hier kaum Verbesserungen zu erzielen sind. Auch die Verluste sind bei modernen Anlagen so gering, dass quantitativ kaum etwas zu optimieren ist.

Zu beachten ist bei diesen Ergebnissen also weniger die Größenordnung der erzielten Ergebnisse als vielmehr die Richtung der Verbesserungen die durch gleichzeitige Opti-

mierung aller Ziele erreicht werden konnten. Es konnte gezeigt werden, dass die in das Suchraummodell integrierten Kennzahlen durch den Dekoder für eine Mehrzieloptimierung verwendet werden können und sich der in dieser Arbeit vorgestellte Ansatz generell eignet um das Wirkleistungsplanungsproblem (6.14) wie es eingangs in der Problemdefinition aufgeworfen wurde vollständig zu lösen. Die Erarbeitung sinnvoll zu integrierender Kennzahlen wird jedoch noch Aufgabe nachfolgender Arbeiten sein müssen.

Gleiches gilt für die Einbindung in weitere Mehrzielansätze. Hier wäre denkbar, die verschiedenen Optimierungsziele gewichtet zu einer einzigen Zielgröße zusammenzufassen, um herkömmliche Einzelverfahren unverändert verwenden zu können. Andererseits könnten auch Mehrzielverfahren zum Einsatz kommen, die statt eines einzelnen Ergebnisses eine Paretofront (Menge Pareto-optimaler Ergebnisse) bestimmen (beispielsweise Beume et al. 2008), aus der in einem zweiten Schritt die je nach Situation günstigste Lösung ausgewählt werden kann. Die Erarbeitung geeigneter Gewichtungen sowie gegebenenfalls geeigneter Entscheidungsregeln für die Wahl einer Lösung bedarf allerdings ebenfalls der vorherigen Entwicklung des konkret zu integrierenden Satzes an Kennzahlen, was aber nicht Ziel dieser Arbeit ist.

### 13.2.2. COHDA

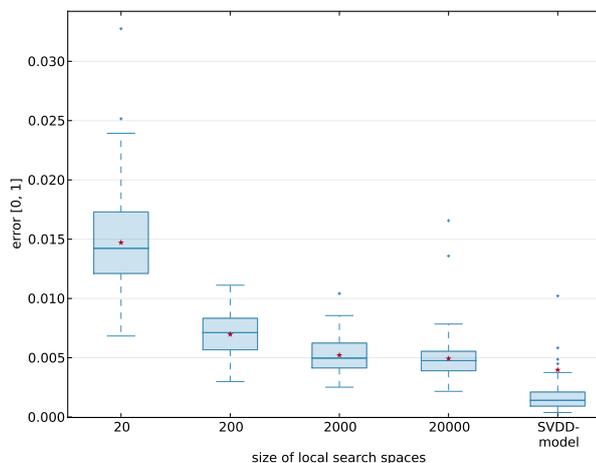
Als Beispiel für den Einsatz des Dekoderansatzes im Rahmen eines asynchron und verteilt arbeitenden Optimierungsansatzes soll hier abschließend die Integration in das Verfahren **C**ombinatorial **O**ptimization **H**euristic for **D**istributed **A**gents (COHDA) betrachtet werden. COHDA (Hinrichs 2014; Hinrichs et al. 2013a; Hinrichs et al. 2013c) wurde von Christian Hinrichs als verteilte Lösung des Wirkleistungsplanungsproblems (in einer verteilten Constraint-Optimierungs-Formulierung) für ein Multiagentensystem entwickelt. Hierbei ist es die Aufgabe eines jeden Agenten, einen Fahrplan für die von dem Agenten verantwortete Anlage zu finden, sodass als Softconstraint die Abweichung des Summenfahrplans von einem vorgegebenen Zielfahrplan möglichst gering wird.

In der Heuristik COHDA beschreibt jeder Agent nicht eine vollständige Lösung (wie häufig in populationsbasierten Verfahren) sondern jeder Agent repräsentiert eine Klasse in einem Multiple-Choice-Kombinatorikproblem (Lust & Teghem 2010). Im Rahmen der Wirkleistungsplanung entspricht jede Klasse dem Lösungsraum der Anlage im Verantwortungsbereich des Agenten. Ein Agent steht in COHDA immer mit einer begrenzten Anzahl an Nachbaragenten in Kontakt, sodass ein Zusammenführen der Lösungsräume möglich wird.

Jeder Agent in COHDA verfügt über sein eigenes (ggf. unvollständiges) Wissen über die von anderen Agenten gewählten Fahrpläne und kann basierend hierauf aus seinem eigenen Lösungsraum den am besten hierzu passenden Fahrplan wählen. Da diese Wahl (bzw. die Änderung einer veralteten vorherigen Auswahl) an die Nachbarschaft des Agenten kommuniziert wird verbreitet sich das Wissen um die gewählten Fahrpläne nach und nach im ganzen Verbund. Hierzu ist es nicht notwendig, dass die Agenten vollständig vernetzt sind, solange das Kommunikationsnetz zusammenhängend ist. In der Praxis hat sich eine Small World Vernetzung (Watts & Strogatz 1998) als sinnvoll erwiesen. Da alle Fahrplanauswahlinformationen mit einem Alter gelabelt sind, kann bei Kommunikation auch über mehrere Nachrichten-Hops hinweg entschieden werden, ob das eigene Wissen

### 13. Umsetzung mit dezentralen Optimierungsansätzen

Abbildung 13.5: Beispielergebnis aus der Integration von Supportvektor-Dekoder und COHDA für ein Szenario mit einem Erzeugerverbund aus 30 BHKW und einem 8-dimensionalen Ziellastgang. Durch Verwendung des vollständigen Lösungsraums mittels Dekoder lässt sich die erreichte Lösungsqualität signifikant verbessern. Der Mittelwert bei Verwendung des SVDD-Modells wird durch einen einzelnen (nicht dargestellten) Ausreißer etwas negativ beeinflusst. Darstellung aus (Hinrichs et al. 2013b).



über die aktuelle Gesamtauswahl angepasst werden muss, was ein Überdenken der eigenen Fahrplanwahl und ggf. deren Kommunikation nach sich ziehen würde. Nach und nach verfügen alle Agenten über vollständiges Wissen und wenn keiner der Agenten mehr in der Lage ist eine Auswahl zu treffen die die Gesamtlösung verbessert, konvergiert das Verfahren.

In einem verteilten kombinatorischen Optimierungsproblem wie der Wirkleistungsplanung treten Constraints in der Regel auf mehreren Ebenen auf: innerhalb eines einzelnen Agenten, zwischen zwei (oder wenigen) Agenten oder auf globaler Ebene, wenn eine Vielzahl von Agenten ein Constraint gemeinsam erfüllen soll. Die Heuristik COHDA ist darauf ausgelegt, vordringlich globale Constraints zu handhaben, welche im Bereich der verteilten Constraint-basierten Optimierung eine besondere Schwierigkeit darstellen, da sie in der Regel eine große Zahl von (oft auch alle) Agenten betreffen und somit zu einer Explosion des Constraintgraphen (mit den Agenten als Knoten und den Constraints als Kante) und damit des Kommunikationsaufwandes führen (Chapman et al. 2011). agentenintern arbeitet COHDA auf einem unbeschränkten Lösungsraum als Voraussetzung, sodass sich hier eine Integrationsmöglichkeit für das Suchraummodells bietet. Hierdurch wird es möglich, die ursprünglich integrierte Variante mit begrenzten Mengen von durch Simulationsmodellen vorab generierten Beispielfahrplänen abzulösen durch ein Modell des unbeschränkten Lösungsraums. Wie in Abbildung 13.5 dargestellt, lassen sich durch die vollständigere und somit größere Auswahl an Lösungsalternativen auch bessere Lösungsgüten erreichen (Hinrichs et al. 2013b). Das betrachtete Szenario umfasst reine Erzeugerverbünde mit 30 BHKW und zufällig generierten Ziellastgängen über 2 Stunden (8 Zeitintervalle je 15 Minuten). Aufgeführt sind die mit COHDA erreichten Ergebnisse als kumulierte absolute Abweichung des Einsatzplans vom Ziellastgang – normiert auf  $[0, 1]$  für bessere Vergleichbarkeit. Verglichen wird der Einsatz des Supportvektor-Dekoders für die lokale Wahl des optimalen Fahrplans eines Agenten mit einem älteren Ansatz, bei dem jeder Agent aus einer Menge vorab bestimmter gültiger Fahrpläne (vergleichbar der Trainingsmenge) wählt, die mit dem Sampling aus Abschnitt 8.2 auf Seite 102 erstellt wurde. Zahlreiche weitere Beispiele finden sich in der Dissertation von Herrn Hinrichs (Hinrichs 2014). Eine vollständige Beschreibung des Integrationsansatzes findet sich auch in (Hinrichs et al. 2013b).

### 13.3. Zusammenfassung

Mit der erfolgreichen Integration auch in dezentrale und verteilte Verfahren konnte nunmehr gezeigt werden, dass der vorgeschlagene Ansatz zur dynamischen Integration von individuell durch Constraints eingeschränkten Suchräumen für die Lösung des Wirkleistungsplanungsproblems geeignet ist.

Die reine Wirkleistungsplanung ist aber nur eine von mehreren Aufgaben bei der Koordination im Smart Grid für die die Erstellung von Suchraummodellen bzw. die Erstellung eines Dekoders für das Generieren zulässiger Fahrpläne hilfreich sein können.

In den nachfolgenden Kapiteln soll noch zunächst eine Erweiterung des Modells für den Einbezug von prognoseunsicheren Trainingsdaten in die Erstellung des Modells betrachtet werden, bevor in Kapitel 15 ab Seite 217 als spezieller Anwendungsfall für den Dekoder die verteilte Bestimmung eines Schlüssels für die faire Gewinnverteilung diskutiert wird. Diese beiden Fälle stellen Erweiterungen des Supportvektor Dekoder Ansatzes dar und sind noch nicht erschöpfend behandelt und werden noch weiter untersucht. An dieser Stelle sollen hier dennoch zunächst bereits die Problemstellung angerissen sowie Wege für mögliche Lösungsansätze unter Verwendung des Supportvektor Dekoders aufgezeigt werden.



# 14. Integration von Prognoseunsicherheiten für eine robuste Planung

*Prediction is very difficult,  
especially about the future.*

---

*(Niels Henrik David Bohr)*

## 14.1. Einleitung

Ein nicht unerheblicher Faktor für die Güte des Suchraummodells sind die Unsicherheiten der Prognosen, die in zahlreichen Modellen in die Erstellung der Beispielfahrpläne eingehen. Als Beispiel seien an dieser Stelle die Prognosen für den zukünftigen thermischen Bedarf zu nennen, der beispielsweise von einem BHKW gedeckt werden soll. Hier gehen als Unsicherheit sowohl der vorhergesagte Verlauf der Außentemperatur ein als auch eine Abschätzung über Höhe und Zeitpunkt diverser Warmwasserzapfungen. Für andere Anlagen gelten ähnliche, insbesondere auch durch menschliche Interaktion hervorgerufene Unsicherheiten. Wenn es gelingt, derartige Unwägbarkeiten zumindest teilweise im Suchraummodell zu integrieren, kann eine Verbesserung bei der Robustheit der durch die Optimierung erarbeiteten Einsatzpläne bezüglich ihrer Umsetzbarkeit trotz geänderter Rahmenbedingungen erwartet werden. Daher wird in diesem Abschnitt ein Vorschlag erarbeitet, wie Informationen über die Zuverlässigkeit einzelner Fahrpläne zielführend mit in das Modell integriert werden können.

## 14.2. Erweiterung des Modells

Für die Handhabung von Datenunsicherheiten in der SVDD sind bisher nur wenige Ansätze vorgeschlagen worden. Für das Anwendungsfeld der Datenclustering haben Zheng et al. (2006) beispielsweise für den Penaltyterm, der das Verhältnis aus Hyperkugelvolumen und Ausreißerakzeptanz steuert, eine Fuzzy-Zugehörigkeitsfunktion definiert, welche die Zugehörigkeit jedes Punktes aus der Trainingsmenge unscharf formuliert. Einen anderen Fuzzy-Ansatz haben GhasemiGol et al. (2010) gewählt. Hier werden die Constraints für die Zugehörigkeit zur Kugel unscharf gewählt. In (Liu et al. 2013) wird ein Konfidenzwert

verwendet. Statt eines fest vorgegebenen Penaltyterms wie in Gleichung 5.14 auf Seite 61 wird der Penaltyterm erweitert zu

$$L(\xi, \mu, \beta) = R^2 - \sum_i (R^2 + \xi_i - \|\Phi(x_i) - a\|^2) \beta_i - \sum_i \xi_i \mu_i + C \sum_i (\kappa[x_i] \xi_i). \quad (14.1)$$

Dieser Penaltyterm (in Gleichung 14.1 fett hervorgehoben) aber auch bereits dessen einfache, originale Form in Gleichung 5.14 der originalen SVDD repräsentieren über den Parameter  $C$  einen Tradeoff zwischen Volumen der Hyperkugel und akzeptiertem Fehler. In der erweiterten Version repräsentiert  $\kappa[x_i] \xi_i$  eine gewichtete Größe für den akzeptierten Fehler (Liu et al. 2013). Der Funktion  $\kappa : \mathbb{R}^d \rightarrow [0, 1]$  ordnet jedem Trainingsvektor einen Konfidenzwert zu. Dieser wird in (Liu et al. 2013) bestimmt, indem das SVDD-Verfahren zweimal durchgeführt wird. Einmal in der klassischen Variante mit konstanter Gewichtung aller Trainingsvektoren. Im Ergebnis liegt dann für den zweiten Durchlauf der Abstand jedes Trainingspunktes vom Zentrum der Hyperkugel vor, der nunmehr als individuelle Gewichtung (skaliert mit dem Maximalabstand) verwendet wird.

All diesen Methoden ist gemein, dass sie einen Einfluss auf die Lagrange'schen Multiplikatoren  $\beta_i$  haben. Statt einer starr vorgegebenen Maximalgröße, die für alle gleich  $0 \leq \beta_i \leq C$  ist, wird für jeden Multiplikator ein individueller Wertebereich  $\beta_i \in [0, c_i]$  zugelassen, wodurch Punkte der Trainingsmenge mit geringer Zuverlässigkeitsbewertung auch mit einem geringeren Gewicht in der Abstandsfunktion  $R$  die die Entscheidungsgrenze definiert zum Tragen kommen. Der Zweck dieser Methoden ist die automatisierte Behandlung von Ausreißern. In dem Anwendungsfall in dieser Arbeit geht es nicht um Ausreißer. Vielmehr haben verschiedene Punkte einer Trainingsmenge einen unterschiedlichen Konfidenzwert, obwohl sie alle zur Klasse der realisierbaren Fahrpläne gehören – zumindest unter der Annahme, dass die Voraussetzung unter der die Trainingsmenge erstellt wurde zum Zeitpunkt der Benutzung des Modells noch gilt. Dies ist aber in der Regel nicht mehr vollständig der Fall. Daher soll hier eine konfidenzgewichtete SVDD (confidence score weighted, csw-SVDD) verwendet werden, die eine Gewichtung anhand von Wahrscheinlichkeiten vornimmt, mit denen Fahrpläne auch bei veränderten Bedingungen unverändert umsetzbar sind.

Zunächst sei der Konfidenzwert eines Fahrplans wie folgt definiert:

**Definition 14.1** (Konfidenz eines Fahrplans). Sei  $p$  ein Fahrplan über  $d$  Zeitintervalle der Einheit  $U$  aus einem Sample  $\mathcal{X}$ , das ausgehend von einem Startzustand  $z_0 \in Z_U$  erstellt wurde. Sei  $\Omega(z_0)$  eine Menge von Variationen des Anfangszustandes und  $\mathcal{F}[\Omega(z_0)] \subseteq \mathbb{R}^d$  der Raum der realisierbaren Fahrpläne, die aus mindestens einem der Startzustände aus  $\Omega(z_0)$  heraus gefahren werden können. Der Konfidenzwert von  $p$  ist definiert als

$$k(p) = P(p \in \mathcal{F}[\Omega(z_0)] | p \in \mathcal{F}[z_0]) = \frac{|\{x | x \in \mathcal{F}[z] \forall z \in \Omega(z_0)\}|}{|\Omega(z_0)|}. \quad (14.2)$$

Hierbei ist zu beachten, dass Prognosen auf denen die Realisierbarkeit und damit der aus dem Startzustand heraus resultierende Lösungsraum  $\mathcal{F}[z_0]$  beruhen in diesem Fall Variablen sind und somit o.B.d.A. als Teil des Startzustandes anzusehen sind.

Der Begriff *Variation des Anfangszustandes* ist in dieser Definition noch recht undeutlich

gefasst. Die genaue Art und der Umfang der Variation des Startzustandes einer Anlage (insbesondere unter Berücksichtigung von Variationen in Prognosen die über die Realisierbarkeit eines Fahrplans mitentscheiden) ist stark vom Anlagentyp abhängig. Im Falle eines BHKW kann hier z. B. von einer gewissen Schwankung im Füllstand des Pufferspeichers ausgegangen werden. Hinzu kommen Prognoseunsicherheiten über den Verlauf (die Höhe) des thermischen Bedarfs. Vollkommen unbekannt sind derzeit auch noch die Verteilungen dieser diversen Variationen. Hier ist im Nachgang noch weitere Forschungsarbeit erforderlich. Zur Überprüfung der Integration auf konzeptioneller Ebene kann zunächst auf die (in der Realität vermutlich nicht zutreffende) Annahme einer Normalverteilung zurückgegriffen werden, da die genaue Kenntnis der Verteilungen zur Überprüfung, ob der Ansatz generell funktioniert oder nicht, nicht erforderlich ist. Generalisierungen der so quantifizierten Aussagen sind dann zwar nicht zulässig – jedoch auch nicht Ziel dieses Kapitels oder dieser Arbeit.

Mit dem oben definierten Konfidenzwert, der hier angibt, mit welcher Wahrscheinlichkeit ein Fahrplan auch dann noch realisierbar ist, wenn sich die Voraussetzungen unter denen die Realisierbarkeit festgestellt wurde bis zu einem gewissen Grad ändern, kann nun das Verfahren aus Gleichung (14.1) verwendet werden. In der Praxis kann der Konfidenzwert mit Hilfe eines Anlagenmodells einfach bestimmt werden indem für jeden Fahrplan aus dem Sample das Anlagenmodell mit einer hinreichend großen Anzahl zufälliger Variationen parametrisiert wird um zu zählen, wie oft er dennoch realisierbar ist. Art der Variation und deren Verteilung müssten vermutlich spezifisch zum Anlagentyp vorab experimentell ermittelt oder aus den Beobachtungen im Laufe des Betriebs der Anlage geschlossen werden.

### 14.3. Simulative Evaluation

Für die Evaluation dieses Ansatzes wurde folgendes Vorgehen gewählt: Für das Modell eines Untertischwarmwassergerätes wurde eine Trainingsmenge an umsetzbaren Fahrplänen unter Annahme einer bestimmten Prognose für die Warmwasserzapfung erstellt. Mit der Trainingsmenge wurde einmal das herkömmliche Suchraummodell trainiert. Zum anderen wurden für die einzelnen Fahrpläne aus der Trainingsmenge Zuverlässigkeitsbewertungen erstellt, um mit derselben Trainingsmenge das erweiterte Suchraummodell trainieren zu können. Die Zuverlässigkeitsbewertung einzelner Fahrpläne wurde mithilfe des Simulationsmodells für die Anlage bestimmt. Für jeden einzelnen Fahrplan wurde mit dem Anlagenmodell geprüft, ob dieser auch mit einer anderen Warmwasserzapfung als der prognostizierten umsetzbar ist, Hierzu wurde als abweichende Zapfung jeweils das prognostizierte Zapfprofil zufällig verändert, indem für jedes Zeitintervall normalverteilte Zufallswerte (mit gegebener, fester Standardabweichung) zu dem Prognosewert addiert wurden (negative Werte wurden auf 0 korrigiert). Eine Normalverteilung der Abweichung wurde wie zuvor erläutert angenommen, da die wahre Verteilung unbekannt ist. Die Größe der Abweichung ist so gewählt, dass das Klassifikationsverfahren an seine Grenzen kommt und somit eher eine Aussage möglich wird, wie groß die Abweichung maximal sein darf (unter der Annahme der Korrektheit der Verteilung), damit das Suchraummodell noch ausreichend gute Aussagen liefert.

#### 14. Integration von Prognoseunsicherheiten für eine robuste Planung

Tabelle 14.1.: Vergleich der Klassifikatoren für einen Warmwasserboiler mit verschiedenen (prognostizierten) Zapfprofilen für die erwartete Warmwasserentnahme. Angegeben ist die Trefferquote für die Klassifikatoren bei verschiedenen großen Prognoseabweichungen. Die Abweichung der Wasserentnahme von der Prognose ist normalverteilt mit Standardabweichung  $\sigma$  in kJ.

$\sigma$ / kJ	Zapfprofil 1		Zapfprofil 2		Zapfprofil 3		$\bar{\Delta}$ / %
	svdd	csw-svdd	svdd	csw-svdd	svdd	csw-svdd	
18	0.969 ± 0.053	0.981 ± 0.048	0.973 ± 0.047	0.984 ± 0.038	0.970 ± 0.053	0.981 ± 0.048	0.91
27	0.943 ± 0.085	0.965 ± 0.072	0.949 ± 0.081	0.969 ± 0.066	0.943 ± 0.085	0.965 ± 0.072	2.15
45	0.841 ± 0.190	0.904 ± 0.149	0.860 ± 0.187	0.918 ± 0.120	0.841 ± 0.190	0.904 ± 0.149	6.99
90	0.527 ± 0.341	0.674 ± 0.319	0.532 ± 0.340	0.699 ± 0.298	0.527 ± 0.341	0.674 ± 0.319	30.24
135	0.313 ± 0.341	0.513 ± 0.381	0.281 ± 0.324	0.501 ± 0.372	0.308 ± 0.339	0.505 ± 0.384	74,37

Mit hinreichend vielen (10000 in den Versuchen hier) zufällig variierten Prognosen kann der Prozentsatz derjenigen Fälle bestimmt werden, in denen ein Fahrplan trotz Prognoseabweichung gegebener Größe noch umsetzbar ist. Je höher dieser Prozentsatz ist, desto zuverlässiger ist der Fahrplan.

Die so ermittelten Zuverlässigkeitsbewertungen wurden als Gewichtung für das Training des erweiterten Modells verwendet. Nun kann für beide Modelle eine Bewertung über deren Klassifikatorfähigkeiten vorgenommen und gegenübergestellt werden. Tabelle 14.1 zeigt das Ergebnis. Hierfür wurden beide Klassifikatortypen für das jeweils gleiche Modell eines Untertischwarmwassergeräts trainiert. Die Realisierbarkeit von Fahrplänen eines solches Gerätes ist stark von der Prognose des zur erwartenden Verlaufs für die Warmwasserzapfung abhängig. Anschließend wurden zufällige Lastgänge generiert und jeweils mit Hilfe des Klassifikators und mit Hilfe des Anlagenmodells (allerdings mit zufälliger Abweichung von der Warmwasserzapfprognose) auf Realisierbarkeit hin überprüft. Hieraus können dann beispielsweise die  $F_1$  Maße aus Recall und Precision (vgl. Abschnitt 11.4 auf Seite 160) o. ä. für den Vergleich berechnet werden. Tabelle 14.1 zeigt die Ergebnisse für einstündige Szenarien mit 15-Minuten-Auflösung für folgende drei (artifizielle) Zapfprofile:  $zp_1 = (180 \text{ kJ}, 0 \text{ kJ}, 0 \text{ kJ}, 720 \text{ kJ})$ ,  $zp_2 = (0 \text{ kJ}, 1440 \text{ kJ}, 180 \text{ kJ}, 540 \text{ kJ})$  und  $zp_3 = (180 \text{ kJ}, 90 \text{ kJ}, 90 \text{ kJ}, 180 \text{ kJ})$ . Die Profile wurden im simulierten Betrieb zufällig verrauscht durch Aufaddieren eines normalverteilten Zufallswertes verschiedener Größe mit Standardabweichungen von 18 bis 135 kJ. Werte kleiner 0 wurden auf 0 korrigiert, um keine negative Zapfung zu haben. Wie aus Tabelle 14.1 ersichtlich variiert die erreichte mittlere Verbesserung durch den csw-SVDD-Klassifikator beachtlich. Dies liegt daran, dass der einfache SVDD-Klassifikator mit geringen Prognoseabweichungen selbst noch sehr gut zurechtkommt. Je größer die Prognoseabweichung, desto mehr spielt der verbesserte Klassifikator seinen Vorteil aus. Dennoch nimmt absolut gesehen, die Leistung beider Klassifikatoren mit zunehmender Prognoseabweichung ab. Bei noch größeren Abweichungen sind praktisch beide Klassifikatoren nicht mehr geeignet. Während der konfidenzgewichtete Klassifikator mittelstarke Prognoseabweichungen gut handhabt, reagiert ein hieraus abgeleiteter Dekoder deutlich sensitiver auf Prognoseabweichungen. Tabelle 14.2 zeigt das Ergebnis einer Versuchsreihe für Zapfprofil 3. Angegeben ist hier jeweils die Fehlerrate.

In den Ergebnissen zur Bewertung des Klassifikators wurde in diesem Kapitel zum Vergleich die Trefferquote (Recall) herangezogen. Die Genauigkeit sinkt in beiden Fällen stark ab. Dies ist unmittelbar einsichtig. Die Genauigkeit beschreibt die Wahrscheinlichkeit, dass ein gefundener Fahrplan auch realisierbar ist. Da hier die Realisierbarkeit unter

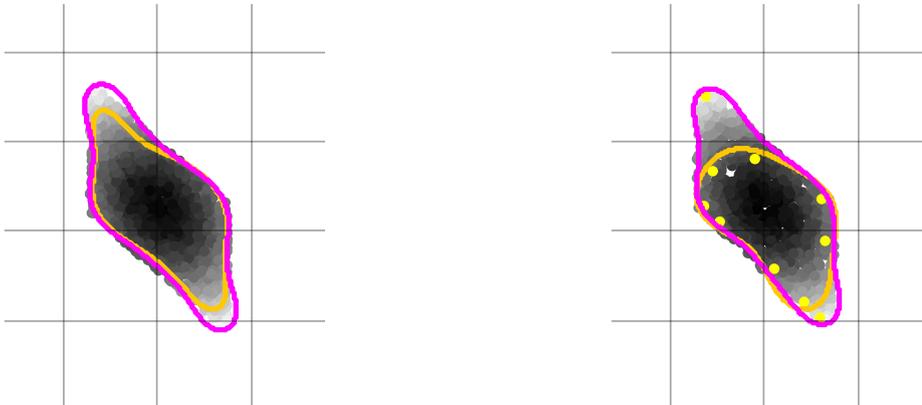


Abbildung 14.1.: Vergleich der Entscheidungsgrenzen an zwei Beispielen mit unterschiedlicher Prognoseabweichung. Die rote Grenze markiert jeweils die Klassenbegrenzung durch die normale SVDD. Die orange Grenze entsteht jeweils durch csw-SVDD.

$\sigma / k_j$	svdd	csw-svdd
9	$0.074 \pm 0.243$	$0.003 \pm 0.056$
18	$0.350 \pm 0.445$	$0.025 \pm 0.152$
27	$0.524 \pm 0.469$	$0.098 \pm 0.293$
45	$0.738 \pm 0.411$	$0.594 \pm 0.468$
67.5	$0.842 \pm 0.355$	$0.804 \pm 0.387$

Tabelle 14.2: Vergleich der Dekoderfehler (Anteil falsch rekonstruierter Fahrpläne) für verschieden große mittlere Prognoseabweichungen am Beispiel eines Boilers.

geänderten Voraussetzungen geprüft wird, ist klar, dass die Genauigkeit abnimmt. Dies ist bei beiden Klassifikatoren etwa in gleichem Umfang der Fall, da die Realisierbarkeit eine Eigenschaft der Fahrpläne ist. Mit dem neuen Klassifikator ist jedoch eine höhere Trefferquote verbunden, die die Wahrscheinlichkeit reflektiert, dass ein Fahrplan der auch unter den geänderten Voraussetzungen immer noch realisierbar ist auch gefunden wird.

Einen visuellen Eindruck von dem, was die Einführung von Konfidenzgewichten in den SVDD-Trainingsprozess bewirkt, gibt Abbildung 14.1. Die Entscheidungsgrenze wird dahingehend angepasst, dass Bereiche mit Anhäufungen niedrig bewerteter Punkte ausgespart werden ohne, dass hierfür explizit ein Schwellwert vorgegeben werden muss, ab dem Punkte nicht mehr zu Trainingsmenge gehören sollen. Für ein Vorgehen, bei dem vorab schlecht bewertete Punkte aus der Trainingsmenge aussortiert werden, müsste auch von Fall zu Fall entschieden werden, ab welcher Bewertung Punkte ausgeschlossen werden. Durch csw-SVDD wird die Verteilung der Konfidenzbewertungen unüberwacht mit einbezogen. Abgebildet ist jeweils der Lösungsraum für ein Warmwassergerät für zwei Zeitperioden. Die Farbe kennzeichnet die Zuverlässigkeit eines Fahrplans; je dunkler desto zuverlässiger. Die Zuverlässigkeit wurde genau so bestimmt wie in dem Experiment aus Tabelle 14.1. Die Entscheidungsgrenze für die modifizierte Variante (orange) ist nach innen in Richtung auf die zuverlässigeren Fahrpläne verschoben. In der rechten Abbildung 14.1(b) sind auch die Supportvektoren, wie sie nach der modifizierten Version entstehen mit eingezeichnet. Ein Teil der Supportvektoren (gelbe Punkte) liegt als bounded Supportvektoren im Datenraum nun außerhalb der Entscheidungsgrenze. Dies liegt allgemein an der mittleren Verkleinerung des Parameters  $C$  (vgl. Abschnitt 5.3.4 auf Seite 68). Da diese auch in die Erstellung des Dekoders eingehen erklärt sich auch, dass der Dekoderteil nicht im gleichen Maße von der Modifikation profitiert.

## 14. Integration von Prognoseunsicherheiten für eine robuste Planung

Tabelle 14.3: Vergleich der Trefferquoten für weitere Szenarien mit unsicheren Prognosen.

Szenario	svdd	csw-svdd
Kühlschrank 1	0.927 ± 0.045	0.994 ± 0.025
Kühlschrank 2	0.747 ± 0.028	0.829 ± 0.085
BHKW	0.838 ± 0.260	0.884 ± 0.226

Tabelle 14.4.: Vergleich der Korrektklassifikationsraten bei der Klassifikation von Warmwasserbereitermodellen.

$\sigma$ / kj	Dim. 6		Dim. 8		Dim. 10	
	SVDD	csw-SVDD	SVDD	csw-SVDD	SVDD	csw-SVDD
60	0.7537 ± 0.1467	0.8983 ± 0.1197	0.6725 ± 0.2121	0.8356 ± 0.2390	0.5992 ± 0.2218	0.7192 ± 0.3348
120	0.5973 ± 0.1241	0.7293 ± 0.1895	0.1374 ± 0.2011	0.2767 ± 0.3460	0.0785 ± 0.1466	0.0992 ± 0.2416

Beispiele für weitere Szenarien zeigt Tabelle 14.3. Betrachtet wurden folgende drei Szenarien:

**Kühlschrank 1:** Zur Nachbildung der nicht vorhersehbaren Benutzung des Kühlschranks wurde das zufällige Hineinstellen von Wärmequellen simuliert. Für zwei jeweils zufällig (gleichverteilt im Simulationszeitraum von 2 Stunden) gewählte Zeitpunkte wurde eine zusätzliche thermische Masse mit  $30 \pm 5$  kJ Energiegehalt (bedingt durch die Temperaturdifferenz zum restlichen Inhalt) hinzugefügt. Das entspricht etwa 500g Lebensmittel mit Zimmertemperatur.

**Kühlschrank 2:** Hier wurde simuliert, dass die Starttemperatur welche als Startbedingung und somit als Voraussetzung für die Berechnung des Lösungsraums angenommen wurde, gerade bei Geräten wie einem Kühlschrank nur schwerlich genau vorhergesagt werden kann. Daher wurde als Prognoseunsicherheit in diesem Szenario eine Variation der Starttemperatur (gleichverteilt aus dem gesamten Bereich der Temperaturspreizung der Zweipunktregelung von 3° bis 5°C) angenommen.

**BHKW:** Hier wurde angenommen, dass vorhergesagte Wetterbedingungen von der Prognose abweichen und so zu einem veränderten Wärmebedarf als Grundlage für die Planung eines BHKW führen. Hierbei wurde eine Variation der Außentemperaturen um  $\pm 5^\circ$  C für einen Zeitraum von 8 Stunden (auf Stundenbasis) simuliert.

Zumindest für hinreichend kurze Zeiträume liefert die Erweiterung des Klassifikators vielversprechende Ergebnisse. Betrachtet werden konnten hier im Wesentlichen nur Zeiträume bis zu 2 Stunden, da ansonsten bei der Bestimmung der Kennzahlen für die Beurteilung der Klassifikatoren dieselben Berechenbarkeitsprobleme beim Raten zufälliger Fahrpläne (für die Klassifikation durch Klassifikator und Anlagenmodell) aufgetreten wären, die bereits in Abschnitt 8.2 auf Seite 102 beschrieben wurden.

Um auch Szenarien mit längeren Produktzeiträumen zu überprüfen, wurde ein Versuch zum Vergleich von Korrektklassifikationsraten vorgenommen. Die Korrektklassifikationsrate wurde hierbei wie folgt bestimmt: mit Hilfe des Samplingverfahrens wurde eine

Tabelle 14.5: Vergleich der Korrektklassifikationsraten für Warmwasserbereiter für 24-dimensionale Fahrpläne. Berücksichtigt sind verschieden große Prognoseabweichungen ( $\sigma$ ) in einer begrenzten Anzahl von  $n$  Zeitintervallen.

$\sigma, n$	svdd	csw-svdd
60, 3	0.8431 ± 0.0939	0.9371 ± 0.0966
120, 1	0.8644 ± 0.1323	0.9507 ± 0.1051
120, 2	0.6802 ± 0.2196	0.9372 ± 0.1062
120, 3	0.5175 ± 0.2658	0.9027 ± 0.1291

weitere Stichprobe durch das Anlagensimulationsmodell bestimmt. Zudem wurde eine gleich große Menge zufälliger Fahrpläne bestimmt, welche mit hoher Wahrscheinlichkeit nicht realisierbar sind (vgl. Abschnitt 8.2 auf Seite 102). Die so erzeugte Fahrplanmenge (jeweils 1 Mio.) wurde mithilfe des Anlagenmodells und des jeweiligen Klassifikators auf Realisierbarkeit geprüft, sodass der prozentuale Anteil korrekt klassifizierter Fahrpläne bestimmt werden kann. Dieses Vorgehen ist auch für höher-dimensionale Fahrpläne noch sinnvoll berechenbar.

Tabelle 14.4 zeigt das Ergebnis für ein Warmwassergerät für Fahrpläne über bis zu 10 Zeitintervallen. Es zeigt sich sehr schnell, dass bei Prognoseunsicherheiten in *jedem* Zeitintervall sehr schnell von beiden Verfahren keine brauchbaren Ergebnisse mehr geliefert werden. Tabelle 14.5 zeigt das Ergebnis für Fahrpläne mit 24 Zeitintervallen, wobei lediglich für  $n = 1$  bis  $n = 3$  Zeitintervalle eine zufällige Prognoseabweichung generiert wurde. In diesem Fall zeigt sich, dass das Verfahren auch für höher-dimensionale Fahrpläne einsetzbar ist.

## 14.4. Fazit

In der vorausschauenden Wirkleistungsplanung bleibt es nicht aus, dass im Rahmen von Modellierungen diverse Prognosen und Annahmen einbezogen werden müssen, die alle mit einem gewissen Fehler behaftet sind. Eine robuste Einsatzplanung sollte bereits in der Planungsphase solche zu erwartenden Abweichungen einbeziehen können. Eine wichtige offene Frage war bisher noch die nach der Integration von Robustheitskriterien in den Supportvektor-Ansatz. Die Robustheit eines Einsatzplans definiert sich darüber, wie gut der Plan auch unter geänderten Begleitumständen noch funktioniert. Ein robuster Plan sollte kleinere Prognosefehler und damit Änderungen der Voraussetzungen unter denen er erstellt wurde verkraften können – d.h. dennoch abgefahren werden können, ohne dass größere Umplanungen erforderlich werden.

Die hier vorgestellte Erweiterung des Klassifikators im Suchraummodell stellt eine erste Richtung vor, wie eine solche Integration von Robustheitskriterien zu bewerkstelligen sein könnte. Erweiterungen des Ansatzes könnten auch eine Modifikation der Trainingsmenge vorsehen und diese aus verschiedenen parametrisierten (mit unterschiedlichen Prognoseabweichungen initialisierten) Anlagenmodellen zusammensetzen. Diese erweiterte Trainingsmenge könnte ebenfalls mit Konfidenzwerten ähnlich wie nach dem aufgezeigten Bewertungsschema versehen werden, um so ein gemeinsames Modell für die überlappenden Samples aus verschiedenen Abweichungen zu generieren. Problematisch ist die Erweiterung des Klassifikators ggf. wenn keine Prognoseabweichung auftritt. In diesem Fall dürfte die Klassifikation in jedem Fall zu konservativ erfolgen, d.h. es werden unnötig viele Fahrpläne abgelehnt, obwohl diese realisierbar wären. Hier scheinen im Nachgang weitere Forschungen zielführend und sinnvoll.

Ziel war es einen Weg für eine mögliche Erweiterung bezüglich des Einbezugs von Robustheitskriterien zu weisen. Dieser konnte in diesem Kapitel aufgezeigt werden.



# 15. Faire Gewinnverteilung

*What has destroyed every previous civilization  
has been the tendency  
to the unequal distribution of wealth and power.*

---

*(Henry George)*

## 15.1. Einleitung

Eine weitere interessante Fragestellung – aufsetzend auf das zuvor beschriebene Problem der dynamischen Verbundbildung – ergibt sich im Zusammenhang mit der fairen Aufteilung des von einer heterogenen Gruppe von Anlagen erwirtschafteten Gewinns durch gemeinschaftliches Erbringen eines Wirkleistungsproduktes.

Statische Virtuelle Kraftwerke, wie sie heute oft beschrieben werden, haben in der Regel genau einen Betreiber. Entweder gehören diesem einen Betreiber alle Anlagen, oder aber er verfügt abgesichert durch ein geeignetes Contracting über die Kontrolle zur Steuerung der Anlagen. Letzteres ist beispielsweise beim Schwarmstrom-Konzept von Lichtblick der Fall (<http://www.lichtblick.de/>). Der Vorteil eines einzelnen Betreibers liegt darin, dass auch ein erwirtschafteter Gewinn nicht aufgeteilt werden braucht. Im Falle dynamischer Virtueller Kraftwerke ist dies nicht mehr der Fall. Hier finden sich Anlagen unterschiedlicher Betreiber dynamisch zu einem Verbund zusammen der sich produktspezifisch ausbildet und somit auch nur für die Dauer des Produktes besteht. Erwirtschafteter Gewinn muss unter den Verbundteilnehmern im Anschluss an die Produktlieferung aufgeteilt werden, da diese dann ihre temporäre Beziehung beenden und sich (vermutlich mit Anlagen anderer Betreiber) für eine neues Produkt zusammenfinden.

Eine solche Gewinnaufteilung muss gewisse Kriterien erfüllen. Sie muss die verschiedenen Beiträge der einzelnen Anlagen zur Erfüllung des Produktes ebenso berücksichtigen, wie das Einbringen einer Flexibilität für mögliche Umplanungen. Letzteres ist für die Zuverlässigkeit der Produkterfüllung und somit bei der Vermeidung von ggf. Strafzahlungen von Bedeutung. Während sich der Produktbeitrag recht einfach beziffern lässt, wirft die Bewertung der Flexibilität Schwierigkeiten auf. Die Gewinnaufteilung muss vor allem fair sein. Für die Bestimmung eines fairen Verteilungsschlüssels ist die Anwendung spieltheoretischer Ansätze unvermeidbar. Hierdurch entsteht meist ein zusätzliches Komplexitätsproblem bei der Bestimmung der Verteilungsschlüssel insbesondere, wenn die Bewertung abstraktere Konstrukte wie die Größe der Flexibilität mit einbeziehen soll.

Ein Lösungsansatz kann auch hier darin bestehen auf eine verteilte Berechnung des Schlüssels in einem dezentralen Verfahren zu setzen. Hierfür soll in diesem Kapitel gezeigt werden, dass die verteilte Wirkleistungsplanung mit einem Näherungsverfahren zur Be-

stimmung so genannter Shapley-Werte (einer elementaren Grundlage zur Ermittlung fairer Verteilungsschlüssel) so kombiniert werden kann, dass verteilt eine Berechnung selbst bei Einbezug individueller Suchräume hinreichend performant durchgeführt werden kann. Hierzu müssen ein Näherungsverfahren für die Bestimmung des Verteilungsschlüssels und das Suchraummodell miteinander kombiniert werden. Eine Möglichkeit das Suchraummodell hierfür einzusetzen soll in diesem Kapitel diskutiert werden. Eine erste Beschreibung der Idee findet sich bereits in (Bremer & Sonnenschein 2013c).

### 15.2. Lösungsansatz

Ein häufig verwendetes Konzept in Multiagentensystemen ist das Zusammenfassen von Agenten zu einer Gruppe, wenn diese gemeinsam ein Ziel besser erreichen können als einzelne Agenten (Mas-Colell et al. 1995; Kahan & Rapoport 1984). Man spricht in diesem Fall von Koalitionen. Oft gibt es auch Aufgabenstellungen in denen die Bildung von Gruppen unverzichtbar ist, da einzelne Agenten allein nicht über die nötigen Mittel verfügen, die Aufgabe zu lösen. Ein Beispiel hierfür ist die Wirkleistungsplanung. Ein einzelner Agent (bzw. die durch ihn vertretene Anlage) verfügt in der Regel nicht über genug Leistung oder hinreichend Flexibilität, um einen geforderten Produktlastgang allein zu erbringen. In heterogenen Gruppen bringt jeder Agent seine eigene Charakteristik ein und trägt auch individuell zum Gesamtergebnis bei. Im Bereich der Multiagentensysteme befasst sich die Theorie der kooperativen Spiele mit der Formierung von Koalitionen unter den Agenten. Die Bildung von Koalitionen mit gewünschten Eigenschaften steht dabei im Vordergrund (Fatima et al. 2008; Rapoport 1970). Ein wichtiger Beitrag der kooperativen Spieltheorie sind Konzepte zur fairen Aufteilung gemeinsam (in der Koalition) erzielter Vorteile bzw. Gewinne. Ein bekannter Vertreter dieser Konzepte ist der Shapley-Wert (Shapley 1953; Hsu & Soo 2009) für faire Gewinnverteilung. Wird kooperatives Verhalten bei den Agenten forciert durch höhere Gewinnaussichten im Falle einer Koalitionszugehörigkeit, so findet die Konkurrenz zwischen den Koalitionen statt und nicht mehr zwischen den Agenten. Man spricht in diesem Fall von einem Koalitionsspiel. Wenn der Nutzen von Agenten untereinander ausgetauscht werden kann, so spricht man von transferierbarem Nutzen. In der Praxis bedeutet dies, dass ein Agent in einer Koalition auch von dem Nutzen anderer Agenten mit profitieren kann. Bringt beispielsweise eine große Anlage genug Kapazität für die Präqualifikation ein, so profitieren alle in der Koalition davon, da sie nun am Markt teilhaben können: der Nutzen der Anlage ist transferierbar. Sei  $N$  eine Menge von Spielern (Agenten), dann ist ein Koalitionsspiel  $(N, v)$  mit transferierbarem Nutzen für die Spieler charakterisiert durch die Funktion  $v(S)$

$$v : 2^N \rightarrow \mathbb{R}, \text{ mit } v(\emptyset) = 0 \quad (15.1)$$

welche den Nutzen (den Wert) einer nicht leeren Koalition  $S \subseteq N$  beziffert. In diesem Sinne gibt  $v(S)$  den Wert einer Koalition an, den es zu verteilen gilt. Shapley (1953) definierte einen Wert für jeden Spieler in einer Koalition anhand dessen die zu erwartende Auszahlung an diesen Spieler beziffert werden kann.

### 15.2.1. Shapley-Werte in der Wirkleistungsplanung

Zunächst soll das Konzept des Shapley-Wertes kurz vorgestellt werden.

Sei  $N$  eine Menge von  $n$  Spielern (Agenten). Der Shapley-Wert  $\varphi_i$  von Spieler  $i$  gemäß einer charakteristischen Funktion  $v$  in einem Koalitionsspiel  $(N, v)$ , der bestimmt wie hoch der Gewinnanteil von  $i$  ist, ist definiert als

$$\varphi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - |S| - 1)!}{n!} (v(S \cup \{i\}) - v(S)). \quad (15.2)$$

Der Term  $v(S \cup \{i\}) - v(S)$  in (15.2) bezeichnet den marginalen Beitrag (marginal contribution) von Spieler  $i$  zu dem Wert der gesamten Koalition (Liben-Nowell et al. 2012). Shapley Werte erfüllen im Speziellen die folgenden Eigenschaften (vgl. z. B. Shapley 1953; Fatima et al. 2008; Bremer & Sonnenschein 2013c):

1. **Pareto-Effizienz (Effektivität):** Der Gesamtwert der Koalition ist immer vollständig verteilt:  $\sum_{i \in N} \varphi_i(v) = v(N)$
2. **Symmetrie:** Der Wert eines Spielers ist unabhängig von seiner Person, d.h. falls für je zwei Spieler  $i$  und  $j$  gilt: aus  $v(S \cup \{i\}) = v(S \cup \{j\})$  für jede Teilmenge  $S \subseteq N$  mit  $S \cap \{i, j\} = \emptyset$  folgt  $\varphi_i(v) = \varphi_j(v)$ . Der Wert eines Spielers, den er zu allen Koalitionen beiträgt ist immer gleich.
3. **Additivität:** Für je zwei Spiele  $\varphi_i(N, v)$ ,  $\varphi_i(N, v^*)$  muss gelten:  $\varphi_i(N, v) + \varphi_i(N, v^*) = \varphi_i(N, v + v^*) \forall i \in N$ .
4. **Zero Spieler:** Wenn der marginale Beitrag eines Spielers zu allen Koalitionen 0 ist, so ist auch sein Wert 0:  $v(S \cup i) = v(S) \forall S \Rightarrow \varphi_i(v) = 0$ .

Shapley-Werte können im Allgemeinen nicht effizient berechnet werden. Die Berechnung gehört zur Klasse der #p-vollständigen Probleme (Deng & Papadimitriou 1994), speziell für so genannte Voting Games (Fatima et al. 2008). Die Berechnungskomplexität hängt von der Bewertungsfunktion  $v$  ab, d.h. es gibt einige Spezialfälle für die effiziente Berechnungsmethoden bekannt sind. Ein Beispiel hierfür sind Matching Games (Aziz & Keijzer 2013).

Gewichtete Voting Games hingegen sind ein Beispiel für nicht effizient berechenbare Shapley-Werte (Fatima et al. 2008). In solchen Spielen bekommt eine Koalition genau dann einen Gewinn von 1, wenn die Summe von individuellen Gewichten, die die Spieler in die Koalition einbringen in Summe einen gegebenen Schwellwert erreichen. Im Wirkleistungsplanungsfall würde das bedeuten, ein Verbund würde belohnt genau dann wenn er in Summe mit den durch die Anlagen eingebrachten Fahrpläne  $p_i$  einen gewünschten Ziellastgang erreicht:  $\sum_i p_i \geq q$ . In der Praxis würde man wohl eher Verlangen, dass die Abweichung vom Ziellastgang unterhalb einer gegebenen Schranke ist. Dies ist aber äquivalent. Im Folgenden sei der Fall betrachtet, bei dem sich der Wert einer Koalition direkt aus der Güte der Annäherung an den geforderten Ziellastgang ergibt.

### 15.2.2. Bestimmung der Verteilungskoeffizienten

In dynamisch gebildeten Verbänden mit regelmäßig wechselnder Zusammensetzung kann nicht wie in statischen Virtuellen Kraftwerken davon ausgegangen werden, dass alle Anlagen im Besitz eines einzelnen Betreibers sind. Das bedeutet aber auch, dass am Ende der gemeinschaftlich erwirtschaftete Überschuss fair unter den beteiligten Betreibern aufgeteilt werden muss. Wenn der gesamte Prozess von der Verbundbildung bis zur Erbringung des Produktes automatisiert abläuft, so muss auch der Verteilungsschlüssel selbsttätig und eindeutig von den Anlagen bestimmt werden können. Ein möglicher Ansatz für einen Schlüssel wäre, dass jeder Beteiligte zunächst seine erbrachten Unkosten erstattet bekommt. Der im Anschluss verbleibende Überschuss wird unter den Anlagen aufgeteilt. Hierfür wird ein Verteilungsschlüssel benötigt. Soll die Verteilung fair erfolgen, so besteht eine Möglichkeit darin, die bereits beschriebenen Shapley-Werte für die Bestimmung eines solchen Verteilungsschlüssels zu verwenden, da auf diese Weise elementare Fairnesskriterien (im spieltheoretischen Sinn) sichergestellt werden können.

Der Wert einer Anlage für die Planungsfähigkeiten des gesamten Verbundes wird charakterisiert durch den Suchraum individueller Möglichkeiten, die eine Anlage in den Verbund einbringt. Soll diese Charakteristik in die Bewertung einer Anlage eingehen, so muss der Suchraum in die Bewertungsfunktion integriert werden können.

Zwei baugleiche Anlagen, die aus einem vergleichbaren Ausgangsbetriebszustand starten sollten für einen Verbund den gleichen Wert haben, auch wenn ein auf heuristischen Ansätzen beruhendes Planungsverfahren einer Anlage einen kleineren Fahrplan zuweist als der anderen. Dies basiert größtenteils auf zufälligen Entscheidungen und könnte genauso gut auch anders herum ausgehen. Einen kleineren Beitrag zum Einsatzplan zu liefern ist aber aus Verbund­sicht gegebenenfalls genauso wertvoll, da es für die Einhaltung des Ziellastgangs notwendig ist.

Benötigt wird also eine charakteristische Funktion, welche für derartige Situationen die Bewertung des Beitrags eines Agenten zu einer Koalition vornehmen kann. Es ist nicht möglich, im Rahmen dieser Arbeit dieses Thema erschöpfend zu behandeln. Dennoch soll zumindest die Möglichkeit des Einbezugs von Suchraumcharakteristiken in diese Bewertung erörtert werden. Um den Wert einer Anlage für eventuell notwendig werdende Umplanungen abzuschätzen, ist die Größe des Suchraums von Bedeutung. Soll jedoch bewertet werden, wie gut eine Anlage einen bereits bestehenden Teilverbund im Hinblick auf die Erreichung der Optimierungsziele der Wirkleistungsplanung (insbesondere im Hinblick auf verschiedene Ziele) voranbringen kann, so ist das entsprechende Optimierungsergebnis zu betrachten (vgl. auch Bremer & Sonnenschein 2013c).

Der Einfluss einer Anlage auf die Verbesserung des Optimierungsergebnisses soll im Folgenden exemplarisch näher betrachtet werden. Eine detailliertere Betrachtung wurde bereits in (Bremer & Sonnenschein 2013c) veröffentlicht. Sei zunächst der marginale

Beitrag eines Agenten zur Koalition in Bezug auf mögliche Verbesserungen des Optimierungsergebnisses aus der Wirkleistungsplanung (Gleichung 6.14) definiert als

$$\begin{aligned} c(j, S) &= \left[ \delta_{max} - \delta_{min} \left( \zeta^{(P)}, \mathbf{p}_j + \sum_{i \in S} \mathbf{p}_i \right) \right] - \left[ \delta_{max} - \delta_{min} \left( \zeta^{(P)}, \sum_{i \in S} \mathbf{p}_i \right) \right] \\ &= \delta_{min} \left( \zeta^{(P)}, \sum_{i \in S} \mathbf{p}_i \right) - \delta_{min} \left( \zeta^{(P)}, \mathbf{p}_j + \sum_{i \in S} \mathbf{p}_i \right). \end{aligned} \quad (15.3)$$

Hierbei kennzeichne  $\delta$  wieder den Abstand zwischen dem Ziellastgang  $\zeta^{(P)}$  und dem durch den Einsatzplan  $\delta_{max}$  maximal möglichen Abstand (der Wert der Koalition ist umso höher je größer der Abstand zu diesem Worst-Case-Ergebnis). Der Wert eines Agenten ergibt sich somit aus der Differenz zwischen dem bestmöglichen Planungsergebnis, das eine Koalition mit bzw. ohne diesen Agenten erzielen könnte. Zu beachten ist hierbei, dass  $\delta_{min}$  ein Optimierungsergebnis (des Wirkleistungsplanungsproblems aus Abschnitt 6.5 auf Seite 86) darstellt. Für jede Auswertung der charakteristischen Funktion  $c$  sind also zwei Optimierungsprobleme zu lösen.

Nun ist bereits die Berechnung der Shapley-Werte im Allgemeinen berechnungsschwer. Verschiedene Näherungsverfahren zur effizienteren Abschätzung der Shapley-Werte wurden bereits vorgeschlagen (beispielsweise Mann & Shapley 1960). Ein kurzer Überblick findet sich auch in (Bremer & Sonnenschein 2013c). Ein solches Näherungsverfahren, welches linear mit der Anzahl der Agenten in der Koalition skaliert, beruht darauf, über einer gewissen Anzahl zufällig gewählter Teilkoalitionen zu mitteln anstatt über der Potenzmenge aller Kombinationsmöglichkeiten (Fatima et al. 2008). Hierzu wird von jeder Koalitionsgröße eine fixe Anzahl an zufälligen Koalitionen gewählt, um für jede dieser Koalitionen den marginalen Beitrag eines Agenten  $j$  zu bestimmen. Sei  $S^X$  eine zufällige Koalition der Größe  $X$ . Dann ergibt sich eine Abschätzung des Shapley-Wertes für einen Spieler  $j$  und Bewertung (15.3) wie folgt (vgl. Fatima et al. 2008; Bremer & Sonnenschein 2013c):

$$\hat{\varphi}_j(c) = \frac{1}{n} \sum_{X=0}^{n-1} c(j, S^X). \quad (15.4)$$

Dies gilt für den einfachsten Fall mit genau einer Koalition von jeder Größe. Offensichtlich sind in diesem Fall  $2n$  Optimierungsprobleme zu lösen. Für den verteilten Greedy-Ansatz aus Kapitel 13 ab Seite 199 konnte empirisch eine lineare Berechnungskomplexität ermittelt werden (Bremer & Sonnenschein 2013c). Vermutlich gilt dies auch für den vorgestellten COHDA-Ansatz (eine Evaluation steht jedoch noch aus), sodass sich für die Approximation des Shapley-Wertes eines Agenten insgesamt eine Komplexität ergibt, die quadratisch mit der Anzahl der Agenten wächst: Es muss eine linear wachsende Anzahl an Optimierungsproblemen gelöst werden, der Größe selbst ebenfalls linear wächst. Abbildung 15.1 aus (Bremer & Sonnenschein 2013c) zeigt die benötigten Berechnungszeiten exemplarisch für verschiedene Koalitionsgrößen. Erste Ergebnisse zur Güte der Approximation finden sich in (Bremer & Sonnenschein 2013c). Tabelle 15.1 fasst die Ergebnisse für aus Shapley-Werten berechnete bzw. approximierte Gewinnverteilungsschlüssel zusammen. Im Szenario *Calculated* wurden die Shapley-Werte exakt berechnet. Da aber auch hier bereits approximierte Optimierungsergebnisse eingehen, weichen die Ergebnisse des oben vorgeschlagenen Approximationsverfahrens (Szenario *Estimated*) kaum hiervon ab. Die übrigen beiden Szenarien zeigen Verteilungsschlüssel für den Fall, dass einige Anlagen einen geeigneteren

## 15. Faire Gewinnverteilung

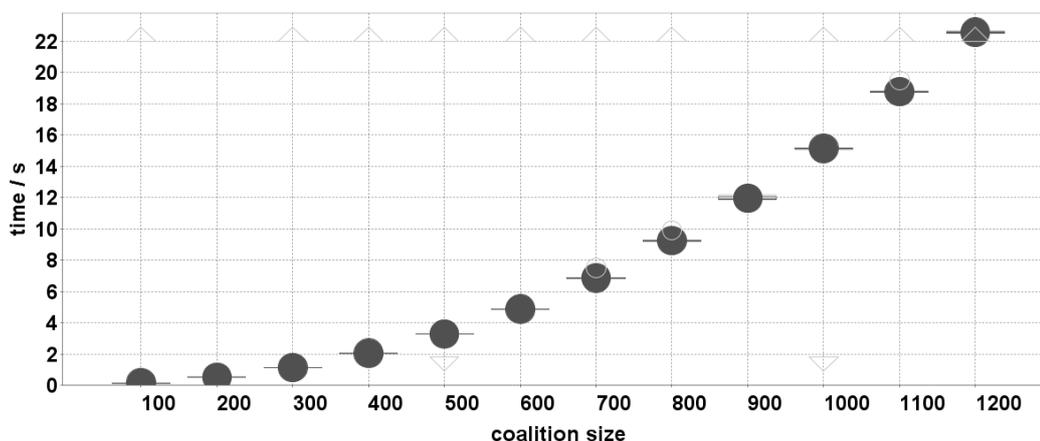


Abbildung 15.1.: Zeitkomplexität bei der verteilten Bestimmung der Shapley-Werte, unter Verwendung des verteilten Greedy-Verfahrens (Bremer & Sonnenschein 2013c).

Tabelle 15.1.: Verteilungsschlüssel für Erzeugerverbünde mit gleichartigen BHKW. Angegeben ist jeweils der auszuschüttende Prozentsatz des gemeinschaftlich erwirtschafteten Überschusses. In den Szenarien 2 bzw. 4 Special hatten jeweils 2 bzw. 4 Anlagen einen höheren Wärmebedarf zu Zeiten zu denen auch eine erhöhte elektrische Leistung nachgefragt war.

Anzahl BHKW	calculated	estimated	2 special	4 special
1	12.498±4.346e-3	12.497±4.990e-3	14.115±1.123e-1	13.156±2.024e-1
2	12.500±6.735e-4	12.500±3.848e-5	14.270±1.229e-1	13.014±2.341e-1
3	12.503±5.809e-3	12.504±6.326e-3	11.856±4.490e-2	13.301±2.517e-1
4	12.501±1.912e-3	12.502±2.597e-3	11.852±4.504e-2	13.056±2.284e-1
5	12.499±2.023e-3	12.499±1.811e-3	11.960±3.897e-2	11.586±3.772e-1
6	12.500±4.805e-3	12.503±5.826e-3	11.834±4.698e-2	11.916±2.637e-1
7	12.497±4.987e-3	12.497±5.608e-3	12.246±1.749e-2	11.953±2.331e-1
8	12.499±1.845e-3	12.499±2.378e-3	11.868±4.403e-2	12.017±3.061e-1

Suchraum aufweisen, weil sie genau zu Zeiten höherer Wirkleistungsnachfrage auch einen erhöhten Wärmebedarf zu befriedigen haben.

Der Vorteil bei der Verwendung der dezentralen Verfahren ist, dass jeder Agent seinen eigenen Shapley-Wert berechnen könnte und somit die Berechnungen für die gesamte Koalition parallelisiert werden können (unter der Annahme, dass agentenintern Wartezeiten durch Kommunikation für anderweitige Berechnungen genutzt werden können). Zwar müsste in diesem Fall jeder Agent auch die Suchraummodellierungen aller anderen Agenten kennen, es müssten aber keine privaten Anlagen-, Constraint- oder Kostenmodellierungen herausgegeben werden. Ein Austausch der Dekoder die von diesen Modellen abstrahieren ist ausreichend.

Eine erste Evaluation der auf diese Weise abgeschätzten Shapley-Werte bzw. der hieraus bestimmten Verteilungsschlüssel wurde bereits in (Bremer & Sonnenschein 2013c) vorgenommen und zeigt die generelle Machbarkeit von derartig zusammengesetzten Verteilungsschlüsseln. Dennoch stellt die Einbettung der Suchraumcharakteristik in den Verteilungsschlüssel nur einen ersten Schritt dar. Weitere Kriterien sind bei der Bewertung zusätzlich zu berücksichtigen. Die Erarbeitung eines vollständigen Verteilungsschlüssels

ist jedoch auch nicht Ziel dieser Arbeit. Daher wurde hier nur die generelle Möglichkeit der Einbettung von in dieser Arbeit erstellten Konzepten in ein solches Vorgehen diskutiert.

### **15.3. Zusammenfassung**

Der Wandel von statischen durch einen Besitzer betriebenen Virtuellen Kraftwerken hin zu dynamischen Verbänden mit Anlagen verschiedener Betreiber die gemeinschaftlich und produktorientiert Energieerzeugung (oder auch -verbrauch) planen, organisieren und gemeinschaftlich vermarkten, lässt eine neue Fragestellung entstehen: die Frage nach einer fairen Aufteilung des gemeinschaftlich erwirtschafteten Gewinns. Seit langem ist das Konzept des Shapley-Wertes ein bekanntes Mittel zur Bestimmung geeigneter Verteilungsschlüssel.

Wie in diesem Kapitel in einem ersten Ansatz gezeigt werden konnte, bietet das Dekoderkonzept auch hier einen Weg, um bei der Beurteilung des Wertes einer Anlage für einen Verbund nicht nur den letztlich zugewiesenen Fahrplan (gegebenenfalls Ergebnis einer Heuristik) zu berücksichtigen, sondern über eine Integration des gesamten Suchraums in die Beurteilung auch Flexibilitäten für eventuelle Umplanungen oder eingebrachte Sicherheitsreserven zu berücksichtigen. Insbesondere aber kann vorab beurteilt werden, wie gut der eingebrachte Suchraum zur Verbesserung der gemeinschaftlichen Lösung des gemeinschaftlichen Optimierungsproblems beiträgt.

Wenn in Zukunft zu erwarten ist, dass in den hier betrachteten Szenarien verteilte Verfahren auf allen Ebenen zum Einsatz kommen, dann wird auch die Bestimmung des Verteilungsschlüssels verteilt stattfinden müssen. Dies ist möglich. Sicherlich kann ein derart bestimmter Schlüssel allein zunächst nur die Richtung für eine Gewinnverteilung geben. Weitere Faktoren wie Reputation o. ä. werden ebenfalls Berücksichtigung finden müssen. Es ist aber auch nicht die Aufgabe hier ein vollständiges Konzept für eine Gewinnverteilung zu erarbeiten. Erweiterungen des Konzeptes auf nicht transferierbaren Nutzen (Shapley 1967) könnten ebenfalls eine Betrachtung wert sein.

Hier konnte zunächst gezeigt werden, dass Shapley-Werte durch den Einsatz von Supportvektor-Dekodern effizient verteilt bestimmt werden können, selbst wenn die Berechnung der wiederholten Integration von Optimierungsergebnissen bedarf, um den Suchraum adäquat zu berücksichtigen.



# 16. Interoperabilität

*If you think of standardization as the best that you know today,  
but which is to be improved tomorrow;  
you get somewhere.*

---

*(Henry Ford)*

## 16.1. Einleitung

Für einen Einsatz in einem späteren Produktivsystem müsste sichergestellt sein, dass die hier entwickelten Konzepte kompatibel mit allen Komponenten (und den dort verwendeten Konzepten, Datenstandards und Kommunikationsprotokollen) sind oder, dass eine solche Kompatibilität zumindest hergestellt werden kann. Ersteres kann im Rahmen dieser Arbeit nicht geleistet werden, daher sollen zumindest einige heute geläufige Standards auf Kompatibilität hin betrachtet werden.

Das Thema Interoperabilität ist im Rahmen dieser Arbeit auf zwei Ebenen zu betrachten. Zum einen ist dies die Integration in im Energiesektor verbreitete und etablierte Standards wie beispielsweise der Standard *Communication networks and systems for power utility automation - Part 7-420: Basic communication structure - Distributed energy resources logical nodes* (IEC 61850-7-420:2009 [2009](#)), mit dem die Verständigung zwischen verschiedenen Anlagen (bzw. mittels der gleichen Standards kommunizierender Agenten) sichergestellt wird.

Zum anderen muss auf Verfahrensebene sichergestellt werden, dass verwendete Kennzahlen bezüglich diverser für die Optimierung verwendeter Kosten untereinander vergleichbar sind oder ineinander umgerechnet werden können.

Welche Standards und Komponenten in einem späteren System genau zum Einsatz kämen ist heute noch spekulativ. Daher können an dieser Stelle die beiden zuvor genannten Aspekte der Interoperabilität wie sie für das hier entwickelte Konzept eine Rolle spielen lediglich exemplarisch betrachtet werden.

## 16.2. Integration in bestehende Standards

Eine Vielzahl von Standards hat Bedeutung im Umfeld der elektrischen Energieversorgung. Ein Großteil hiervon stammt aus dem Gebiet der Elektrotechnik. Für diese Arbeit von Interesse sind jedoch diejenigen Standards, die einen Bezug zur Informationsverarbeitung im Zusammenhang mit der Energieversorgung bzw. mit der Kommunikation zwischen

verschiedenen Beteiligten Akteuren im Netz haben. Zusätzlich zu den elektrotechnisch orientierten Kommunikationsstandards haben im Rahmen Agenten-basierter Steuerungsansätze auch Standards auf dem Gebiet von Agentensystemen eine Bedeutung für die Umsetzung im Anwendungsfall Smart Grid. Die wichtigsten Standards, ihre Bedeutung sowie die Integrationsaspekte für den in dieser Arbeit entwickelten Ansatz werden in den folgenden Unterabschnitten spezifisch für den jeweiligen Standard diskutiert.

Unabhängig von einem konkreten Standard muss es in jedem Fall möglich sein, das Suchraummodell als Datenmodell zu modellieren und gegebenenfalls zu übertragen. Abgesehen von einfachen, die Einheit beschreibenden Parameterwerten (z. B. maximale und minimale Wirkleistung) müssen im Wesentlichen Konstrukte existieren, die die Modellierung von Datenarrays für den Gewichtungsvektor und für eine Menge von Fahrplänen ermöglichen.

### 16.2.1. 61850

Die wichtigste Normenreihe für den Informationsaustausch in der elektrischen Energieversorgung stellt der Standard IEC 61850 dar, der von der International Electrotechnical Commission (IEC) herausgegeben wird. Die Schaltanlagenleittechnik (Substation Automation) benötigt zum Steuern, Schützen und Überwachen eine standardisierte Kommunikation für die kompatible Integration von Anlagen verschiedener (weltweiter) Hersteller (Brand & Wimmer 2002).

Um unterschiedlichen Geschwindigkeiten bei der Weiterentwicklung von Technologien Rechnung zu zollen und Modularität zu gewährleisten, ist beim Standard IEC 61850 das domänenspezifische Modell (Daten und deren Kommunikation aus funktionaler Sicht) vom Kommunikationsstack für die eigentliche Übertragung getrennt. Mithilfe eines standardisierten Mappings können die einmal modellierten Daten über verschiedene Kommunikationswege (und -protokolle) übertragen werden. Die Datenobjekte und ihre Parameter sind im Standard als so genannte Common Data Classes (CDC) definiert. Die Erweiterung auf verteilte Energieanlagen (IEC 61850-7-420:2009 2009) ergänzt CDCs für die Einbindung dezentraler Energieerzeugung. Dies umfasst unter anderem auch die folgenden beiden CDCs:

**Array (ARY):** Diese Klasse definiert ein Array (eigentlich als zeitliche Abfolge) von Messwerten. Es ist bereits vorgesehen, für jeden Wert ein Qualitätsattribut mitzuführen (evtl. geeignet für den Konfidenzwert eines Fahrplans; siehe Kapitel 14 ab Seite 209). Möglicherweise könnten hiermit Gewichtung und Konfidenz eines Fahrplans modelliert werden, allerdings sind die nicht benötigten Zeitstempel als zwingend anzugeben vorgesehen.

**Schedule (SCH):** Die Klasse Schedule definiert das Modell für einen Fahrplan und wäre somit auch geeignet, einen Supportvektor-Fahrplan im Modell zu beschreiben. Allerdings ist diese Klasse derzeit nur für das Setzen eines Fahrplans gedacht, nicht für die Abfrage oder die Übertragung innerhalb anderer Funktionen.

Somit liegen zwei wesentliche Klassen zur Modellierung des Suchraummodells in 61850 im Prinzip bereits vor. Notwendige Anlagenparameter, z. B. für die Skalierung der Fahrpläne (minimale bzw. maximale Leistung), sind im Datenmodell ohnehin vorgesehen, sodass

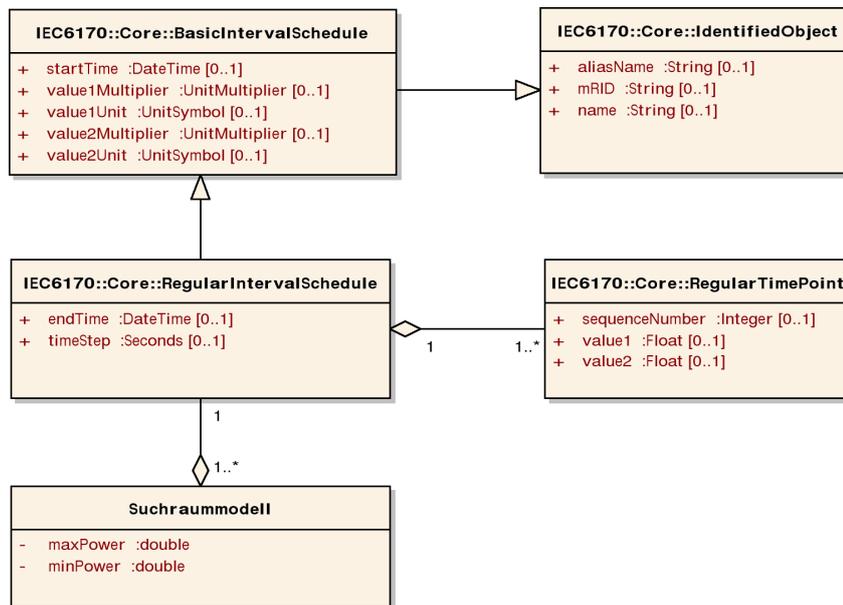


Abbildung 16.1.: Minimalbeispiel für ein Suchraummodell in IEC61970 CIM.

die Umsetzbarkeit des Suchraummodells mit dem Standard IEC 61850 zumindest bereits eine Grundlage hat. Erweiterungen wären definitiv notwendig bezüglich der Definition von Methoden zum Abruf sowie für das Zusammenfassen von Schedule-Objekten zu einer Menge von Schedules.

### 16.2.2. Common Information Model

Mit dem Standard IEC 61970-301 wurde ein semantisches Modell für die Beschreibung eines Stromversorgungssystems und seiner Komponenten geschaffen, das heute bei vielen Energieversorgern als Standard im Einsatz ist. Mit der Erweiterung IEC 61968-11 existiert ein semantisches Datenmodell für den Austausch von Informationen und die Kommunikation zwischen verschiedenen Softwareanwendungen bei den Energieversorgern. Zusammen bilden diese zwei Standards das Common Information Model (CIM) (McMorran 2007).

CIM ist nicht auf die Energiebranche spezialisiert, kommt aber dort häufig zum Einsatz. Technisch gesehen ist CIM eine Domänenontologie und wird bereits seit Mitte der 90er Jahre maßgeblich vom Electric Power Research Institute (EPRI) entwickelt.

CIM beinhaltet alle üblicherweise innerhalb eines Energieversorgungsunternehmens gehandhabten Datenobjekte und wird gepflegt als UML-Datenmodell (Uslar et al. 2005). CIM bietet auch die Möglichkeit, dynamische Daten zu modellieren, wie Verbraucher und ihre zeitvariable Lastcharakteristik (Uslar et al. 2005). Damit dürften auch hier bereits alle Basisdatentypen vorhanden sein, um das Suchraummodell abbilden zu können.

Eine erste Möglichkeit für ein Minimalbeispiel einer auf CIM basierenden Modellierung zur Kommunikation von Suchräumen ist in Abbildung 16.1 dargestellt.

### 16.2.3. OPC-UA

Die Object Linking and Embedding for Process Control (OPC) Foundation ist ein Industriekonsortium für offene Verbindungsprotokolle in der Automationstechnik. Die aktuelle Spezifikation OPC Unified Architecture (OPC-UA) vereinigt diverse Vorgängerstandards und ermöglicht den Wechsel von einer ursprünglich COM/DCOM-basierten Kommunikation zu einem für alle Plattformen offenen und serviceorientierten Ansatz. Eine gute Einführung in OPC-UA findet sich beispielsweise in (Mahnke et al. 2009).

Die Modellierung von zu übertragender Information erfolgt in OPC-UA mittels Objektorientierter Technologien und unterstützt auch Typhierarchien und Vererbung. Ein komplexes Modell kann für andere für den Zugriff einfach offengelegt werden, sodass für den Zugriff kein Mapping erforderlich wird. Hierdurch ergeben sich quasi keine Begrenzungen für die adäquate Modellierung eigener Datenmodelle (Mahnke et al. 2009); somit auch nicht für das Suchraummodell.

Der Standard OPC-UA eignet sich prinzipiell für alle Bereiche der Automation. Betrachtungen für den Sektor Smart Grid finden sich beispielsweise in (Rohjans & Specht 2013). Erste Ansätze für eine Integration von IEC 61850 und OPC-UA aber perspektivisch auch mit CIM finden sich z. B. in (Lehnhoff et al. 2011).

## 16.3. Vergleichbarkeit der Kostenkennzahlen

In Kapitel 10 ab Seite 133 wurde beschrieben, wie verschiedene Kostenkennzahlen einfach in das in dieser Arbeit entwickelte Konzept der Suchraummodellierung integriert werden können. Der Vektor mit den Werten für die Wirkleistung in den einzelnen Zeitintervallen kann erweitert werden, sodass zusätzliche Werte mit zugehörigen Kostenkennzahlen als weitere Elemente an das Ende des Vektors angehängt werden. Auf diese Weise erfolgt eine direkte Zuordnung der Kennzahlen zum Lastgang, den sie jeweils beschreiben.

Jedoch liegen diese Kennzahlen als reine Zahlenwerte vor. Sie haben für einen Algorithmus, der mit diesen Zahlen arbeitet, also zunächst keine semantische Bedeutung. Auf algorithmischer Ebene scheint klar, dass eine Integration in die Bewertung einzelner Lösungskonfigurationen über separate Zieldefinitionen in einem Mehrzielansatz erfolgen sollte.

Ein üblicher Ablauf in einem Mehrzielszenario wird die folgenden Schritte beinhalten (siehe Bremer & Sonnenschein 2013d): Jeder Agent erstellt zunächst ein Suchraummodell für die von ihm verantwortete Einheit anhand eines Anlagenmodells. Für jeden der Fahrpläne, die für das Training des Suchraummodells verwendet werden, können individuelle, einheitenspezifische Kennzahlen je Fahrplan berechnet werden. Mit dieser Trainingsmenge kann sodann das Modell trainiert werden, wodurch sowohl der Unterraum der realisierbaren Lastgänge (Domäne des Einheitenmodells) als auch der funktionale Zusammenhang zwischen Fahrplan und den Kennzahlen gelernt und im Modell kodiert werden kann.

Anschließend existiert je Einheit ein Suchraummodell, welches seine Fahrpläne individuell beschreibt. Zu jedem Fahrplan steht eine Menge von Kennzahlen zur Verfügung, die in der Zielfunktion des Verfahrens individuell als Nebenbedingung verwendet werden kann.

Somit wird es möglich, vordringlich das eigentliche Optimierungsziel (möglichst genau die notwendige Wirkleistung erreichen) zu bearbeiten, aber aus der Paretofront diejenigen Wirkleistungslösung zu wählen, welche die übrigen Ziele gut befriedigt.

Ein solches Modell erlaubt zwar je Suchraum lediglich eine 1:1 Beziehung zwischen Lastgang und der Menge der kodierten Kennzahlen, da Lastgang und zugehörige Kennzahlen in einem gemeinsamen Vektorraum definiert sind, Einheiten- und Suchraum-übergreifend sind jedoch verschiedene Kennzahlenmengen möglich. Jede Einheit kann eine eigene (Teil-)Menge von Kennzahlen kommunizieren und diese auch individuell variieren, falls notwendig. Unmittelbar einsichtig ist, dass nicht alle Kennzahlen auf alle Einheiten zutreffen; oder möglicherweise anlagenspezifisch angepasst sind. Ein Beispiel aus dem Umweltbereich sind die statischen Speicherverluste, die nur bei Vorhandsein eines Speichers zum Tragen kommen und je nach Speicherart unterschiedlich ausgeprägt sein können. Hier stellt sich noch die Frage, wie verschiedene Kennzahlen entweder auf eine festgelegte Menge von Kennzahlenklassen zurückgeführt oder aber geeignet ineinander umgerechnet werden können, um gemeinsam als Optimierungsziel (oder Menge von Zielen) genutzt werden zu können (Bremer & Sonnenschein 2013d). Gegebenenfalls müssen zusammen mit den Kennzahlen entsprechende Hinweise zur Integration mit übermittelt werden.

Eine Möglichkeit das Problem auf semantischer Ebene anzugehen wurde bereits in (Bremer 2012) vorgestellt. Hier wurde ein Weg zur Sicherstellung der Operabilität am Beispiel Umweltkennzahlen vorgestellt.

Basierend auf einer Ontologie (Löschner 2013), welche die semantischen Zusammenhänge der Kennzahlen beschreibt (einschließlich Vergleichbarkeit, mögliche Umrechnungen und Aggregation von Kennzahlen) wird eine Erweiterung des Suchraummodells vorgeschlagen (Bremer 2012). Hierzu muss lediglich für jede Kennzahl die im jeweiligen Suchraummodell enthalten ist die Klassenzugehörigkeit annotiert werden. Dies kann außerhalb des Fahrplan-Kennzahlen-Featurevektors geschehen, da innerhalb eines Suchraummodells die Kennzahlen (und ihre Reihenfolge in Vektor) für alle Fahrpläne gleich sind und diese Information somit nur einmal vorliegen muss. Daher ist es ausreichend, dem Suchraummodell eine Liste der Klassen (über welche sie in der Ontologie beschrieben sind) in der Reihenfolge in der sie auch in den Vektoren enthalten sind mitzugeben. Ein Planungsalgorithmus kann dann (Kenntnis der Ontologie vorausgesetzt) während der Optimierung hiervon Gebrauch machen, um Kennzahlen korrekt zu behandeln – auch im dezentralen Fall, da auch hier Suchraumerstellung und die Planung auf dem Suchraummodell logisch voneinander getrennt sind. Mögliche Änderungen der benutzten Kennzahlen im Falle einer geänderten Anlagenkonfiguration würden auch hier keine Änderungen an der Planungsimpementierung nach sich ziehen

## 16.4. Zusammenfassung

Interoperabilität bzw. deren konkrete Umsetzung ist nicht essentieller Bestandteil dieser Arbeit – insbesondere nicht da der Ansatz nicht auf konkrete Anwendungsfälle und damit im Zusammenhang stehende konkrete Interoperabilitätsfragen (falls überhaupt vorhanden) festlegt ist. Davon ungeachtet stellt sich die Frage nach der Interoperabilität zwischen verschiedenen Anlagen im Fall Smart Grid sehr wohl, sodass eine Diskussion zumindest der

## 16. Interoperabilität

---

Möglichkeiten hier unumgänglich ist. Es konnten Wege aufgezeigt werden, wie eventuelle Probleme bei der Verständigung im Rahmen der Optimierung und damit einhergehender Kommunikation angegangen werden können.

Aufgrund der Verwendung von lediglich geläufigen Konzepten für Datenstrukturen und Übertragung scheint einer Verwendung im Rahmen verschiedener Standards nichts im Wege zu stehen, wenngleich auch die konkrete Umsetzung einer standardkonformen Umsetzung im Rahmen dieser Arbeit nicht sinnvoll leistbar (und auch nicht Ziel der Arbeit) war.

Wie die konkrete Umsetzung der Standardisierung von Bewertungskennzahlen (bzw. deren semantische Interpretation) aussehen könnte ist ebenfalls noch unklar und muss derzeit Thema weiterer Arbeiten bleiben. Ein möglicher, genereller Weg für die Herstellung der Interoperabilität konnte in diesem Abschnitt aufgezeigt werden.

# 17. Generalisierung des Ansatzes

*Every word or concept,  
clear as it may seem to be,  
has only a limited range of applicability.*

---

*(Werner Heisenberg)*

## 17.1. Einleitung

In dieser Arbeit wurde ein Supportvektor-Ansatz für den Umgang mit Constraints in der verteilten Optimierung vorgestellt und zur Evaluierung vornehmlich im Kontext der Wirkleistungsplanung im Smart Grid betrachtet. Der vorgestellte Ansatz ist aber nicht auf diesen Anwendungsfall beschränkt. Daher soll abschließend in diesem Kapitel noch der Aspekt der Generalisierbarkeit diskutiert werden. Hierzu sind weitere Anwendungsfälle auf zwei Ebenen zu betrachten: zum einen weitergehende Fragestellungen, die sich im Zusammenhang mit der Wirkleistungsplanung im Smart Grid weiterhin ergeben, sowie Anwendungsfälle außerhalb des Anwendungskontextes Smart Grid.

Zwei weiterführende Fragestellungen aus dem Bereich der Wirkleistungsplanung wurden mit der Integration von Prognoseunsicherheiten und der Frage nach der Bestimmung einer fairen Gewinnverteilung bereits ausführlicher angerissen. An dieser Stelle sollen noch einige weitere Operationen auf dem Suchraum betrachtet werden, ohne diese im Detail auszuarbeiten. Zudem werden die mögliche Ausdehnung des Ansatzes auf Anwendungsfelder anderer Domänen betrachtet, sowie Grenzen der Anwendbarkeit aufgezeigt.

## 17.2. Erweiterungen des Ansatzes

Als erstes sollen an dieser Stelle mögliche Erweiterungen aus dem Umfeld des Smart Grid diskutiert werden. Diese umfassen (hier am Beispiel komplexer Leistungswerte) sowohl die Kodierung weiterer Informationen als auch die Umsetzung weiterer Operationen auf dem Suchraummodell.

### 17.2.1. Einsatz für komplexe Leistungswerte

Ein weiterer interessanter Aspekt für eine Erweiterung des Dekoderansatzes ergibt sich durch die Betrachtung komplexer Leistungswerte. Bei fortschreitender Ersetzung herkömmlicher Großkraftwerke müssen auch mehr und mehr netzstützende Maßnahmen

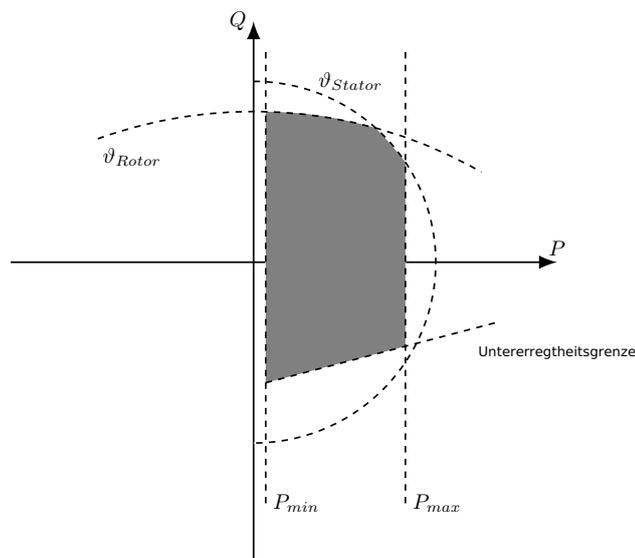


Abbildung 17.1.: Zulässiger Bereich für Wirk- und Blindleistungswerte einer Synchronmaschine in der komplexen Ebene; nach (Milano 2010), verändert.

durch dezentrale Anlagen übernommen werden. Hierzu zählt neben der Frequenz- u. a. auch die Spannungshaltung, die beispielsweise über die Bereitstellung von Blindleistung beeinflusst werden kann. So kann etwa ein am Leitungsende einspeisender Erzeuger die durch ihn verursachte Spannungsanhebung verringern indem er untererregt, d.h. induktiv, betrieben wird und so gleichzeitig Blindleistung aus dem Netz aufnimmt. Abbildung 17.1 zeigt am Beispiel eines Synchrongenerators die zulässigen Möglichkeiten, die Kombination aus Wirk- und Blindleistung zu variieren. In der komplexen Ebene können die Leistungswerte technisch bedingt immer nur innerhalb des grau markierten Bereiches liegen, da verschiedene technische Einschränkungen zu beachten sind. So ergeben sich im Beispiel in 17.1 aus (Milano 2010) neben den Einschränkungen durch minimale bzw. maximale Wirkleistung thermische Beschränkungen ( $\vartheta_{Rotor}$  und  $\vartheta_{Stator}$ ) welche die möglichen Stator- und Rotor-Ströme beschränken. Durch Variation des Erregerstroms kann das Verhalten des Generators bezüglich Aufnahme oder Abgabe von Blindleistung beeinflusst werden. Ähnliche Diagramme lassen sich auch für den Fall der Wechselrichtereinspeisung aufstellen (siehe beispielsweise Kühn 2007).

Um den in dieser Arbeit vorgestellten Dekoderansatz bezüglich der Verwendung komplexer Leistungswerte zu erweitern, müsste statt eines Wirkleistungswertes je Zeitintervall ein komplexer Leistungswert vorgesehen werden. Ein Fahrplan ist dann definiert als  $\mathbf{p}_{\mathbb{C}} = (p_0, p_1, \dots, p_d) \in \mathbb{C}^d$ . Hierfür sind verschiedene Ansätze denkbar. Im einfachsten Fall wird die Leistung in jedem Zeitintervall durch zwei reelle Werte dargestellt, wodurch sich die Dimension des Fahrplanvektors verdoppelt. Alle weiteren Ansätze könnten unverändert verwendet werden. Es gibt aber auch bereits erste Forschungsansätze, die sich mit der Erweiterung von Supportvektor Maschinen bezüglich komplexwertiger Datenräume befassen (z. B. Bouboulis et al. 2013). Hier könnten sich Adaptionmöglichkeiten ergeben. In jedem Fall müsste das Samplingverfahren erweitert werden. Hier kommen je nach Anlage

zusätzliche Nebenbedingungen für mögliche Kombinationen aus Wirk- und Blindleistung hinzu, die bei der Erstellung der Trainingsmenge einbezogen werden müssten.

Alternativ wäre auch denkbar, weiterhin nur Wirkleistungswerte zu verwenden und mögliche Blindleistungswerte anhand des (statischen) Diagramms im Nachgang zu den gegebenen Wirkleistungswerten zuzuordnen. Der Dekoder könnte in diesem Fall jedoch keine Fahrpläne erzeugen, die direkt aus gültigen Wirk- und Blindleistungswerten bestehen.

### 17.2.2. Operationen auf dem Suchraum

Sind Suchraummodelle mit zugehörigem Dekoder erstellt worden, so sind hierauf weitere als die bisher betrachteten Funktionen denkbar, die im Rahmen dieser Arbeit nicht mehr eingehend untersucht werden konnten. Einige hiervon sollen in diesem Abschnitt als mögliche Erweiterung kurz aufgeworfen werden.

**Teilweise Umplanung** In bestimmten Szenarien kann es vorkommen, dass eine Anlage lediglich einen Teil ihres bei der initialen Planung zugewiesenen Fahrplans neu planen muss. Die ist z. B. der Fall wenn eine Anlage über den Tag verteilt an mehreren sich zeitlich nicht überschneidenden (Block-) Stromprodukten (vgl. z. B. <http://www.epexspot.com/en/market-data/auction>) in verschiedenen Verbänden beteiligt ist. Wenn in dem Verbund, der für die Lieferung eines frühen Produktes zuständig ist eine Umplanung notwendig wird, dann hat diese Umplanung auch Auswirkungen auf die für nachfolgende Zeiträume zuständigen Verbände. Die initiale Planung eines Verbandes beruht auf Annahmen über den Startzustand, den eine Anlage zu Beginn der jeweiligen Produktlieferung haben wird. Eine Umplanung in einem früher liefernden Verbund kann aber Auswirkungen auf den Anlagenzustand haben. Bekommt eine Anlage einen anderen Fahrplan zum abfahren zugewiesen, so hat sie am Ende des Lieferzeitraums (was gleichzeitig der Beginn eines nachfolgenden Lieferzeitraums sein kann) gegebenenfalls einen anderen Anlagenzustand als ursprünglich vorgesehen. Somit kann sie eventuell den für den nachfolgenden Verbund initial geplanten Fahrplan nicht abfahren, was die Notwendigkeit der Umplanung auch des nachfolgenden Verbandes zur Folge hätte. Andere Anlagen, die nachfolgend in weiteren, anderen Verbänden beteiligt sind würden eventuell noch weitere Umplanungen erforderlich machen. Eine Umplanung zu einem frühen Zeitpunkt könnte somit eine große Zahl weiterer Umplanungen in anderen (an der ursprünglichen Ursache unschuldigen) Verbänden notwendig machen.

Um diesen Effekt zu verhindern, müsste eine Funktion auf dem Suchraum definiert sein, die im Falle einer Umplanung möglichst nur solche Fahrpläne berücksichtigt, die keine Auswirkung auf die Durchführbarkeit nachfolgender Fahrpläne haben. In einfachen Fällen kann dies indirekt erreicht werden, indem der finale Anlagenzustand, in den die Abarbeitung eines gegebenen Fahrplans führt, als Kennzahl in das Suchraummodell integriert wird. Dann kann eine große Zahl an zufällig generierten Fahrplänen daraufhin überprüft werden, ob sie in den gleichen oder zumindest einen hinreichend ähnlichen Endzustand führen, sodass die Durchführbarkeit nachfolgender Planungen als hinreichend wahrscheinlich angenommen werden kann. Dies ist aber nur für einfach kodierbare Anlagenzustände

## 17. Generalisierung des Ansatzes

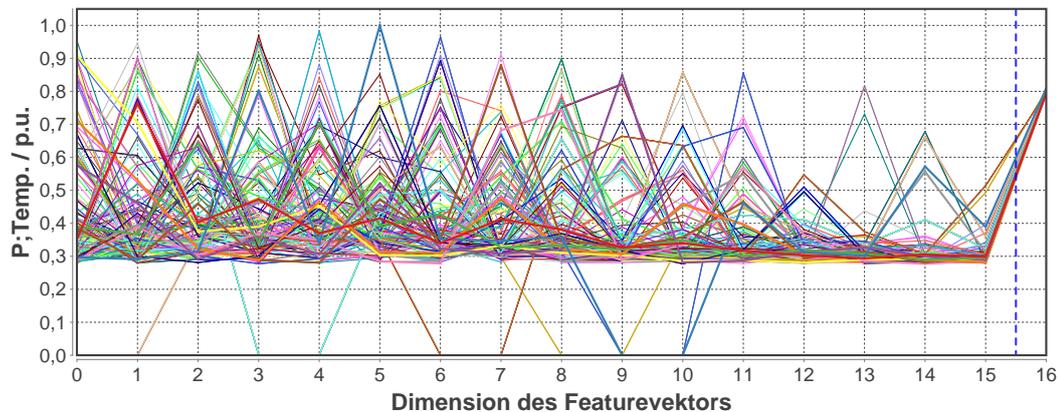


Abbildung 17.2.: Menge von Fahrplänen mit gleichem Anlagenendzustand am Beispiel eines BHKW.

möglich. Abbildung 17.2 zeigt das am Beispiel eines BHKW. Aufgeführt ist eine Auswahl an durch das BHKW realisierbaren Fahrplänen aus einem Suchraummodell, die alle zu einer ähnlichen ( $\pm 0.1$  °C) Pufferspeichertemperatur führen.

Durch Verwendung des Endzustandes als Kennzahl kann es also möglich sein, im Falle einer Umplanung diese gezielt als Optimierungsziel einzubeziehen. Wie dies unabhängig vom konkreten Typ einer Anlage geschehen könnte, ist jedoch noch eine völlig offene Fragestellung.

**Flexibilitäten** Im Rahmen dieser Arbeit bezog sich der Begriff Flexibilität einer Anlage bisher immer auf die Menge alternativer Fahrpläne, die eine Anlage für einen gegebenen Zeithorizont fahren kann. Häufig wird in Smart-Grid-Planungssituationen der Begriff Flexibilität aber auch in strukturierterer Form benötigt. Wenn es beispielsweise darum geht, im Rahmen einer Umplanung bestimmte Energiemengen zeitlich zu verschieben, so werden Alternativfahrpläne bestimmter Struktur benötigt. Der Suchraum hier besteht dann aus Fahrplänen mit der Eigenschaft, dass sie sich von einem gegebenen Referenzfahrplan nur durch zwei komplementäre Abweichungen zu zwei verschiedenen Zeitpunkten unterscheiden. Eine geringere Erzeugung zu einem Zeitpunkt  $t_1$  wird durch eine betragsgleiche Mehrererzeugung zum Zeitpunkt  $t_2 > t_1$  ausgeglichen, sodass die Gesamterzeugung gleich bleibt und lediglich eine bestimmte Energiemenge zeitlich verschoben wird; oder umgekehrt. Um derartige Suchraummodelle für Flexibilitäten zu erzeugen kämen prinzipiell zwei Ansätze in Betracht.

Zum einen könnte aus einem bereits existierenden Suchraummodell (so wie in dieser Arbeit beschrieben) eine Trainingsmenge für ein neues Suchraummodell extrahiert werden indem systematisch einige Fahrpläne generiert werden die sich von einem gegebenen Referenzfahrplan (der den die Anlage aktuell fährt) genau in zwei Zeitintervallen durch betragsgleiche aber mit unterschiedlichem Vorzeichen behaftete Leistungswerte unterscheiden. Das Klassifikatormodell des bereits existierenden Suchraummodells könnte als Entscheidungsfunktion dienen, welche dieser Fahrpläne von der Anlage tatsächlich auch umsetzbar sind. Mit dieser neuen Trainingsmenge könnte ein weiteres Suchraummodell konstruiert werden, das nunmehr nur noch Fahrpläne enthält, die jeweils die Verschiebung

einer bestimmten Energiemenge repräsentieren. Der Dekoder würde somit zulässige Verschiebungen erzeugen. Eine zunächst naheliegendere Alternative wäre daher, die Konstruktion einer entsprechenden Trainingsmenge direkt mit dem Anlagensimulationsmodell vorzunehmen. Entsprechende Untersuchungen wurden bis dato aber noch nicht vorgenommen.

**Suchräume von Anlagenensembles** Eine weitere Erweiterungsmöglichkeit könnte die Ausdehnung des Suchraums auf mehrere Anlagen darstellen. Dies könnte immer dann zu einem Anwendungsfall werden, wenn beispielsweise die Gesamtheit der Anlagen in einem Verteilnetz (oder einem anderen Unterabschnitt) bzw. ihre (unter Berücksichtigung von Netzrestriktionen) zulässige kumulierte Erzeugung betrachtet werden soll. Dies ist beispielsweise bei der Abwägung verschiedener Energiemixoptionen der Fall wie sie z. B. aktuell im Projekt DFlex (<http://www.offis.de/struktur/projekte/d-flex.html>) betrachtet werden. Aber auch beim gemeinschaftlichen Management einer Menge von Anlagen (ein Haushalt, eine Produktionsstraße o. ä. ) durch einen einzelnen Agenten muss der gemeinschaftliche Suchraum aller Anlagen aufgestellt werden.

Diese Erweiterungsmöglichkeit stellt gleichzeitig aber auch ein Beispiel dar, welches bereits Grenzen des derzeitigen Ansatzes aufzeigt. Es scheint zunächst logisch, die möglichen Fahrpläne für das Ensemble von Anlagen aus den Fahrplänen der einzelnen Anlagen zusammensetzen (vermutlich existieren in einem solchen System später auch nur Einzelmodelle der Anlagen). Verschiedene Anlagenfahrpläne werden addiert und sofern nicht globale Nebenbedingungen (z. B. Netzrestriktionen) gegen den so entstandenen Einsatzplan sprechen zur Trainingsmenge für das Ensemble von Anlagen hinzugefügt und ansonsten verworfen. Mit der so entstandenen Ensembletrainingsmenge werden wie gehabt Suchraummodell und Dekoder trainiert. Dieses Vorgehen weist aber zwei Schwachstellen auf. Zum einen wächst die Zahl an Kombinationsmöglichkeiten der Trainingsfahrpläne exponentiell an. Hier könnte anstatt alle Kombinationsmöglichkeiten zu betrachten, vielleicht eine zufällige Auswahl betrachtet werden. Es ist aber wohl dennoch zu erwarten, dass die Trainingsmenge für ein Ensemble von Anlagen sehr groß sein müsste. Zum anderen ergibt sich ein Problem mit der statistischen Verteilung der neuen Trainingsmenge.

Die Verteilung der Summe von unabhängigen Zufallsvariablen ergibt sich durch die Faltung der Verteilungen der einzelnen Zufallsvariablen. Zwar sind die möglichen, umsetzbaren Fahrpläne verschiedener Anlagen in einem Ensemble nicht zwangsläufig unabhängig (Korrelation im Verhalten z. B. durch ähnliche Witterungsbedingungen), das Sampling zum Erstellen der einzelnen Trainingsmengen basiert jedoch auf voneinander unabhängigen Zufallsreihen. Unter der Annahme, dass die Samples in den einzelnen Trainingsmengen in jedem Zeitintervall gleichverteilt sind, ergäbe sich für die Ensembletrainingsmenge als Summe der Einzeltrainingsmengen eine multimodale Irvin-Hall-Verteilung (siehe z. B. Hall 1927). Approximativ kann diese durch eine Normalverteilung (je Zeitintervall) beschrieben werden. Was dies exemplarisch für die Summe von Trainingsmengen von mehreren BHKW (die alle unabhängig von einander arbeiten können) bedeutet zeigt Abbildung 17.3.

Das Problem, das hierdurch entsteht, liegt in der im Vergleich zum Rand sehr großen Dichte der Trainingspunkte in mittleren Bereich der zulässigen Fahrpläne. Das hat zur Folge, dass von dem SVDD-Verfahren die zu dünn besiedelten Randbereiche als Ausreißer

## 17. Generalisierung des Ansatzes

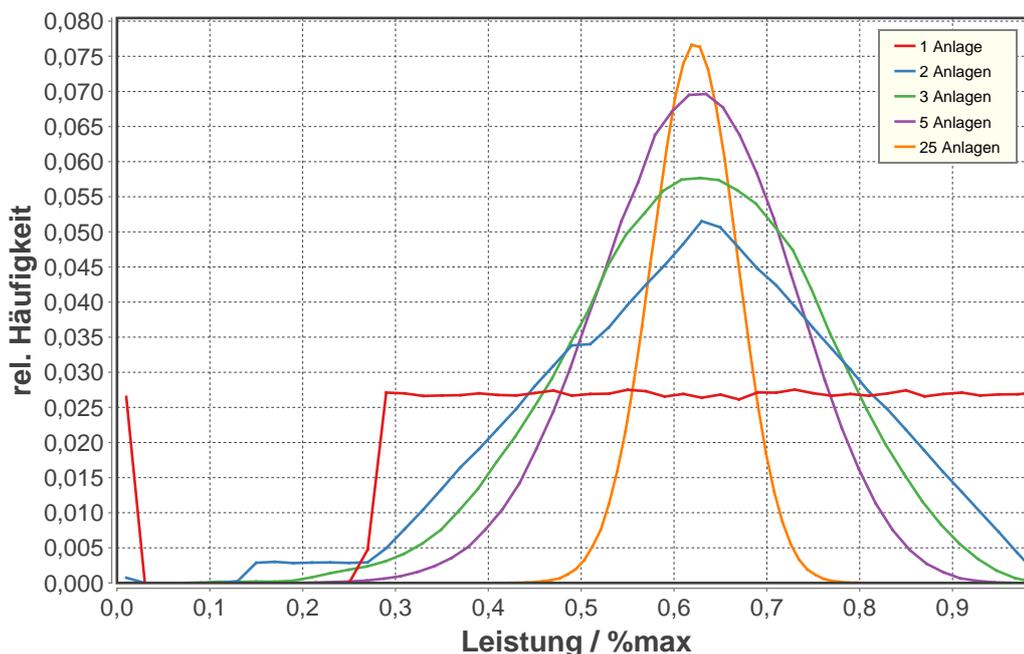


Abbildung 17.3.: Verteilung kombinierter BHKW-Trainingsmengen für verschiedene Anlagenzahlen. Die kombinierten Trainingsmengen sind additiv aus Einzeltrainingsmengen zusammengesetzt.

ignoriert werden. Somit wird ein Modell gelernt, dass nur einen relativ kleinen Bereich aus der Mitte des tatsächlich zulässigen Bereichs umfasst. Dieser Effekt ist umso stärker ausgeprägt je mehr Anlagen zusammengefasst werden. Abbildung 17.4 zeigt die Auswirkungen an einem Beispiel mit lediglich zwei BHKW.

Das Beispiel zeigt auf der linken Seite in blau eine Menge von Fahrplänen (für zwei Zeitintervalle) die aus separat für zwei BHKW erstellten Trainingsmengen erzeugt wurde indem zufällig aus diesen zwei Trainingsmengen je ein Fahrplan ausgewählt wurde. Zwei zufällig gewählte Fahrpläne wurden jeweils zu einem gemeinschaftlichen Fahrplan addiert. Der rechte Teil 17.4(b) zeigt die gleiche Trainingsmenge sowie anhand der Farbintensität der Graustufen ihre Dichte. Die rote Linie kennzeichnet die gelernte Grenze des zulässigen Bereiches für gemeinschaftliche Fahrpläne. Wie erwartet umfasst das Modell lediglich eine zu kleine Menge aus dem dichteren Bereich. Im Falle mehrere Anlagen wäre der Effekt noch stärker.

Zudem könnte ein hiermit trainierter Dekoder nur Gesamtfahrpläne für ein Ensemble generieren ohne dass die genaue Aufteilung auf die einzelnen Anlagen bekannt wäre. Es wäre lediglich sichergestellt, dass ein Gesamtfahrplan von dem Ensemble gefahren werden kann, eine mögliche Partitionierung in Einzelfahrpläne müsste aber immer noch bestimmt werden. Dies wäre ein eigenes Wirkleistungsplanungsproblem für jedes Ensemble.

Der in dieser Arbeit präsentierte Ansatz ist also nicht direkt auf Anlagenensembles anwendbar. Hier ist im Nachgang noch weitere Forschung erforderlich.

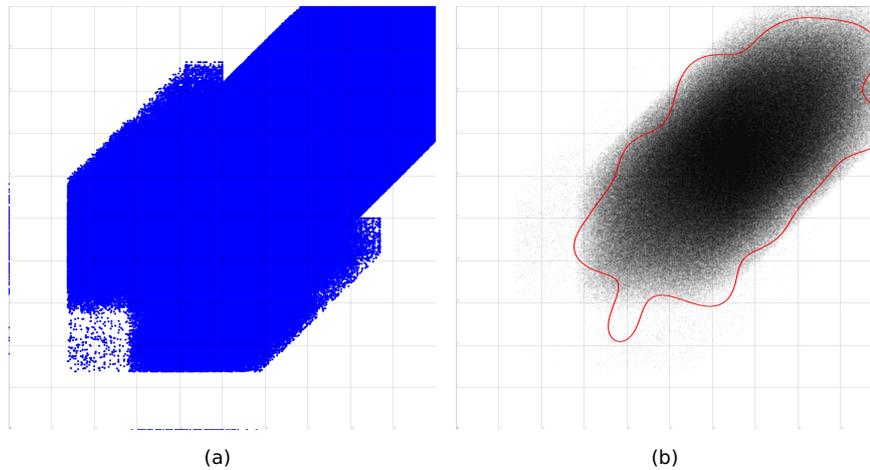


Abbildung 17.4.: Beispiel für kombinierte Suchräume zweier BHKW. Die blaue Fläche links zeigt näherungsweise den tatsächlich zulässigen Bereich der kombinierten Fahrpläne beider Anlagen. Da jedoch die Trainingsmenge einer Irvin-Hall-Verteilung unterliegt (rechts durch halbtransparente Darstellung der Fahrpläne an der Farbintensität zu erkennen), wird vom SVDD-Verfahren nur der innere Teil als Modell (rote Linie) erfasst.

### 17.2.3. Verbessertes Sampling

Potenzial zur Verbesserung dürfte auch das Sampling für die Erstellung der Trainingsmengen bieten. Jede Trainingsmenge diskretisiert aufgrund der beschränkten Anzahl an Trainingsfahrplänen den realen, kontinuierlichen Suchraum immer nur unvollständig und (zumindest bei höherdimensionalen Fahrplänen) eher grob.

Neben Verbesserungen des Samplings selbst bzgl. Geschwindigkeit, Verteilung u.ä. sei an dieser Stelle vor allem folgende Idee noch aufgeführt: die des produktspezifischen Samplings. Beim Erstellen der Trainingsmenge ist grundsätzlich immer ein Kompromiss einzugehen aufgrund von Berechnungsperformanz und Speicherplatzrestriktionen. Die Trainingsmenge wird daher den realen Suchraum immer nur in Teilen reflektieren können. Wenn aber a priori Informationen über die Struktur des Stromproduktes, das es nachher zu liefern gilt bekannt sind, so liegt es vielleicht nahe, diese bei der Erstellung der Trainingsmenge bereits zu nutzen.

Ist beispielsweise vorab bekannt, dass das Produkt auf das geboten werden soll ein Blockprodukt mit möglichst konstantem Leistungsverlauf ist, wie es heute an der Strombörse noch üblich ist, so kann bereits beim Sampling darauf geachtet werden möglichst nur solche Fahrpläne in die Trainingsmenge aufzunehmen, die einen möglichst flachen Verlauf haben.

Abbildung 17.5 auf der nächsten Seite zeigt ein Beispiel für ein über 4 Stunden konstantes Wirkleistungsprodukt für 5 BHKW. Das Sampling der BHKW wurde so verändert, dass passend zum konstanten Produkt Fahrpläne mit möglichst flachen (allerdings nicht konstanten) Lastverläufen erzeugt wurden. Hiermit wurden die Suchraummodelle trainiert und die Optimierung durchgeführt. Im Ergebnis wird ein Einsatzplan gefunden, der sich

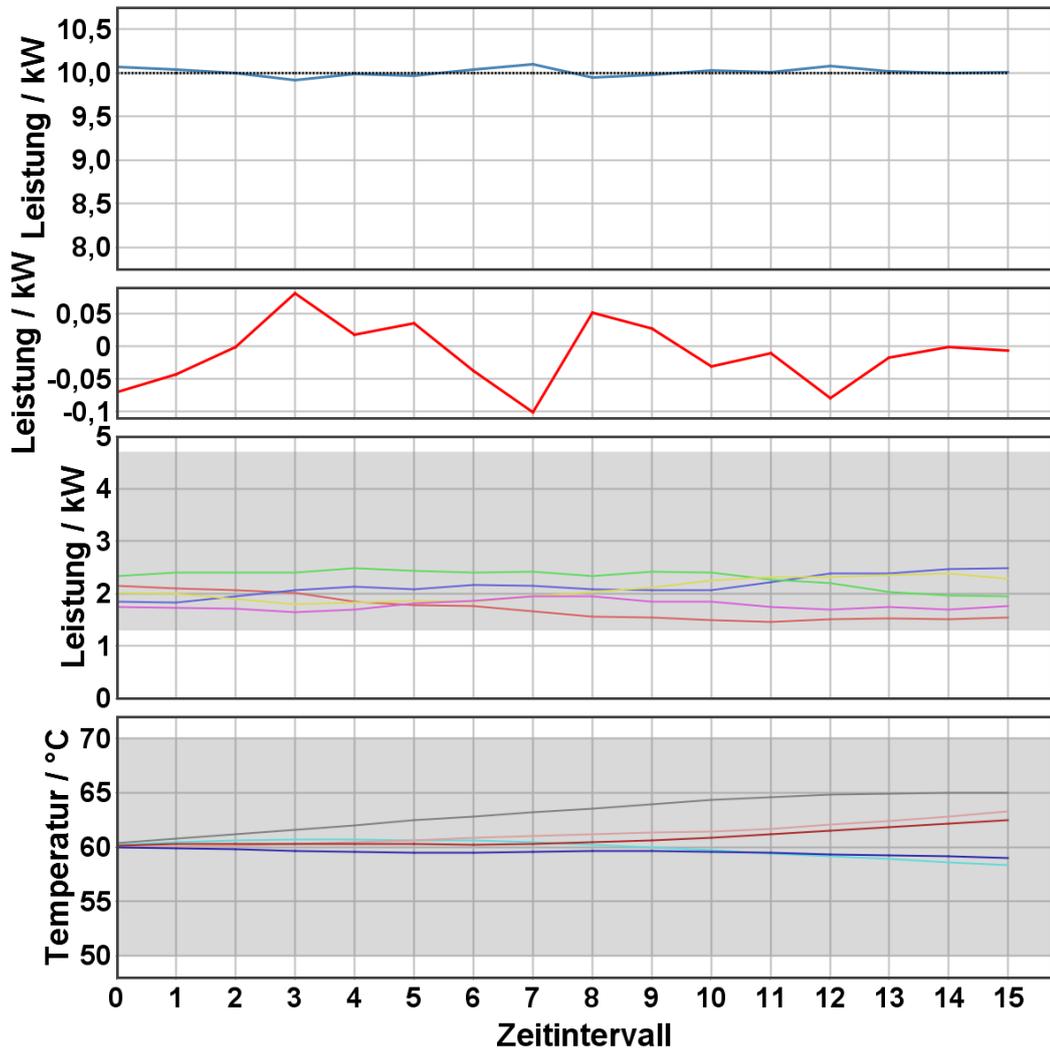


Abbildung 17.5.: Beispiel eines Optimierungsergebnisses erzielt mit produktspezifischen Trainingsmengen. Dargestellt ist v.o.n.u. Ziellastgang und Lastverlauf des Einsatzplans, der verbleibende Restfehler, die verwendeten Einzelfahrpläne und die resultierenden Temperaturverläufe im BHKW-Pufferspeicher (jeweils mit zulässigem Bereich).

deutlich besser an den konstanten Ziellastgang annähert als dies mit einem Suchraum aus vorwiegend stark in der Leistung schwankenden Fahrplänen der Fall ist.

Allein durch das produktspezifische Sampling lässt sich der verbleibende Restfehler in dem Beispiel in 17.5 auf etwa  $4.234 \pm 1.412$  Prozent des ursprünglichen Fehlers reduzieren (durch Optimierung mit Modellen, die mit einer nicht speziell angepassten Trainingsmenge erstellt wurden). Umgekehrt ist aber die flache Trainingsmenge nicht gut geeignet für die Erzielung schwankender Zielvorgaben.

Liegt ein Vorwissen über die Struktur der gewünschten Ziellastgänge vor, so ist es aber hilfreich, dieses Wissen bei der Erstellung der Trainingsmengen bereits einfließen zu lassen.

## 17.3. Generalisierung des Ansatzes

Was noch bleibt ist die Frage, ob sich der Ansatz des Supportvektro-Dekoders auch auf andere Anwendungsdomänen ausweiten lässt als das bisher ausschließlich betrachtete Smart Grid. Generell ist das Verfahren überall dort sinnvoll einsetzbar, wo Lösungsräume einen dichten Unterraum eines durch Constraints eingeschränkten Vektorraums bilden. Lassen sich anhand geeigneter Modelle Trainingsmengen erzeugen mit dem das SVDD-Modell gelernt werden kann, so kann hierauf aufbauend auch ein Dekoder zur systematischen Erzeugung gültiger Lösungen aus dem Modell abgeleitet werden. Supportvektor-Methoden haben sich in der Vergangenheit bereits als geeignet erwiesen für eine Vielzahl vielfältigster Problemstellungen sowohl für kontinuierliche als auch diskrete Daten, da verschiedene Kerne verwendet werden können (z. B. Ruiz et al. 2006). Dabei ist der Einsatz des Dekoders keineswegs auf verteilte Planungs- oder Optimierungsprobleme beschränkt.

Zwei Beispiele für weitere Anwendungsfelder könnten sein:

**Energieoptimale Wasserversorgung:** Hier geht es darum, für eine Menge von Pumpen (und gegebenenfalls anderen Betriebsmitteln) in einem Wasserversorgungsnetz Fahrpläne zu finden, die einerseits den verteilt benötigten Wasserbedarf decken und andererseits bezüglich der benötigten Energiekosten optimiert sind. Das Optimierungsproblem ist beispielsweise beschrieben in (Costa et al. 2010). Zu beachten sind hier in Analogie zur Wirkleistungsplanung Constraints welche sowohl das Leistungsvermögen der Pumpen als auch das Fassungsvermögen diverser Zwischenspeicher betreffen. Aber auch Constraints bezüglich der Aufrechterhaltung des lokalen hydraulischen Gleichgewichts, welche nicht in expliziter Form gegeben sind sondern während der Optimierung fortwährend anhand von Simulationen des Systems ermittelt werden müssen, sind zu beachten (vgl. Costa et al. 2010). Eine recht ähnliche Fragestellung ergibt sich bei der Steuerung von Gasnetzen.

**Dynamische Verkehrsflusssteuerung:** Bei der verteilten Verkehrsflusssteuerung – wie beispielsweise in dem White Paper von Masterson & Topiwala (2008) beschrieben – geht es darum, über einen größeren Bezirk verteilte Signalanlagen so zu steuern, dass der gesamte Durchsatz des Straßenverkehrs maximiert bzw. die durchschnittliche Wartezeit für die Fahrzeuge minimiert wird. Naturgemäß sind hierbei für einzelne Signalanlagen lokale Constraints für individuelle Signalabfolgen zu beachten. Eine Formulierung als verteiltes Constraint Satisfaction Problem findet sich u. a. in (Mizuno et al. 2008). Als Besonderheit können bei der verteilten Verkehrsflusssteuerung auch verschiedene Typen von Steuerungsanlagen (z. B. Ampeln, variable Geschwindigkeitsbeschränkungen o. ä. ) beteiligt sein. Auch der Wegfall (im Falle einer Störung) oder das Hinzukommen neuer Anlagen (automatische Leitsysteme sind gegebenenfalls nicht immer in Betrieb) ist bei dieser Problemstellung möglich.

## 17.4. Grenzen des Ansatzes

Für den in dieser Arbeit konzipierten Supportvektor Ansatz wurden nunmehr bereits diverse Anwendungsfelder und Erweiterungsmöglichkeiten aufgezeigt. Schließlich sol-

## 17. Generalisierung des Ansatzes

---

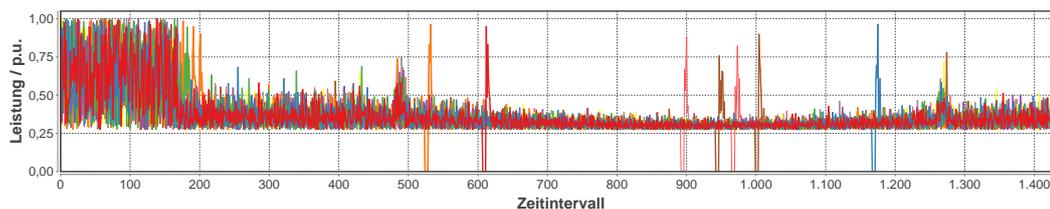


Abbildung 17.6.: Beispielfahrpläne für ein BHKW für einen ganzen Tag in Minutenauflösung. Die Fahrpläne wurden zufällig mit einem Dekoder generiert.

len an dieser Stelle aber auch noch die Grenzen der Einsetzbarkeit angesprochen bzw. rekapituliert werden.

**Auflösung und Dimensionalität** Das Verfahren so wie es in dieser Arbeit vorgestellt wurde ist zwar prinzipiell unabhängig von der verwendeten zeitlichen Auflösung, jedoch gibt es in der Praxis Grenzen. Abbildung 17.6 zeigt ein Beispiel für Ganztagesfahrpläne in Minutenauflösung, die zufällig mithilfe eines Dekoders erzeugt wurden. Das Lernen eines Modells für derart hochaufgelöste Fahrpläne (1440-dimensionale Vektoren) scheint zunächst möglich und der entsprechende Dekoder generiert auch zulässige Ganztagesfahrpläne. Betrachtet man jedoch die Werte der Entscheidungsfunktion (vgl. Abschnitt 5.3 auf Seite 60) so fällt ein erstes Problem auf. Während sich der Abstand vom Zentrum der Hypersphäre bei 96-dimensionalen Fahrplänen noch etwa ab der achten Nachkommastelle unterscheidet, ist bei den 1440-dimensionalen Fahrplänen auf einem System mit doppelt genauen Fließkommawerten (12 Stellen) kein Unterschied zwischen den verschiedenen Fahrplänen festzustellen. Eine Klassifikation ist so nicht möglich; zumindest nicht mit dieser Zahlendarstellung. Dies ist ein generelles Problem – auch bekannt als Fluch der Dimensionalität – der unter anderem dazu führt, dass mit wachsender Dimension der Daten beliebige Punkte paarweise immer denselben Abstand haben (Köppen 2000; Evangelista et al. 2006). Zum anderen wächst die Zahl der benötigten Trainingsvektoren und damit auch die notwendige Berechnungszeit stark an. Dies gilt insbesondere auch für die Dekoderberechnungen. An dem Beispiel in Abbildung 17.6 sind z. B. etwa 1200 Supportvektoren beteiligt, deren Zahl quadratisch in den Berechnungsaufwand für die Abbildung auf gültige Fahrpläne eingeht.

Ein anderes Problem tritt bei dem Versuch auf, einen längeren Zeitraum (z. B. mehrere Tage) in ein einzelnes Suchraummodell zu integrieren. Hier stellt sich die Frage, ob die Prognosemodelle beispielsweise zum Wärmebedarf, zum Wetter o. ä. hinreichend gut sind oder ob zu erwarten ist, dass ein solches Suchraummodell ohnehin häufiger (bei Vorliegen neuer Prognosen) durch ein neues Modell ersetzt werden müsste.

Eine genaue, quantitative Untersuchung zur möglichen Problemdimension wurde im Rahmen dieser Arbeit nicht durchgeführt. Welche Vorhersagezeiträume sinnvoll und mit welcher zeitlichen Auflösung abgedeckt werden können hängt zum einen von den konkreten Anforderungen (Genauigkeit, Qualität der Simulationsmodelle, usw.) der jeweiligen Problemstellung ab; zum anderen aber auch stark von den Fähigkeiten der verwendeten Hard- und Software.

**Genauigkeit** Wie bereits in Kapitel 9 ab Seite 117 ausgeführt kommen an verschiedenen Stellen der Prozesskette Näherungsverfahren zum Einsatz. Alle Einsatzbereiche in denen eine hochgenaue, exakte Lösung benötigt wird, sind daher vorab auf Anwendbarkeit zu prüfen. Die erreichte Genauigkeit (mit Ungenauigkeiten im Bereich weniger Watt bei mehreren kW Nennleistung) stellt für das in dieser Arbeit betrachtete Wirkleistungsplanungsproblem kein Problem dar. In anderen Anwendungsbereichen ist gegebenenfalls Vorsorge durch nachgeschaltete Fehlerkorrekturen zu treffen. Anwendungsbereiche der mathematischen Optimierung mit benötigten hohen Genauigkeiten können das hier vorgeschlagene Verfahren vermutlich nur zur Generierung einer ersten Schätzung als Ausgangspunkt für die exakte Suche nach dem Optimum einbinden. Dies ist aber auch nicht der vordringlich angestrebte Anwendungsfall.

**Hardwareanforderungen** Eine systematische Erhebung der Anforderungen an die spätere Hardware wurde im Rahmen dieser Arbeit nicht durchgeführt. Dennoch lassen sich bereits einige Aussagen treffen. Die technische Ansteuerung der Anlage über vorhandene Regelungssysteme kann bei ersten Systemen bereits als gegeben und somit (nach Anpassungen) zumindest als machbar angesehen werden, da die technischen Voraussetzungen vorhanden sind (vgl. insbesondere auch Hinrichs 2014). Wesentlich interessanter für den in dieser Arbeit betrachteten Ansatz ist die notwendige Rechenleistung, über die die Anlagengenhardware verfügen müsste. Eine allgemeingültige Aussage ist hier nicht zu treffen, da dies stark vom tatsächlich umgesetzten Szenario abhängt. Hohe Genauigkeit verlangt große Trainingsmengen und damit mehr Speicherbedarf und Rechenleistung für die Bestimmung der Supportvektoren usw. Für Szenarien mit eher niedrigeren Genauigkeitsanforderungen wurden exemplarisch die in Tabelle 17.1 gelisteten Berechnungszeiten auf einem Desktoprechner (Intel i3 3.1 GHz, Java 64 Bit) ermittelt. Tabelle 17.2 zeigt einige Zeiten für die Benutzung des Dekoders während der Optimierung.

Tabelle 17.1.: Verschiedene Rechenzeiten für die Dekodererstellung für Fahrpläne über  $d$  Zeitintervalle und verschiedene Trainingsmengengrößen. Alle Angaben sind in  $ms$  CPU-Zeit für Sampling (inkl. Kerndichteschätzung), Modelltraining und Dekodererstellung.

$d$	50	100	500
16	$38.12 \pm 8.06$	$229.43 \pm 8.26$	$12010.97 \pm 1836.14$
32	$48.14 \pm 5.64$	$253.44 \pm 10.28$	$16814.86 \pm 1690.53$
96	$89.13 \pm 8.03$	$335.71 \pm 12.85$	$26047.44 \pm 3277.21$

Tabelle 17.2.: Verschiedene Rechenzeiten für die Benutzung vom Dekoder für Fahrpläne über  $d$  Zeitintervalle erstellt aus verschiedenen Trainingsmengengrößen. Alle Angaben sind in  $ms$  CPU-Zeit für das Abbilden eines Fahrplans auf einen gültigen Fahrplan.

$d$	50	100	500
16	$0.196 \pm 0.128$	$0.432 \pm 0.283$	$1.103 \pm 0.821$
32	$0.261 \pm 0.182$	$0.536 \pm 0.373$	$1.523 \pm 1.282$
96	$0.425 \pm 0.303$	$0.945 \pm 0.689$	$4.814 \pm 6.264$

Auf einem eingebetteten System sind derartige Zeiten wohl nicht zu erreichen; ein Testsystem stand nicht zur Verfügung. Ansätze den Grafikprozessor zur Beschleunigung der Berechnungen einzusetzen (z. B. Liao et al. 2009) funktionieren auf einem eingebette-

ten System mangels Hardware ebenfalls nicht. Andererseits existieren aber bereits erste Forschungsarbeiten mit dem Ziel Supportvektor Maschinen vollständig in Hardware umzusetzen. Beispiele hierfür sind u. a. gegeben in (Papadonikolakis & Bouganis 2010; Ruiz-Llata et al. 2010). Die Integration derartiger Hardware in gegebene Anlagensteuerungssysteme ist aber wohl derzeit nicht ohne weiteres möglich.

**Anlagenübergreifende Constraints** Constraints, die mehr als eine Anlage betreffen sind mit dem Ansatz, wie er hier vorgestellt wurde zunächst nicht eins zu eins umsetzbar. Als Beispiel hierfür mag die Minimierung des Gleichzeitigkeitsfaktors dienen. Bei der Minimierung des Gleichzeitigkeitsfaktors wird versucht, bei der Erbringung einer gegebenen Energiemenge in einem Zeitintervall, die Leistung möglichst gleichmäßig über das Intervall zu verteilen. Es ist nicht wünschenswert, dass eine große Zahl an Anlagen gleichzeitig ein- bzw. ausschaltet, wodurch große Leistungsgradienten verursacht werden.

Die Gleichzeitigkeit von Schaltereignissen in den Fahrplänen verschiedener Anlagen kann aber nicht im Suchraum der einzelnen Anlagen kodiert werden. Hierzu müssten die Anlagen im Ensemble modelliert werden; mit den bereits angesprochenen (wenngleich lösbaren) Problemen.

### 17.5. Zusammenfassung

Mit den Betrachtungen in diesem Kapitel schließt nunmehr der inhaltliche Teil dieser Arbeit. Die aufgezeigten Erweiterungsmöglichkeiten machen deutlich, dass durchaus noch Raum für nachfolgende Arbeiten geblieben ist; sowohl für die Ausdehnung auf weitere Anwendungsdomänen als auch für weitere Aufgaben im Smart Grid Umfeld selbst. Aber auch die Integration in weitere Optimierungsansätze und deren Parametrierung lässt weiteren Forschungsbedarf vermuten.

Wie häufig der Fall, sind auch dem in dieser Arbeit aufgezeigten Konzept in der praktischen Anwendung Grenzen gesetzt. Soweit bekannt, wurde versucht auch diese aufzuzeigen. Was noch bleibt, ist ein abschließendes Resümee mit dem diese Arbeit im nächsten Kapitel schließen wird.

# 18. Fazit und Ausblick

*Difficult to see.  
Always in motion is the future.*

---

*(Yoda)*

## 18.1. Zusammenfassung

Ziel aktueller politischer Entscheidungen ist die schrittweise Transformation einer bisher eher zentralistisch orientierten elektrischen Energieversorgung durch wenige Großkraftwerke hin zu einem dezentralen Aufbau. Bedingt durch die dringend notwendige Integration einer Vielzahl kleiner, schlecht prognostizier- und steuerbarer regenerativer Erzeuger wird auch ein Wandel bei den Planungs- und Steuerungsalgorithmen zu beobachten sein. Kleine Anlagen werden sich situationsbedingt zu Verbänden aggregieren, um gemeinsam hinreichend Aktionsmöglichkeiten zu haben mit denen sie eine Aufgabe der Netzsteuerung verantwortlich übernehmen können. Planungsalgorithmen werden in der Lage sein müssen, ad hoc mit ihnen bisher unbekanntem Anlagen zusammenzuarbeiten, wenn sie Einsatzpläne für Verbände erstellen und somit für jede Anlage einen nach gegebenen Kriterien optimalen Fahrplan bestimmen. Wichtig für die verantwortliche Übernahme ist die Sicherstellung der Umsetzbarkeit der gefundenen Lösungen. Hierfür ist es aber notwendig, dass alle Anlagen die Information darüber zur Verfügung stellen, welche Fahrpläne für einen gegebenen Zeithorizont realisierbar sind ohne Restriktionen zu verletzen. Für eine systematische Planung muss der entsprechende Algorithmus jedoch nicht nur wissen, wo im Suchraum die gültigen Lösungen zu finden sind, es bedarf zusätzlich auch einer Bewertungsmöglichkeit hinsichtlich (je nach Anwendungsfall) diverser Optimierungsziele.

In dieser Arbeit wurde ein Ansatz entwickelt zur Modellierung und zur Integration von Constraints in bestehende Optimierungsansätze. Auch wenn im Rahmen des zur Evaluation verwendeten Anwendungsfalls der Wirkleistungsplanung im Smart Grid beinahe alle Teilaspekte anwendungsspezifisch umgesetzt wurden, ist das zugrundeliegende Konzept dennoch problemoffen und auf andere Domänen leicht anpassbar. Eine Reihe von spezifischen Anforderungen wurde eingangs identifiziert an denen sich der entwickelte Supportvektor-Dekoder messen musste.

Ziel der vorliegenden Arbeit war es, eine abstrakte Beschreibung von Suchräumen (hier zunächst als Menge realisierbarer Fahrpläne einer elektrischen Anlage) bzw. von beschränkenden Constraints zu erarbeiten, die eine abstrakte Formulierung des Wirkleistungsplanungsproblems ermöglicht, bei der nicht a priori die Menge der Anlagen in einem Verbund (und damit auch nicht deren Modellierungen für die Planung) bekannt sein muss. Es sollte möglich sein, die zulässigen Handlungsalternativen verschiedener Akteure in einer Gruppe auf abstrakter Ebene so zu erfassen, dass diese ad hoc und insbesondere

automatisiert zu einem Optimierungsmodell für ein kombinatorisches Problem integriert werden können, bei dem verschiedene Handlungsoptionen der einzelnen Akteure sich zu einer gewünschten Gesamthandlungsweise der Gruppe als Ganzes ergänzen.

Zu diesem Zweck wurde eine Prozesskette erarbeitet, die eine Modellierung der Suchräume auf abstrakter, geometrischer Ebene vornimmt und ein System bereitstellt, welches eine systematische Bewegung in diesem Suchraum ermöglicht. Zentrales Element ist ein geometrisches Modell der individuellen Suchräume basierend auf einem 1-Klassen-Klassifikator. Realisierbare Fahrpläne können somit unabhängig vom Anlagentyp erfasst und angesprochen werden. Ein integrierendes Suchverfahren muss die zugrundeliegende Systematik (Modellierung sowie einschränkende Constraints) nicht kennen, da aufgrund des abstrakten Modells alle Anlagentypen gleich behandelt werden können. Die Integration kann automatisiert über einen algorithmisch ableitbaren Dekoder erfolgen, der eine einheitliche Navigation innerhalb der zulässigen Suchräume erlaubt.

Für die Modellierung der Constraints mittels Supportvektor-Methoden konnte anhand von verschiedenen Anlagenmodellen gezeigt werden, dass die geforderte Abstraktion der Modellierung mit dem vorgeschlagenen Machine Learning Ansatz erreicht werden kann. Fahrplanmengen verschiedener Anlagen können unabhängig von konkreter Anlagenmodellierung, gegebenen Constraints, Anlagenselbstbeschreibung, aktuellem Anlagenzustand o. ä. einheitlich beschrieben werden allein durch Mengen von umsetzbaren Fahrplänen und einem hierauf trainierten Supportvektor-Verfahren.

Die Evaluation des zugehörigen Supportvektor-Dekoders anhand von verschiedenen Simulationsmodellen konnte zeigen, dass ein systematisches Durchsuchen des Fahrplanraums beliebiger Anlagen möglich ist, auch ohne die konkrete Anlagenmodellierung zu kennen. Hiermit wird die Transformation des durch Constraints beschränkten Optimierungsproblems zu einem unbeschränkten Problem möglich. Verschiedene Typen von Anlagen können einheitlich angesprochen und gleich behandelt werden ohne Kenntnis eines Anlagenmodells oder einer Anlagenbeschreibung.

Mit dem Supportvektor-Dekoder steht nunmehr ein universeller Ansatz zur Verfügung, für die ad hoc Integration der Suchräume verschiedener Anlagen zu einem gemeinsamen Optimierungsmodell. Die abstrakte Beschreibung der Suchräume bzw. der beschränkenden Constraints ermöglicht es nunmehr, eine allgemeingültige Formulierung des Wirkleistungsproblems umzusetzen, bei der nicht bereits a priori die Menge der Anlagen in einem Verbund bekannt sein muss. Damit müssen auch deren Modellierungen vorab nicht bekannt sein. Somit können nun auch sich dynamisch zusammenfindende Verbünde vorab unbekannte Anlagen (auch unbekanntem Typs) aufnehmen und durch automatisierte Ableitung eines Gesamtoptimierungsmodells in die Planung integrieren.

Abbildung 18.1 fasst noch einmal die Ergebnisse dieser Arbeit zusammen. Ausgehend vom zentralen Konzept des Dekoders ermöglicht die Formulierung als constraintfreies Problem die einheitliche Verwendung verschiedener Heuristiken, ohne dass für diese eine Anpassung auf die eingeschränkten Suchräume der verschiedenen Anlagen notwendig wird. Aber auch in dezentralen Verfahren kann der Dekoder zur Generierung zulässiger Lösungen eingesetzt werden. Ermöglicht wird dies durch die Darstellung der eingeschränkten Suchräume der Anlagen als Supportvektor-Modell. Das hierfür entwickelte Sampling zur Erstellung der Trainingsmenge aus Beispielfahrplänen ermöglicht zugleich sowohl die

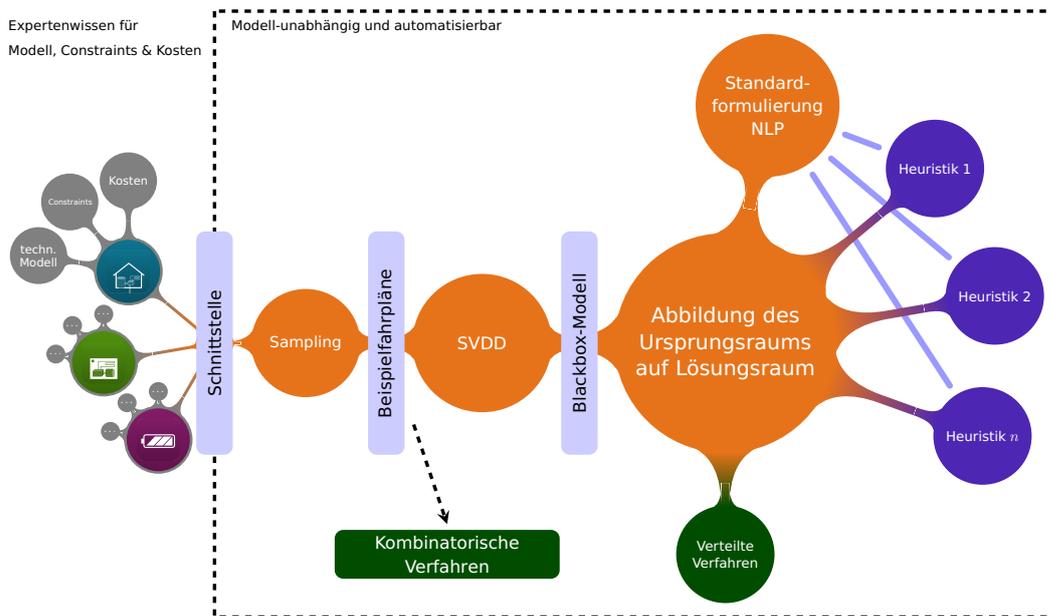


Abbildung 18.1.: Zusammenfassung der erreichten Ziele.

Darstellung beliebiger auf Fahrplänen basierender Handlungsalternativen der Anlagen als auch die Definition einer Schnittstelle zu den Anlagen auf Basis einer Verhaltenssimulation. Somit kann der gesamte Prozess – wie er in dieser Arbeit vorgeschlagen wurde – unabhängig von jedweder konkreter Anlagenmodellierung und ohne Kenntnis der Constraints arbeiten. Der Prozess ist somit automatisierbar und kann sich für neue Optimierungsfragestellungen mit bisher unbekanntem Anlagentypen selbst konfigurieren. Expertenwissen muss für jeden Anlagentyp nur einmal vorab bei der Erstellung eines Simulationsmodells einfließen.

## 18.2. Fazit

Mit dieser Zusammenfassung endet nunmehr diese Arbeit. Die eingangs gesteckten Ziele konnten erreicht werden. Der Forschungsverbund Smart Nord (<http://smartnord.de>) bot und bietet weiterhin eine ideale Testumgebung für die Ergebnisse dieser Arbeit. Die erarbeiteten Konzepte aber insbesondere auch die umgesetzten Implementierungen sind hierüber auch einer Nachnutzung zugänglich. Verbleibender Forschungsbedarf für zukünftige Arbeiten konnte identifiziert und aufgezeigt werden. Insbesondere im Anwendungsfeld der Wirkleistungsplanung in dynamischen Anlagenverbänden stellen die faire Gewinnverteilung und die systematische Integration von Prognoseunsicherheiten spannende Forschungsfragen dar. Für das entwickelte Verfahren an sich befinden sich Erweiterungsmöglichkeiten auf aggregierte Verbände oder Teilnetze in aufsetzender Arbeit in der Erforschung.

Der Bedarf an intelligenten, selbsttätig arbeitenden Methoden zur Integration einer stetig wachsenden Zahl dezentraler, schwer prognostizierbarer Stromerzeugers in die

Planung und Steuerung der jederzeitigen Abstimmung zwischen elektrischer Erzeugung und Verbrauch wird in Zukunft weiter zunehmen. Eine Heerschar verteilter, individuell konfigurierter, für sich selbst entscheidende Anlagen muss in der Lage sein, sich selbsttätig untereinander abzustimmen. Hierfür werden Algorithmen benötigt, die in der Lage sind, die alternativen Möglichkeiten dezentraler Energieanlagen systematisch zu betrachten und zu bewerten.

Der Bedarf an neuen Methoden für die Koordination der stetig steigenden Zahl dezentraler Energieerzeuger in eine stabile und verlässliche Stromversorgung ist ohne Zweifel gegeben. Eine vorausschauende Planung der Stromerzeugung in Day-ahead Szenarien ist jedoch wertlos, wenn nicht bei der Planung bereits die Umsetzbarkeit der Einsatzpläne als Planungsergebnis sichergestellt werden kann. Ein Beitrag zur möglichen Sicherstellung der praktischen Umsetzbarkeit der Ergebnisse von Planungsalgorithmen konnte in dieser Arbeit geleistet werden.

# Literatur

- Abdel-raouf, O. & M. A. baset Metwally (Mai 2013). „Article: A Survey of Harmony Search Algorithm“. In: *International Journal of Computer Applications* 70.28. Published by Foundation of Computer Science, New York, USA, S. 17–26.
- Agarwal, P. K., S. Har-Peled, M. Sharir & Y. Wang (Sep. 2010). „Hausdorff Distance Under Translation for Points and Balls“. In: *ACM Trans. Algorithms* 6.4, 71:1–71:26. issn: 1549-6325. DOI:[10.1145/1824777.1824791](https://doi.org/10.1145/1824777.1824791).
- Ahmad, A. (Juni 2010). „Unit Commitment using Hybrid Approaches“. PhD Theses. Taxila, Pakistan: Department of electrical engineering, University of engineering und technology. <http://prp.hec.gov.pk/Thesis/910S.pdf>.
- Aizerman, M. A., E. A. Braverman & L. Rozonoer (1964). „Theoretical foundations of the potential function method in pattern recognition learning.“ In: *Automation and Remote Control*, 25, S. 821–837.
- Akella, R., F. Meng, D. Ditch, B. McMillin & M. Crow (Okt. 2010). „Distributed Power Balancing for the FREEDM System“. In: *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*. Gaithersburg, MD, USA: IEEE, S. 7–12. isbn: 978-1-4244-6510-1. DOI:[10.1109/smartgrid.2010.5622003](https://doi.org/10.1109/smartgrid.2010.5622003).
- Anders, G., F. Siefert, J.-P. Steghöfer, H. Seebach, F. Nafz & W. Reif (2010). „Structuring and Controlling Distributed Power Sources by Autonomous Virtual Power Plants“. In: *IEEE Power and Energy Student Summit (PESS 2010)*. IEEE Power & Energy Society.
- Angantyr, A., J. Andersson & J.-O. Aidanpaa (2003). „Constrained optimization based on a multiobjective evolutionary algorithm“. In: *IEEE Congress on Evolutionary Computation (3)*. IEEE, S. 1560–1567. isbn: 0-7803-7804-0.
- Aronszajn, N. (1950). „Theory of reproducing kernels“. In: *Transactions of the American Mathematical Society* 68.
- Audi AG (2013). *Energiewende im Tank*. Webseite. [http://www.audi.de/de/brand/de/vorsprung\\_durch\\_technik/content/2013/08/energiewende-im-tank.html](http://www.audi.de/de/brand/de/vorsprung_durch_technik/content/2013/08/energiewende-im-tank.html) (besucht am 29.01.2014).
- Avriel, M. (2003). *Nonlinear Programming: Analysis and Methods*. Dover Books on Computer Science Series. Dover Publications. isbn: 9780486432274. <http://books.google.de/books?id=byF4Xb1QbvMC>.
- Awerbuch, S. & A. M. Preston, (Hrsg.) (1997). *The Virtual Utility: Accounting, Technology & Competitive Aspects of the Emerging Industry*. Bd. 26. Topics in Regulatory Economics and Policy. Kluwer Academic Publishers. isbn: 0-7923-9902-1.
- Aziz, H. & B. de Keijzer (2013). „Shapley Meets Shapley“. In: *CoRR* abs/1307.0332.

- BMWi (Apr. 2012). *Zahlen und Fakten Energiedaten: Nationale und Internationale Entwicklung*. Techn. Ber. Berlin: Bundesministerium für Wirtschaft und Technologie, Referat III C 3. <http://www.bmwi.de/Navigation/Technologie-und-Energie/Energiepolitik/energiedaten.html>.
- Bach, F. R. & M. I. Jordan (März 2003). „Kernel Independent Component Analysis“. In: *J. Mach. Learn. Res.* 3, S. 1–48. issn: 1532-4435.
- Bäck, T. (1997). „Handbook of Evolutionary Computation“. In: Department of Industrial Engineering, University of Pittsburgh, USA: Oxford University Press und IOP Publishing. Kap. Introduction, Section B1.1.
- Bäck, T., D. B. Fogel & Z. Michalewicz, (Hrsg.) (1997). *Handbook of Evolutionary Computation*. 1st. Bristol, UK, UK: IOP Publishing Ltd. isbn: 0750303921.
- Bakr, M. H., J. W. Bandler, R. M. Biernacki, S. H. Chen & K. Madsen (1998). „A trust region aggressive space mapping algorithm for EM optimization“. In: *IEEE Transactions on Microwave Theory and Techniques* 46 (12), S. 2412–2425. DOI:10.1109/22.739229.
- Bandler, J. W., R. M. Biernacki, S. H. Chen, P. A. Grobelny & R. H. Hemmers (1994). „Space mapping technique for electromagnetic optimization“. In: *IEEE Transactions on Microwave Theory and Techniques* 42.12, S. 2536–2544.
- Beer Sebastian; Appelrath, H.-J. (Okt. 2013). „A Formal Model for Agent-Based Coalition Formation in Electricity Markets“. In: *IEEE PES Conference on Innovative Smart Grid Technologies Europe (ISGT Europe 2013)*. IEEE Power & Energy Society.
- Bemporad, A. & M. Morari (1999). „Robust Model Predictive Control: A Survey“. In: *Robustness in Identification and Control* 245, S. 207–226.
- Ben-Hur, A., H. T. Siegelmann, D. Horn & V. Vapnik (2001). „Support Vector Clustering“. In: *Journal of Machine Learning Research* 2, S. 125–137.
- Bendel, C., D. Nestle & J. Ringelstein (Nov. 2006). „BEMI – Energiemanagement-Interface zur dezentralen Stromspeisung in Niederspannungsnetze“. In: *3. VWEW-Fachtagung "Photovoltaik und Netzintegration"*. Hrsg. von F. VWEW Energieverlag GmbH. Fulda: VWEW Energieverlag GmbH, Frankfurt.
- Bendrat, M. (2010). „Analyse der Spannungsqualität in Verteilnetzen auf der Niederspannungsebene sowie Realisierung geeigneter Kompensationsmaßnahmen“. Dissertation. Hagen: Fakultät für Mathematik und Informatik der Fern-Universität in Hagen.
- Beume, N., B. Naujoks & G. Rudolph (2008). „SMS-EMOA - Effektive evolutionäre Mehrzieloptimierung (SMS-EMOA - Effective Evolutionary Multiobjective Optimization).“ In: *Automatisierungstechnik* 56.7, S. 357–364.
- Bhardwaj, A., V. Kamboj, V. Shukla, B. Singh & P. Khurana (Juni 2012). „Unit commitment in electrical power system-a literature review“. In: *Power Engineering and Optimization Conference (PEDCO) Melaka, Malaysia, 2012 Ieee International*, S. 275–280. DOI:10.1109/PEOC0.2012.6230874.
- Bigler, T., G. Gaderer, P. Loschmidt & T. Sauter (2011). „SmartFridge: Demand Side Management for the device level“. In: *ETFA*. Hrsg. von Z. Mammeri. IEEE, S. 1–8. isbn: 978-1-4577-0018-7.

- Binder, I. & M. Braverman (2009). „The complexity of simulating Brownian Motion“. In: *SODA*. Hrsg. von C. Mathieu. SIAM, S. 58–67.
- Bitsch, R., W. Feldmann & G. Aumayr (2002). „Virtuelle Kraftwerke – Einbindung dezentraler Energieerzeugungsanlagen“. In: *etz* 9.123, S. 16–23. issn: 0948-7387.
- Blank, M., S. Gerwin, O. Krause & S. Lehnhoff (2011). „Support Vector Machines for an efficient Representation of Voltage Band Constraints“. In: *Innovative Smart Grid Technologies*. IEEE PES.
- Blank, M., T. Breithaupt, A. Dammasch, S. Garske & A. Nieße (Okt. 2013). *Evaluationsszenarien für die Teilprojekte 1 – 4*. technical report. OFFIS – Institut für Informatik, Universität Oldenburg, Universität Hannover und Technische Universität Braunschweig.
- Boßmann, T., W. Eichhammer & R. Elsland (Juni 2012). *Policy Report: Contribution of Energy Efficiency Measures to Climate Protection within the European Union until 2050*. Hrsg. von S. Karcher & M. Schäfer. Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit (BMU). Berlin: BMU in-house printing office.
- Borenstein, S., M. Jaske & A. Rosenfeld (2002). *Dynamic Pricing, Advanced Metering and Demand Response in Electricity Markets*. CSEM Working Papers. <http://www.ucei.berkeley.edu/PDF/csemwp105.pdf>.
- Boser, B. E., I. M. Guyon & V. N. Vapnik (1992). „A training algorithm for optimal margin classifiers“. In: *Proceedings of the fifth annual workshop on Computational learning theory*. COLT '92. New York, NY, USA: ACM, S. 144–152. isbn: 0-89791-497-X. DOI:10.1145/130385.130401.
- Bottou, L., O. Chapelle, D. DeCoste & J. Weston (2007). *Large-Scale Kernel Machines (Neural Information Processing)*. The MIT Press. isbn: 0262026252.
- Bouboulis, P., S. Theodoridis & C. Mavroforakis (2013). „Complex Support Vector Machines for Regression and Quaternary Classification“. In: *CoRR* abs/1303.2184.
- Bowman, A. W. & A. Azzalini (Nov. 1997). *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations (Oxford Statistical Science Series)*. Oxford University Press, USA. isbn: 0198523963.
- Boyd, S. & L. Vandenberghe (März 2004). *Convex Optimization*. Cambridge University Press. isbn: 0521833787.
- Brand, K. & W. Wimmer (2002). *Der Standard IEC 61850: Offene Kommunikation in Schaltanlagen im deregulierten Strommarkt*. Bulletin. SEV/ VSE H.1. [http://www.nettedautomation.com/download/mannheim-2003-03/SEV-Bulletin\\_Brand\\_2002\\_06.pdf](http://www.nettedautomation.com/download/mannheim-2003-03/SEV-Bulletin_Brand_2002_06.pdf).
- Bremer, J., B. Rapp, F. Jellinghaus & M. Sonnenschein (2009). „Tools for Teaching Demand-Side Management“. In: *Environmental Informatics and Industrial Environmental Protection - 23rd International Conference on Informatics for Environmental Protection*. Hrsg. von V. Wohlgemuth, B. Page & K. Voigt, S. 455–463. isbn: 978-3-8322-8397-1.
- Bremer, J. (Apr. 2012). „Ontology based Description of DER's Learned Environmental Performance Indicators“. In: *Proceedings of the 1st International Conference on Smart*

- Grids and Green IT Systems – SmartGreens 2012*. Hrsg. von B. Donnellan, J. P. Lopes, J. Martins & J. Filipe. Porto, Portugal: SciTePress, S. 107–112. isbn: 978-989-8565-09-9.
- Bremer, J. & M. Sonnenschein (2012a). „A Distributed Greedy Algorithm for Constraint-based Scheduling of Energy Resources“. In: *FedCSIS*. Hrsg. von M. Ganzha, L. A. Maciaszek & M. Paprzycki, S. 1285–1292. isbn: 978-83-60810-51-4.
- (2012b). „A Distributed Greedy Algorithm for Constraint-based Scheduling of Energy Resources“. In: *Federated Conference on Computer Science and Information Systems - FedCSIS 2012, Wroclaw, Poland, 9-12 September 2012, Proceedings*. Hrsg. von M. Ganzha, L. A. Maciaszek & M. Paprzycki, S. 1285–1292.
  - (2013a). „Automatic Reconstruction of Performance Indicators from Support Vector based Search Space Models in Distributed Real Power Planning Scenarios“. In: *GI-Jahrestagung*. Hrsg. von M. Horbach. Bd. 220. LNI. GI, S. 1441–1454. isbn: 978-3-88579-614-5.
  - (2013b). „Constraint-handling for Optimization with Support Vector Surrogate Models – A Novel Decoder Approach“. In: *ICAART 2013 – Proceedings of the 5th International Conference on Agents and Artificial Intelligence*. Hrsg. von J. Filipe & A. Fred. Bd. 2. Barcelona, Spain: SciTePress, S. 91–105. isbn: 978-989-8565-38-9.
  - (2013c). „Estimating Shapley Values for Fair Profit Distribution in Power Planning Smart Grid Coalitions“. In: *Multiagent System Technologies*. Hrsg. von M. Klusch, M. Thimm & M. Paprzycki. Bd. 8076. Lecture Notes in Computer Science. Springer Berlin Heidelberg, S. 208–221. isbn: 978-3-642-40775-8. DOI:10.1007/978-3-642-40776-5\_19. [http://dx.doi.org/10.1007/978-3-642-40776-5\\_19](http://dx.doi.org/10.1007/978-3-642-40776-5_19).
  - (Apr. 2013d). „Kommunikation von Umweltkennzahlen im Smart Grid und deren Integration in die verteilte Wirkleistungsplanung“. In: *5. BUIS-Tage: IT-gestütztes Ressourcen- und Energiemanagement*. Hrsg. von J. Marx Gómez, C. Lang & V. Wohlgemut. Oldenburg, Germany: Springer.
  - (2013e). „Sampling the Search Space of Energy Resources for Self-organized, Agent-based Planning of Active Power Provision“. In: *EnviroInfo*. Hrsg. von B. Page, A. G. Fleischer, J. Göbel & V. Wohlgemuth. Berichte aus der Umweltinformatik. Shaker, S. 214–222. isbn: 978-3-8440-1676-5.
- Bremer, J., B. Rapp & M. Sonnenschein (Apr. 2011a). „Encoding distributed Search Spaces for Virtual Power Plants“. In: *IEEE Symposium Series on Computational Intelligence 2011 (SSCI 2011)*. Paris, France.
- (2011b). „Including Environmental Performance Indicators into Kernel based Search Space Representations“. In: *Information Technologies in Environmental Engineering (ITEE 2011)*.
- Bremer, J. (März 2006). *Agenten-basierte Simulation des Planungsverhaltens adaptiver Verbraucher in Stromversorgungssystemen mit Real-Time-Pricing*. Diplomarbeit. C.v.O. Universität Oldenburg, Department für Informatik (Abteilung Umweltinformatik).
- Bremer, J. & M. Sonnenschein (2013f). „Model-based integration of constrained search spaces into distributed planning of active power provision.“ In: *Comput. Sci. Inf. Syst.* 10.4, S. 1823–1854.

- 
- (2013g). „Model-based integration of constrained search spaces into distributed planning of active power provision.“ In: *Comput. Sci. Inf. Syst.* 10.4, S. 1823–1854. <http://dblp.uni-trier.de/db/journals/comsis/comsis10.html#BremerS13>.
  - Bremer, J., B. Rapp & M. Sonnenschein (2010). „Support Vector based Encoding of Distributed Energy Resources' Feasible Load Spaces“. In: *IEEE PES Conference on Innovative Smart Grid Technologies Europe*. Chalmers Lindholmen, Gothenburg, Sweden.
  - Brock, P. (2004). „Optimization Using Space Mapping“. Supervised by Prof. Kaj Madsen, Assoc. Prof. Hans Bruun Nielsen, and Ph.D.stud. Frank Pedersen. Magisterarb. Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby: Informatics und Mathematical Modelling, Technical University of Denmark, DTU. <http://www2.imm.dtu.dk/pubdb/p.php?3259>.
  - Brown, F. W. M., R. A. Pielke Sr & J. D. Annan (2008). „Is there agreement amongst climate scientists on the IPCC AR4 WG1“. <http://pielke-climatesci.wordpress.com/2008/02/22/is-there-agreement-amongst-climate-scientists-on-the-ipcc-ar4-wg1>.
  - Broyden, C. G. (1965). „A class of methods for solving nonlinear simultaneous equations“. In: *Mathematics of Computation* 19, S. 577–593. <http://www.ams.org/journals/mcom/1965-19-092/S0025-5718-1965-0198670-6/home.html>.
  - Buchholz, B., V. Bühner, H. Frey & u.a. (2008). *VDE-Studie: Smart Distribution 2020 – virtuelle Kraftwerke in Verteilungsnetzen; technische, regulatorische und kommerzielle Rahmenbedingungen*. Studie der Energietechnischen Gesellschaft im VDE (ETG). Verband der Elektrotechnik Elektronik Informationstechnik e.V. (VDE).
  - Bundesnetzagentur (Dez. 2011). *Smart Grid und Smart Market*. Eckpunktepapier der Bundesnetzagentur zu den Aspekten des sich verändernden Energieversorgungssystems. Bonn: Bundesnetzagentur.
  - (2012). *'Smart Grid' und 'Smart Market' Eckpunktepapier der Bundesnetzagentur zu den Aspekten des sich verändernden Energieversorgungssystems*. [http://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Sachgebiete/Energie/Unternehmen\\_Institutionen/NetzzugangUndMesswesen/SmartGridEckpunktepapier/SmartGridPapier\\_EN.pdf](http://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Sachgebiete/Energie/Unternehmen_Institutionen/NetzzugangUndMesswesen/SmartGridEckpunktepapier/SmartGridPapier_EN.pdf) (besucht am 03.02.2014).
  - Burges, C. J. C. (Juni 1998). „A Tutorial on Support Vector Machines for Pattern Recognition“. In: *Data Min. Knowl. Discov.* 2.2, S. 121–167. issn: 1384-5810. DOI:10.1023/A:1009715923555.
  - Caron, S. & G. Kesidis (2010a). *Distributed Scheduling Algorithms for the Smart Grid*. <http://www.cse.psu.edu/~kesidis/papers/smartgridcomm10-extended.pdf> (besucht am 18.07.2013).
  - (Okt. 2010b). „Incentive-Based Energy Consumption Scheduling Algorithms for the Smart Grid“. In: *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, S. 391–396. DOI:10.1109/SMARTGRID.2010.5622073.
  - Carroll, C. W. (1961). „The Created Response Surface Technique for Optimizing Nonlinear, Restrained Systems“. In: *Operations Research* 9, S. 169–185. <http://www.ams.org/mathscinet-getitem?mr=23:B2057>.

- Casey-Lefkowitz, S., K. Colarulli & B. Baizel (Mai 2010). „Tar Sands Invasion: How Dirty and Expensive Oil from Canada Threatens America’s New Energy Economy“. In: <http://www.nrdc.org/energy/tarsandsinvasion.asp> (besucht am 04. 10. 2012).
- Chaitin, G. J. (1987). *Algorithmic information theory*. Cambridge tracts in theoretical computer science. Cambridge, Cambridgeshire, New York: Cambridge University Press. isbn: 0-521-34306-2. <http://opac.inria.fr/record=b1086413>.
- Chang, W.-C., C.-P. Lee, & C.-J. Lin (2013). *A revisit to support vector data description (SVDD)*. technical report. Taipei 10617, Taiwan: Department of Computer Science, National Taiwan University.
- Chapman, A. C., A. Rogers, N. R. Jennings & D. S. Leslie (2011). „A unifying framework for iterative approximate best-response algorithms for distributed constraint optimization problems.“ In: *Knowledge Eng. Review* 26.4, S. 411–444. <http://dblp.uni-trier.de/db/journals/ker/ker26.html#ChapmanRJJ11>.
- Charles River Associates, (Hrsg.) (Feb. 2005). *Primer on Demand-Side Management (with an emphasis on price-responsive programs)*. <http://siteresources.worldbank.org/INTENERGY/Resources/PrimeronDemand-SideManagement.pdf>. i.A. von: The World Bank, Washington, DC.
- Charpentier, A. D., J. A. Bergerson & H. L. MacLean (2009). „Understanding the Canadian oil sands industry’s greenhouse gas emissions“. In: *Environmental Research Letters* 4.1, S. 014005. <http://stacks.iop.org/1748-9326/4/i=1/a=014005>.
- Chassin, D., K. Schneider & C. Gerkenmeyer (Apr. 2008). „GridLAB-D: An open-source power systems modeling and simulation environment“. In: *Transmission and Distribution Conference and Exposition, 2008. T & D. IEEE/PES*, S. 1–5. DOI:10.1109/TDC.2008.4517260.
- Chua, K., S. Chou & W. Yang (2010). „Advances in heat pump systems: A review“. In: *Applied Energy* 87.12, S. 3611–3624. issn: 0306-2619. DOI:<http://dx.doi.org/10.1016/j.apenergy.2010.06.014>.
- Cloete, T., A. P. Engelbrecht & G. Pampara (2008). „Clib: A collaborative framework for Computational Intelligence algorithms - Part II.“ In: *IJCNN. IEEE*, S. 1764–1773.
- Cochran, W. G. (Jan. 1977). *Sampling Techniques, Third Edition*. New York, NY: Wiley & Sons. isbn: 047116240X.
- Coello Coello, C. A. (Jan. 2002). „Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art“. In: *Computer Methods in Applied Mechanics and Engineering* 191.11-12, S. 1245–1287. DOI:10.1016/S0045-7825(01)00323-1.
- Coll-Mayor, D., R. Picos & E. Garcíá-Moreno (2004). „State of the art of the virtual utility: the smart distributed generation network“. In: *International Journal of Energy Research* 28.1, S. 65–80. issn: 1099-114X. DOI:10.1002/er.951.
- Conrad, M., J. Dinger, H. Hartenstein, M. Schöller, M. Zitterbart & D. Rolli (2005). „30. A Peer-to-Peer Framework for Electronic Markets“. In: *Peer-to-Peer Systems and Applications*. Hrsg. von R. Steinmetz & K. Wehrle. Bd. 3485. Lecture Notes in Computer

- Science. Springer Berlin / Heidelberg, S. 509–525. isbn: 978-3-540-29192-3. DOI:10.1007/11530657\\_30.
- Consentec GmbH & r2b energy consulting GmbH (Juni 2010). *Voraussetzungen einer optimalen Integration erneuerbarer Energien in das Stromversorgungssystem*. Studie im Auftrag des Bundesministeriums für Wirtschaft und Technologie (BMWi), Endbericht. Aachen, Köln: Consentec Consulting für Energiewirtschaft und -technik GmbH und r2b energy consulting GmbH. <http://www.bmwi.de/DE/Mediathek/publikationen,did=349366.html>.
- Corkill, D. (Jan. 1991). „Blackboard Systems“. In: *AI Expert* 6.9. <http://mas.cs.umass.edu/paper/218>.
- Cortes, C. & V. Vapnik (Sep. 1995). „Support-Vector Networks“. In: *Mach. Learn.* 20.3, S. 273–297. issn: 0885-6125. DOI:10.1023/A:1022627411411.
- Costa, L. H. M., H. M. Ramos & M. A. H. de Castro (2010). „Hybrid genetic algorithm in the optimization of energy costs in water supply networks“. In: *Water Science & Technology: Water Supply – WSTWS* 10.3, pp. 315–326. issn: 1606-9749.
- Cristianini, N. & J. Shawe-Taylor (2000). *An Introduction to Support Vector Machines: And Other Kernel-based Learning Methods*. New York, NY, USA: Cambridge University Press. isbn: 0-521-78019-5.
- Dada, A. (2013). „Sustainable Sourcing and Procurement“. English. In: *Organizations' Environmental Performance Indicators*. Hrsg. von A. Dada, K. Stanoevska & J. Marx Gómez. Environmental Science and Engineering. Springer Berlin Heidelberg, S. 53–62. isbn: 978-3-642-32719-3. DOI:10.1007/978-3-642-32720-9\_5.
- Dadios, E. & J. Ashraf (2006). „Genetic algorithm with adaptive and dynamic penalty functions for the selection of cleaner production measures: a constrained optimization problem“. English. In: *Clean Technologies and Environmental Policy* 8.2, S. 85–95. issn: 1618-954X. DOI:10.1007/s10098-006-0036-9.
- Dasgupta, D. & Z. Michalewicz (1997). *Evolutionary algorithms in engineering applications*. Springer, S. I–XXI, 1–554. isbn: 978-3-540-62021-1.
- Davies, J. & M. Haubensstock (2002). „Building effective indicators to monitor operational risk“. In: *The RMA Journal* 5, S. 40–43.
- De Jong, K. (2006). *Evolutionary Computation: A Unified Approach*. Prentice-Hall Of India Pvt. Limited. isbn: 9788120330023. <http://books.google.de/books?id=bVHPgAACAAJ>.
- De Jong, K., D. B. Fogel & H. Schwefel (1997). „Handbook of Evolutionary Computation“. In: Department of Industrial Engineering, University of Pittsburgh, USA: Oxford University Press und IOP Publishing. Kap. A history of evolutionary computation, Section A2.3.
- Dean, T. & M. Boddy (Aug. 1988). „An Analysis of Time-Dependent Planning“. In: *Proceedings of the Seventh National Conference on Artificial Intelligence (AAAI-88)*. Saint Paul, Minnesota, USA: AAAI Press/MIT Press, S. 49–54. isbn: 0-262-51055-3.

- Delarue, E., D. Cattrysse & W. D'haeseleer (2013). „Enhanced priority list unit commitment method for power systems with a high share of renewables“. In: *Electric Power Systems Research* 105.0, S. 115–123. issn: 0378-7796. DOI:[10.1016/j.epsr.2013.07.014](https://doi.org/10.1016/j.epsr.2013.07.014).
- Dena (Nov. 2010). *dena-Netzstudie II - Integration erneuerbarer Energien in die deutsche Stromversorgung im Zeitraum 2015 - 2020 mit Ausblick 2025*. Deutsche Energie-Agentur GmbH (dena). Berlin.
- Deng, X. & C. H. Papadimitriou (Mai 1994). „On the complexity of cooperative solution concepts“. In: *Math. Oper. Res.* 19.2, S. 257–266. issn: 0364-765X. DOI:[10.1287/moor.19.2.257](https://doi.org/10.1287/moor.19.2.257).
- Dethlefs, T. & W. Renz (Nov. 2013). „A distributed registry for service-based energy management systems“. In: *Industrial Electronics Society, IECON 2013 - 39th Annual Conference of the IEEE*, S. 4710–4714. DOI:[10.1109/IECON.2013.6699896](https://doi.org/10.1109/IECON.2013.6699896).
- Dethlefs, T., T. Preisler & W. Renz (2014). „Multi-Agent-based Distributed Optimization for Demand-Side-Management Applications“. In: *Proceedings of the 2014 Federated Conference on Computer Science and Information Systems*. Hrsg. von M. P. M. Ganzha L. Maciaszek. Bd. 2. Annals of Computer Science and Information Systems. IEEE, pages 1489–1496. DOI:[10.15439/2014F251](https://doi.org/10.15439/2014F251). <http://dx.doi.org/10.15439/2014F251>.
- Energieeinsparverordnung. Verordnung zur Änderung der Energieeinsparverordnung vom 29.04.2009 (BGBl. I S. 954)*.
- National Renewable Energy Action Plan in accordance with Directive 2009/28/EC on the promotion of the use of energy from renewable sources. Renewables Directive*.
- Gesetz für den Vorrang Erneuerbarer Energien. Erneuerbare-Energien-Gesetz vom 25. Oktober 2008 (BGBl. I S. 2074), das zuletzt durch Artikel 5 des Gesetzes vom 20. Dezember 2012 (BGBl. I S. 2730) geändert worden ist*.
- Diaconis, P., S. Holmes & M. Shahshahani (2012). „Sampling from a manifold“. In: *arXiv preprint arXiv:1206.6913v2*.
- Donnellan, B., J. F. Martins, M. Helfert & K.-H. Krempels, (Hrsg.) (2013). *SMARTGREENS 2013 - Proceedings of the 2nd International Conference on Smart Grids and Green IT Systems, Aachen, Germany, 9-10 May, 2013*. SciTePress. isbn: 978-989-8565-55-6.
- Doran, P. T. & M. K. Zimmerman (2009). „Examining the Scientific Consensus on Climate Change“. In: *EOS Trans. AGU* 90.3. DOI:[10.1029/2009E0030002](https://doi.org/10.1029/2009E0030002).
- Dorigo, M. & T. Stützle (2004). *Ant Colony Optimization*. Scituate, MA, USA: Bradford Company. isbn: 0262042193.
- EPRI (1999). *Electricity Technology Roadmap: Powering Progress, 1999 Summary and Synthesis*. Bd. 2. Palo Alto, California: Electric Power Research Institute. <http://books.google.de/books?id=eL0yHAAACAAJ>.
- Eiben, A. (2001). „Evolutionary Algorithms and Constraint Satisfaction: Definitions, Survey, Methodology, and Research Directions“. In: *Theoretical Aspects of Evolutionary Computing*. Hrsg. von L. Kallel, B. Naudts & A. Rogers. Natural Computing Series. Berlin Heidelberg: Springer. Kap. 1, S. 13–58.

- Eiben, A. & Z. Ruttkay (1997). „Handbook of Evolutionary Computation“. In: Hrsg. von T. Bäck, D. B. Fogel & Z. Michalewicz. 1st. Bristol, UK, UK: IOP Publishing Ltd. Kap. Constraint satisfaction problems, C5.7:1–C5.7:8. isbn: 0750303921.
- Eichhorn, A, W Romisch & I Wegner-Specht (2005). „Simultane Optimierung von Risiko und Ertrag am Beispiel eines Strombeschaffungsportfolios eines kleineren Marktteilnehmers“. In: *VDI BERICHTE* 1908, S. 41.
- Ekanayake, J., N. Jenkins, K. Liyanage, J. Wu & A. Yokoyama (2012). *Smart Grid: Technology and Applications*. John Wiley & Sons. isbn: 9781119968689. <http://books.google.de/books?id=AmpSx0DiF7sC>.
- Directive 2009/28/EC of 23 April 2009 on the promotion of the use of energy from renewable sources and amending and subsequently repealing Directives 2001/77/EC and 2003/30/EC. Renewables Directive.*
- European SmartGrids Technology Platform* (2006). Techn. Ber. EUR 22040. European Commission Directorate-General for Research.
- Evangelista, P. F., M. J. Embrechts & B. K. Szymanski (2006). „Taming the Curse of Dimensionality in Kernels and Novelty Detection“. In: *Applied Soft Computing Technologies: The Challenge of Complexity*. Hrsg. von A. Abraham, B. de Baets, M. Köppen & B. Nickolay. Bd. 34. Advances in Soft Computing. Springer Berlin Heidelberg, S. 425–438. isbn: 978-3-540-31649-7. DOI:10.1007/3-540-31662-0\_33.
- Eysholdt, M., U. Denecke, M. Belasus, C. Bley, M. Eilers, M. Erlmann, C. Hinrichs, M. Hurrelmann, J. Bremer, B. Rapp et al. (2008). „A Tool for Modeling and Optimization of Residential Electricity Consumption“. In: *Environmental Informatics and Industrial Ecology - 22th International Conference on Informatics for Environmental Protection*. Shaker Verlag, ISBN, S. 978–3.
- Faria, P., Z. Vale, J. Soares & J. Ferreira (Juli 2013). „Demand Response Management in Power Systems Using Particle Swarm Optimization“. In: *Intelligent Systems, IEEE* 28.4, S. 43–51. issn: 1541-1672. DOI:10.1109/MIS.2011.35.
- Fatima, S. S., M. Wooldridge & N. R. Jennings (2008). „A linear approximation method for the Shapley value“. In: *Artif. Intell.* 172.14, S. 1673–1699.
- Fenn, B., D. Metz, T. Fiedler & A. Röglin (Nov. 2010). „Der Einsatz dezentraler Stromspeicher in Verteilnetzen“. In: *VDE-Kongress 2010 - E-Mobility: Technologien - Infrastruktur - Märkte*. Leipzig, Germany: VDE Verlag.
- Fiacco, A., G. McCormick & R. A. C. M. VA. (1965). *The Sequential Unconstrained Minimization Technique for Convex Programming with Equality Constraints*. Defense Technical Information Center. <http://books.google.de/books?id=1Sx5NwAACAAJ>.
- Fletcher, R. (1987). *Practical methods of optimization*. 2. Aufl. New York, NY, USA: Wiley-Interscience. isbn: 0-471-91547-5.
- Fogel, L., A. Owens & M. Walsh (1966). *Artificial intelligence through simulated evolution*. Chichester, WS, UK: Wiley.

- Fonseca, C. M. & P. J. Fleming (März 1995). „An Overview of Evolutionary Algorithms in Multiobjective Optimization“. In: *Evol. Comput.* 3.1, S. 1–16. issn: 1063-6560. DOI:[10.1162/evco.1995.3.1.1](https://doi.org/10.1162/evco.1995.3.1.1).
- Franch, T., M. Scheidt & G. Stock (2009). „Current and Future Challenges for Production Planning Systems“. In: *Optimization in the Energy Industry*. Hrsg. von J. Kallrath, P. M. Pardalos, S. Rebennack, M. Scheidt & P. M. Pardalos. Energy Systems. Springer Berlin Heidelberg, S. 5–17. isbn: 978-3-540-88965-6.
- Franz, O., M. Wissner, F. Büllingen, C.-I. Gries, C. Cremer, M. Klobasa, F. Sensfuß, S. Kimpeler, E. Baier, T. Lindner, H. Schäffler, W. Roth & M. Thoma (Dez. 2006). *Potenziäle der Informations- und Kommunikations-Technologien zur Optimierung der Energieversorgung und des Energieverbrauchs (eEnergy)*. Bundesministerium für Wirtschaft und Technologie (BMWi). Bad Honnef.
- Friedberg, R. M. (Jan. 1958). „A Learning Machine: Part I“. In: *IBM J. Res. Dev.* 2.1, S. 2–13. issn: 0018-8646. DOI:[10.1147/rd.21.0002](https://doi.org/10.1147/rd.21.0002).
- Friedman, T. (2009). *Hot, Flat, and Crowded 2.0: Why We Need a Green Revolution—and How It Can Renew America*. Picador. isbn: 9780312428921. <http://books.google.de/books?id=BpkALHFTnhUC>.
- Frisch, A. M., C. Jefferson, B. M. Hernández & I. Miguel (2005). „The Rules of Constraint Modelling.“ In: *IJCAI*. Hrsg. von L. P. Kaelbling & A. Saffiotti. Professional Book Center, S. 109–116. isbn: 0938075934. <http://dblp.uni-trier.de/db/conf/ijcai/ijcai2005.html#FrischJMM05>.
- Gano, S. E., H. Kim & D. E. Brown II (2006). „Comparison of three surrogate modeling techniques: Datascape, kriging, and second order regression“. In: *Proceedings of the 11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, AIAA-2006-7048*. Portsmouth, Virginia.
- Gao, C., E. Bompard, R. Napoli & H. Cheng (Aug. 2007). „Price forecast in the competitive electricity market by support vector machine“. In: *Physica A: Statistical Mechanics and its Applications* 382.1, S. 98–113.
- Garcke, J. (2004). „Maschinelles Lernen durch Funktionsrekonstruktion mit verallgemeinerten dünnen Gittern.“ Diss. University of Bonn.
- Gatterbauer, W. (2010). „Economic Efficiency of Decentralized Unit Commitment from a Generator’s Perspective“. In: *Engineering Electricity Services of the Future*. Hrsg. von M. Ilic. Springer.
- Gellings, C. W. (2009). *The Smart Grid: Enabling Energy Efficiency and Demand Response*. The Fairmont Press, Inc.
- Gellings, C. W. & J. H. Chamberlin (1988). *Demand-Side Management: Concepts and Methods*. The Fairmont Press, Inc.
- Gen, M. & R. Cheng (1997). *Genetic Algorithms and Engineering Design*. Wiley Series in Engineering Design and Automation. John Wiley & Sons. isbn: 9780471127413. <http://books.google.de/books?id=MCHCaJAHFJAC>.

- Genton, M. G. (März 2002). „Classes of Kernels for Machine Learning: A Statistics Perspective“. In: *J. Mach. Learn. Res.* 2, S. 299–312. issn: 1532-4435. <http://dl.acm.org/citation.cfm?id=944790.944815>.
- GhasemiGol, M., R. Monsefi & H. Yazdi (2009). „Ellipse Support Vector Data Description“. In: *Engineering Applications of Neural Networks*. Hrsg. von D. Palmer-Brown, C. Draganova, E. Pimenidis & H. Mouratidis. Bd. 43. Communications in Computer and Information Science. Springer Berlin Heidelberg, S. 257–268. isbn: 978-3-642-03968-3. DOI:10.1007/978-3-642-03969-0\_24. [http://dx.doi.org/10.1007/978-3-642-03969-0\\_24](http://dx.doi.org/10.1007/978-3-642-03969-0_24).
- GhasemiGol, M., M. Sabzekar, R. Monsefi, M. Naghibzadeh & H. S. Yazdi (2010). „A New Support Vector Data Description with Fuzzy Constraints“. In: *Proceedings of the 2010 International Conference on Intelligent Systems, Modelling and Simulation*. ISMS '10. Washington, DC, USA: IEEE Computer Society, S. 10–14. isbn: 978-0-7695-3973-7. DOI:10.1109/ISMS.2010.13.
- Gieseke, F. & O. Kramer (2013). „Towards Non-linear Constraint Estimation for Expensive Optimization“. In: *Applications of Evolutionary Computation*. Hrsg. von A. Esparcia-Alcázar. Bd. 7835. Lecture Notes in Computer Science. Springer Berlin Heidelberg, S. 459–468. isbn: 978-3-642-37191-2. DOI:10.1007/978-3-642-37192-9\_46.
- Gieseke, F., A. Airola, T. Pahikkala & O. Kramer (2012). „Sparse Quasi-Newton Optimization for Semi-supervised Support Vector Machines.“ In: *ICPRAM (1)*. Hrsg. von P. L. Carmona, J. S. Sánchez & A. L. N. Fred. SciTePress, S. 45–54. isbn: 978-989-8425-98-0. <http://dblp.uni-trier.de/db/conf/icpram/icpram2012-1.html#GiesekeAPK12>.
- (2014). „Fast and simple gradient-based optimization for semi-supervised support vector machines“. In: *Neurocomputing* 123.0. Contains Special issue articles: Advances in Pattern Recognition Applications and Methods, S. 23 –32. issn: 0925-2312. DOI:10.1016/j.neucom.2012.12.056.
- Glover, F. & M. Laguna (1997). *Tabu Search*. Norwell, MA, USA: Kluwer Academic Publishers. isbn: 079239965X.
- Goudarzi, H., S. Hatami & M. Pedram (2011). „Demand-side load scheduling incentivized by dynamic energy prices“. In: *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on*. IEEE, S. 351–356.
- Grötschel, M. (2013). *Einführung in die Lineare und Kombinatorische Optimierung (Algorithmische Diskrete Mathematik I)*. Vorlesungsskript. Technische Universität Berlin. [www.zib.de/groetschel/teaching/WS1213/Skriptum\\_ADM\\_I\\_aktuell.pdf](http://www.zib.de/groetschel/teaching/WS1213/Skriptum_ADM_I_aktuell.pdf) (besucht am 16.07.2014).
- Grüne, L. & J. Pannek (2011). *Nonlinear Model Predictive Control: Theory and Algorithms*. 1st ed. Communications and Control Engineering. Springer, S. xi+359. <http://www.springer.com/978-0-85729-500-2>.
- Guan, X., Q. Zhai & A. Papalexopoulos (2003). „Optimization based methods for unit commitment: Lagrangian relaxation versus general mixed integer programming“. In: Bd. 2, 1100 Vol. 2. [http://ieeexplore.ieee.org/xpls/abs/\\_all.jsp?arnumber=1270468](http://ieeexplore.ieee.org/xpls/abs/_all.jsp?arnumber=1270468).

- Guan, X., S. Guo & Q. Zhai (Nov. 2005). „The conditions for obtaining feasible solutions to security-constrained unit commitment problems“. In: *Power Systems, IEEE Transactions on* 20.4, S. 1746–1756. issn: 0885-8950. DOI:[10.1109/TPWRS.2005.857399](https://doi.org/10.1109/TPWRS.2005.857399).
- Gunn, S. R. (Mai 1998). *Support vector machines for classification and regression*. Techn. Ber. Faculty of Engineering, Science, Mathematics School of Electronics und Computer Science.
- Gutsch, C. & H. Stigler (2006). „Verbraucherseitiges Spitzenlastmanagement zur Optimierung des Gesamtsystems von Erzeugern und Verbrauchern“. In: *9. Symposium Energieinnovation*. Technische Universität Graz.
- Hall, P. (1927). „The Distribution of Means for Samples of Size N Drawn from a Population in which the Variate Takes Values Between 0 and 1, All Such Values Being Equally Probable“. English. In: *Biometrika* 19.3/4, pp. 240–245. issn: 00063444.
- Hammersley, J. & D. Handscomb (1964). *Monte Carlo Methods*. Methuen's monographs on applied probability and statistics. Methuen. isbn: 9780416523409.
- Hansen, N. (2000). „Invariance, Self-Adaptation and Correlated Mutations and Evolution Strategies“. In: *Proceedings of the 6th Conference on Parallel Problem Solving from Nature*. Hrsg. von M Schoenauer, K Deb, G Rudolph, X Yao, E Lutton, J. Merelo & H.-P Schwefel. Bd. 1917. Lecture Notes in Computer Science 1917. Springer Berlin Heidelberg New York, S. 355–364. isbn: 3-540-41056-2.
- (2006). „The CMA evolution strategy: a comparing review“. In: *Towards a new evolutionary computation. Advances on estimation of distribution algorithms*. Hrsg. von J. Lozano, P. Larranaga, I. Inza & E. Bengoetxea. Springer, S. 75–102.
- Hansen, N. (2011). *The CMA Evolution Strategy: A Tutorial*. Techn. Ber. [www.lri.fr/~hansen/cmatutorial.pdf](http://www.lri.fr/~hansen/cmatutorial.pdf) (besucht am 19.06.2014).
- Harmeling, S., A. Ziehe, M. Kawanabe & K.-R. Müller (2002). *Kernel-based Nonlinear Blind Source Separation*. Technical Report IST-1999-14190. Berlin, Potsdam, Germany: Fraunhofer FIRS.IDA und University of Potsdam.
- (Mai 2003). „Kernel-based Nonlinear Blind Source Separation“. In: *Neural Comput.* 15.5, S. 1089–1124. issn: 0899-7667. DOI:[10.1162/089976603765202677](https://doi.org/10.1162/089976603765202677).
- Hausdorff, F. (1914). *Grundzüge der Mengenlehre*. Das Hauptwerk von Felix Hausdorff. Leipzig: Veit und Company.
- Hayton, P., B. Schölkopf, L. Tarassenko & P. Anuzis (2000). „Support Vector Novelty Detection Applied to Jet Engine Vibration Spectra“. In: *Advances in Neural Information Processing Systems 13, NIPS 2000*. Hrsg. von T. K. Leen, T. G. Dietterich & V. Tresp. Denver, CO, USA: MIT Press, S. 946–952.
- Heindler, F. (2001). *Energiemanagement und Stromeinkauf im liberalisierten Markt*. Diplomarbeit. Technische Universität Graz.
- Helwig, S. & R. Wanka (Apr. 2007). „Particle Swarm Optimization in High-Dimensional Bounded Search Spaces“. In: *Swarm Intelligence Symposium, SIS*, S. 198–205. DOI:[10.1109/sis.2007.368046](https://doi.org/10.1109/sis.2007.368046).

- Herbrich, R., T. Graepel & C. Campbell (1999). „Bayes Point Machines : Estimating the Bayes Point in Kernel Space“. In: *Proceedings of IJCAI Workshop Support Vector Machines*, S. 23–27. <http://www.herbrich.me/papers/ijcai99.pdf>.
- Heuck, K., K. Dettmann & D. Schulz (2007). *Elektrische Energieversorgung*. Programm elektrische Energietechnik. Vieweg. isbn: 9783834802170. <http://books.google.de/books?id=4GE9BJw30t0C>.
- Hevner, A. R., S. T. March, J. Park & S. Ram (März 2004). „Design Science in Information Systems Research“. In: *MIS Q.* 28.1, S. 75–105. issn: 0276-7783. <http://dl.acm.org/citation.cfm?id=2017212.2017217>.
- Hinrichs, C. (2014). „Selbstorganisierte Einsatzplanung dezentraler Akteure im Smart Grid“. in Begutachtung. Dissertation. Oldenburg: Carl von Ossietzky Universität Oldenburg.
- Hinrichs, C., S. Lehnhoff & M. Sonnenschein (2013a). „A Decentralized Heuristic for Multiple-Choice Combinatorial Optimization Problems“. In: *Operations Research 2012 – Selected Papers of the International Conference on Operations Research (OR 2012)*. Hannover, Germany: Springer. <http://www-ui.informatik.uni-oldenburg.de/download/Publikationen/HLS12.pdf>.
- Hinrichs, C., J. Bremer & M. Sonnenschein (2013b). „Distributed Hybrid Constraint Handling in Large Scale Virtual Power Plants“. In: *IEEE PES Conference on Innovative Smart Grid Technologies Europe (ISGT Europe 2013)*. IEEE Power & Energy Society. <http://www-ui.informatik.uni-oldenburg.de/download/Publikationen/HBS13.pdf>.
- Hinrichs, C., M. Sonnenschein & S. Lehnhoff (2013c). „Evaluation of a Self-Organizing Heuristic for Interdependent Distributed Search Spaces“. In: *International Conference on Agents and Artificial Intelligence (ICAART 2013)*. Hrsg. von J. Filipe & A. L. N. Fred. Bd. Volume 1 – Agents. SciTePress, S. 25–34. isbn: 978-989-8565-38-9. DOI:10.5220/0004227000250034. <http://www-ui.informatik.uni-oldenburg.de/download/Publikationen/HSL13.pdf>.
- Hinrichs, C., S. Lehnhoff & M. Sonnenschein (2014). „A Decentralized Heuristic for Multiple-Choice Combinatorial Optimization Problems“. In: *Operations Research Proceedings 2012. Selected Papers of the International Annual Conference of the German Operations Research Society (GOR)*. Springer. Kap. XIII, S. 297–302. isbn: 978-3-319-00795-3. DOI:10.1007/978-3-319-00795-3\_43. <http://www-ui.informatik.uni-oldenburg.de/download/Publikationen/HLS14a.pdf>.
- Hoffmann, H. (März 2007). „Kernel PCA for Novelty Detection“. In: *Pattern Recogn.* 40.3, S. 863–874. issn: 0031-3203. DOI:10.1016/j.patcog.2006.07.009.
- Holland, J. H. (1967). *Nonlinear environments permitting efficient adaptation*. Computer and Information Sciences II. New York: Academic.
- Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. second edition, 1992. Ann Arbor, MI: University of Michigan Press.
- Hommelberg, M. P. F., B. J. van der Velde, C. J. Warmer, I. G. Kamphuis & J. K. Kok (2008). „A novel architecture for real-time operation of multi-agent based coordination of demand and supply“. In: *Power Engineering Society, IEEE General Meeting*, S. 1–5. DOI:10.1109/PES.2008.4596531.

- Hordijk, W. (Dez. 1996). „A Measure of Landscapes“. In: *Evol. Comput.* 4.4, S. 335–360. issn: 1063-6560. DOI:[10.1162/evco.1996.4.4.335](https://doi.org/10.1162/evco.1996.4.4.335).
- Horenkamp, W., W. Hube, J. Jäger, M. Kleimaier & u. a. (2007). *VDE-Studie Dezentrale Energieversorgung 2020 – Gesamtttext*. Forschungsbericht. Frankfurt: Energietechnische Gesellschaft im VDE (ETG).
- Hřebíček, J., P. Misařová & J. Hyršlová (2007). „Environmental Key Performance Indicators and Corporate Reporting“. In: *Environmental Accounting and Sustainable Development Indicators (EA-SDI 2007)*. Universita Jana Evangelisty Purkyně, S. 978–980.
- Hsu, M.-C. & V.-W. Soo (2009). „Multi-Agent Systems for Society“. In: Hrsg. von D. Lukose & Z. Shi. Berlin, Heidelberg: Springer-Verlag. Kap. Fairness in Cooperating Multi-agent Systems — Using Profit Sharing as an Example, S. 153–162. isbn: 978-3-642-03337-7. DOI:[10.1007/978-3-642-03339-1\\_13](https://doi.org/10.1007/978-3-642-03339-1_13).
- Huber, D., Z. Taylor & S. Knudsen (Jan. 2011). *Environmental Impacts of Smart Grid*. technical report DOE/NETL-2010/1428. Morgantown, West Virginia; Pittsburgh, Pennsylvania; Sugar Land, Texas; Albany, Oregon; und Anchorage, Alaska: National Energy Technology Laboratory; Office of Strategic Energy Analysis und Planning.
- Huttenlocher, D. P., G. A. Klanderman & W. Rucklidge (1993). „Comparing Images Using the Hausdorff Distance.“ In: *IEEE Trans. Pattern Anal. Mach. Intell.* 15.9, S. 850–863.
- IEC 61850-7-420:2009 (2009). *Communication networks and systems for power utility automation - Part 7-420: Basic communication structure - Distributed energy resources logical nodes*. IEC, Geneva, Switzerland.
- IPCC (2011). *Special Report on Renewable Energy Sources and Climate Change Mitigation*. Hrsg. von O. Edenhofer, R. Pichs-Madruga, Y. Sokona, K. Seyboth, P. Matschoss, S. Kadner, T. Zwickel, P. Eickemeier, G. Hansen, S. Schlömer & C. von Stechow. United Kingdom und New York, NY, USA: Cambridge University Press.
- Ilić, M. D. (2007). „From Hierarchical to Open Access Electric Power Systems“. In: *Proceedings of the IEEE* 95.5, S. 1060–1084. issn: 0018-9219. DOI:[10.1109/JPROC.2007.894711](https://doi.org/10.1109/JPROC.2007.894711).
- International Energy Agency (2002). *Distributed Generation in Liberalised Electricity Markets*. OECD Publishing. isbn: 978-9-2641-7597-6. DOI:[10.1787/9789264175976-en](https://doi.org/10.1787/9789264175976-en).
- Jamous, N. & K. Müller (2013). „Environmental Performance Indicators“. English. In: *Organizations' Environmental Performance Indicators*. Hrsg. von A. Dada, K. Stanoevska & J. Marx Gómez. Environmental Science and Engineering. Springer Berlin Heidelberg, S. 3–18. isbn: 978-3-642-32719-3. DOI:[10.1007/978-3-642-32720-9\\_4](https://doi.org/10.1007/978-3-642-32720-9_4).
- Jarboui, B., N. Damak, P. Siarry & A. R. Rebai (2008). „A combinatorial particle swarm optimization for solving multi-mode resource-constrained project scheduling problems.“ In: *Applied Mathematics and Computation* 195.1, S. 299–308. <http://dblp.uni-trier.de/db/journals/amc/amc195.html#JarbouiDSR08>.
- Jiang, L. & S. Low (Dez. 2011). „Multi-period optimal energy procurement and demand response in smart grid with uncertain supply“. In: *Decision and Control and European*

- Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, S. 4348–4353. DOI:[10.1109/CDC.2011.6161320](https://doi.org/10.1109/CDC.2011.6161320).
- Jiménez, Á. B., J. L. Lázaro & J. R. Dorronsoro (2008). „Finding Optimal Model Parameters by Discrete Grid Search“. In: *Innovations in Hybrid Intelligent Systems*. Hrsg. von E. Corchado, J. M. Corchado & A. Abraham. Bd. 44. Advances in Soft Computing. Springer, S. 120–127. isbn: 978-3-540-74971-4.
- Juszczyk, P., D. Tax & R. P. W. Duin (2002). „Feature scaling in support vector data description“. In: *Proc. ASCI 2002, 8th Annual Conf. of the Advanced School for Computing and Imaging*. Hrsg. von E. Depreitere, A. Belloum, J. Heijnsdijk & F. van der Stappen, S. 95–102.
- Kaestle, G. (2007). „Das virtuelle Heizkraftwerk als realer Cluster stromerzeugender Heizungen“. In: *Tagungsband Internationaler ETG-Kongress*. ETG-Fachbericht 108. Berlin: VDE-Verlag, S. 473–480.
- Kahan, J. & A. Rapoport (1984). *Theories of coalition formation*. Basic Studies in Human Behavior. L. Erlbaum Associates. isbn: 9780898592986. <http://books.google.de/books?id=TABSQAIAAJ>.
- Kahlen, C., H. Krings, P. Wittlinger, B. S. von der Brelie, S. Federlein, C. Matrose, R. Frings & T. Smolka (Aug. 2010). „Einflüsse und Auswirkungen von Elektromobilität auf das Energieversorgungsnetz sowie Anforderungen und Voraussetzungen für eine interoperable Netzintegration“. In: *VDE-Kongress 2010 - E-Mobility: Technologien - Infrastruktur - Märkte*. Leipzig, Germany: VDE Verlag.
- Kalenda, N. (2003). *Dynamische Leistungsregelung von Kleinst-Stirling-BHKW's im Inselbetrieb*. Berichte aus der Energietechnik. Shaker. isbn: 9783832214760. <http://books.google.de/books?id=0dlsAgAACA>.
- Kallrath, J. (2013). *Gemischt-ganzzahlige Optimierung: Modellierung in der Praxis: mit Fallstudien aus Chemie, Energiewirtschaft, Papierindustrie, Metallgewerbe, Produktion und Logistik*. Springer Fachmedien. isbn: 9783658006907.
- Kamphuis, I. G., J. K. Kok, C. J. Warmer & M. P. F. Hommelberg (2007a). *Massive coordination of residential embedded electricity generation and demand response using the PowerMatcher approach*. Techn. Ber. Energy Research Centre of the Netherlands, ECN.
- Kamphuis, R., C. Warmer, M. Hommelberg & K. Kok (Mai 2007b). „Massive Coordination of Dispersed Generation using Powermatcher based Software Agents“. In: *19th International Conference on Electricity Distribution*. Vienna.
- Karaboga, D. & B. Basturk (Nov. 2007). „A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm“. In: *Journal of Global Optimization* 39.3, S. 459–471. DOI:[10.1007/s10898-007-9149-x](https://doi.org/10.1007/s10898-007-9149-x).
- Kennedy, J. & R. Eberhart (Nov. 1995). „Particle swarm optimization“. In: *Neural Networks, 1995. Proceedings., IEEE International Conference on*. Bd. 4. Perth, WA, Australia: IEEE, 1942–1948 vol.4. isbn: 0-7803-2768-3. DOI:[10.1109/ICNN.1995.488968](https://doi.org/10.1109/ICNN.1995.488968).
- Khan, S. S. & M. G. Madden (2010). „A Survey of Recent Trends in One Class Classification“. In: *Proceedings of the 20th Irish Conference on Artificial Intelligence and Cognitive*

- Science, Dublin, Ireland*. AICS'09. Berlin, Heidelberg: Springer-Verlag, S. 188–197. isbn: 978-3-642-17079-9. <http://dl.acm.org/citation.cfm?id=1939047.1939070>.
- Kim, D. G. (1998). „Riemann mapping based constraint handling for evolutionary search“. In: SAC, S. 379–385. DOI:<http://doi.acm.org/10.1145/330560.330844>.
- Kim, P. J., H. J. Chang & J. Y. Choi (2008). „Fast incremental learning for one-class support vector classifier using sample margin information.“ In: *19th International Conference on Pattern Recognition (ICPR 2008), December 8-11, 2008, Tampa, Florida, USA*. IEEE, S. 1–4. isbn: 978-1-4244-2175-6.
- Kim, P., H. Chang, D. Song & J. Choi (2007). „Fast Support Vector Data Description Using K-Means Clustering“. In: *Advances in Neural Networks – ISNN 2007*. Hrsg. von D. Liu, S. Fei, Z. Hou, H. Zhang & C. Sun. Bd. 4493. Lecture Notes in Computer Science. 10.1007/978-3-540-72395-0\_64. Springer Berlin / Heidelberg, S. 506–514. isbn: 978-3-540-72394-3.
- Kirkpatrick, S., C. D. Gelatt & M. P. Vecchi (1983). „Optimization by Simulated Annealing“. In: *Science*. New Series 220.4598, S. 671–680. issn: 00368075. DOI:[10.1126/science.220.4598.671](https://doi.org/10.1126/science.220.4598.671). <http://www.jstor.org/stable/1690046>.
- Knorre, J. (Okt. 2011). *Erforschung und Erprobung neuer Fahrzeugkonzepte zur Elektromobilität*. Abschlussbericht zum Vorhaben im Rahmen des FuE-Programms "Förderung von Forschung und Entwicklung im Bereich der Elektromobilität" 16EM0037. Unter Mitarbeit von K. Wippel, Z. Hagen und M. Treblin, mit einem Beitrag von Lothar Bach und Ulf Rahmel. München: Bayerische Motoren Werke AG. [http://www.pt-elektromobilitaet.de/projekte/foerderprojekte-aus-dem-konjunkturpaket-ii-2009-2011/pkw-feldversuche/abschlussberichte/abschlussbericht-activee\\_bmw.pdf](http://www.pt-elektromobilitaet.de/projekte/foerderprojekte-aus-dem-konjunkturpaket-ii-2009-2011/pkw-feldversuche/abschlussberichte/abschlussbericht-activee_bmw.pdf).
- Kok, K., Z. Derzsi, J. Gordijn, M. Hommelberg, C. Warmer, R. Kamphuis & H. Akkermans (2008). „Agent-Based Electricity Balancing with Distributed Energy Resources, A Multiperspective Case Study“. In: *Hawaii International Conference on System Sciences 0*, S. 173. issn: 1530-1605. DOI:<http://doi.ieeecomputersociety.org/10.1109/HICSS.2008.46>.
- Königsberger, K. (1991). *Analysis 2*. Springer-Lehrbuch. Berlin, Heidelberg, New York: Springer-Verlag. isbn: 3-540-54723-1.
- Köppen, M. (2000). „The Curse of Dimensionality“. In: *5th Online World Conference in Industrial Applications*. WSC '05. On the Internet.
- Koziel, S. & Z. Michalewicz (1998). „A decoder-based evolutionary algorithm for constrained parameter optimization problems“. In: *Parallel Problem Solving from Nature – PPSN V*. Hrsg. von A. n. Eiben, T. Bäck, M. Schoenauer & H.-P. Schwefel. Bd. 1498. Lecture Notes in Computer Science. Springer Berlin Heidelberg, S. 231–240. isbn: 978-3-540-65078-2. DOI:[10.1007/BFb0056866](https://doi.org/10.1007/BFb0056866).
- (März 1999). „Evolutionary algorithms, homomorphous mappings, and constrained parameter optimization“. In: *Evol. Comput.* 7 (1), S. 19–44. issn: 1063-6560. DOI:<http://dx.doi.org/10.1162/evco.1999.7.1.19>.
- Kramer, O. (Jan. 2010). „A review of constraint-handling techniques for evolution strategies“. In: *Appl. Comp. Intell. Soft Comput.* 2010, S. 1–19. issn: 1687-9724. DOI:<http://dx.doi.org/10.1155/2010/185063>.

- Kramer, O. & F. Gieseke (2011). „Short-Term Wind Energy Forecasting Using Support Vector Regression.“ In: *SOCO*. Hrsg. von E. Corchado, V. Snásel, J. Sedano, A. E. Hassanien, J. L. Calvo-Rolle & D. Slezak. Bd. 87. *Advances in Soft Computing*. Springer, S. 271–280. isbn: 978-3-642-19643-0.
- Kramer, O., B. Satzger & J. Lässig (2010). „Power Prediction in Smart Grids with Evolutionary Local Kernel Regression.“ In: *HAIS (1)*. Hrsg. von M. G. Romay, E. Corchado & M. T. García-Sebastián. Bd. 6076. *Lecture Notes in Computer Science*. Springer, S. 262–269. isbn: 978-3-642-13768-6. DOI:[10.1007/978-3-642-13769-3](https://doi.org/10.1007/978-3-642-13769-3). <http://dblp.uni-trier.de/db/conf/hais/hais2010-1.html#KramerSL10>.
- Kuhn, H. W. & A. W. Tucker (1950). „Nonlinear Programming“. In: *Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability*. Hrsg. von J. Neyman. University of California Press, Berkeley, CA, USA, S. 481–492.
- Kühn, W. (2007). „Systemdienstleistungen unter besonderer Berücksichtigung von Wechselrichtereinspeisungen“. In: *Aktive Energienetze im Kontext der Energiewende*. Frankfurt: Aktive Energienetze im Kontext der Energiewende.
- Kurban, M. & Ü. B. Filik (2009). „A Comparative Study of Three Different Mathematical Methods for Solving the Unit Commitment Problem“. In: *Mathematical Problems in Engineering* 2009. Article ID 368024. DOI:[10.1155/2009/368024](https://doi.org/10.1155/2009/368024).
- Kurscheid, E. (2009). „Randbedingungen für die Lieferung von Regelleistung aus BHKW-Pools“. In: *5. BHKW-Infotage*. Goslar, Germany.
- Kütz, M. (2008). *Kennzahlen in der IT: Werkzeuge für Controlling und Management*. 3. Aufl. Heidelberg: dpunkt. isbn: 978-3-89864-579-9.
- Kwok, J. T. Y. & I. W. H. Tsang (Nov. 2004). „The pre-image problem in kernel methods“. In: *Neural Networks, IEEE Transactions on* 15.6, S. 1517–1525. DOI:[10.1109/TNN.2004.837781](https://doi.org/10.1109/TNN.2004.837781). <http://www.hpl.hp.com/conferences/icml2003/papers/345.pdf>.
- Lang, C. B. & N. Pucker (1998). *Mathematische Methoden in der Physik*. ger. Hochschul Taschenbuch. Heidelberg [u.a.]: Spektrum, Akad. Verl., XVII, 625 S. isbn: 3-8274-0225-5 ; 978-3-8274-0225-7.
- Lee, J., H.-J. Kim, G.-L. Park & M. Kang (2012). „Energy Consumption Scheduler for Demand Response Systems in the Smart Grid.“ In: *J. Inf. Sci. Eng.* 28.5, S. 955–969.
- Lee, K. S. & Z. W. Geem (2005). „A new meta-heuristic algorithm for continuous engineering optimization: harmony search theory and practice“. In: *Computer Methods in Applied Mechanics and Engineering* 194.36–38, S. 3902–3933. issn: 0045-7825. DOI:<http://dx.doi.org/10.1016/j.cma.2004.09.007>. <http://www.sciencedirect.com/science/article/pii/S0045782504004682>.
- Lee, S.-H. & K. M. Daniels (2012). „Gaussian kernel width exploration and cone cluster labeling for support vector clustering.“ In: *Pattern Anal. Appl.* 15.3, S. 327–344. <http://dblp.uni-trier.de/db/journals/paa/paa15.html#LeeD12>.
- Leguizamón, G. & C. A. C. Coello (2009). „Boundary Search for Constrained Numerical Optimization Problems With an Algorithm Inspired by the Ant Colony Metaphor.“ In: S. 350–368.

- Lehnhoff, S. (2010). *Dezentrales vernetztes Energiemanagement: Ein Ansatz auf Basis eines verteilten adaptiven Realzeit-Multiagentensystems*. Vieweg + Teubner research. Vieweg+Teubner Verlag. isbn: 9783834812704. <http://books.google.de/books?id=nQScngEACAAJ>.
- Lehnhoff, S., W. Mahnke, S. Rohjans & M. Uslar (Aug. 2011). „IEC 61850 based OPC UA Communication – The Future of Smart Grid Automation“. In: *17th Power Systems Computation Conference (PSCC 2011), Stockholm*.
- Liang, J., S. Liu & D. Wu (2009). „Fast Training of SVDD by Extracting Boundary Targets“. In: *Iranian Journal of Electrical and Computer Engineering (IJECE)* 8.2, S. 133–137.
- Liao, Q., J.-B. Wang, Y. W. Webster & I. A. Watson (2009). „GPU Accelerated Support Vector Machines for Mining High-Throughput Screening Data“. In: *Journal of Chemical Information and Modeling* 49.12, S. 2718–2725.
- Liben-Nowell, D., A. Sharp, T. Wexler & K. Woods (2012). „Computing Shapley Value in Supermodular Coalitional Games“. In: *COCOON*. Hrsg. von J. Gudmundsson, J. Mestre & T. Viglas. Bd. 7434. Lecture Notes in Computer Science. Springer, S. 568–579. isbn: 978-3-642-32240-2.
- Liu, B., Y. Xiao, L. Cao, Z. Hao & F. Deng (2013). „SVDD-based outlier detection on uncertain data“. English. In: *Knowledge and Information Systems* 34.3, S. 597–618. issn: 0219-1377. DOI:10.1007/s10115-012-0484-y. <http://dx.doi.org/10.1007/s10115-012-0484-y>.
- Lopes, J. P., N. Hatziaargyriou, J. Mutale, P. Djapic & N. Jenkins (2007). „Integrating distributed generation into electric power systems: A review of drivers, challenges and opportunities“. In: *Electric Power Systems Research* 77.9. Distributed Generation, S. 1189–1203. issn: 0378-7796. DOI:<http://dx.doi.org/10.1016/j.epsr.2006.08.016>. <http://www.sciencedirect.com/science/article/pii/S0378779606001908>.
- Löschner, E. (2013). „OEPI Ontology“. English. In: *Organizations' Environmental Performance Indicators*. Hrsg. von A. Dada, K. Stanoevska & J. Marx Gómez. Environmental Science and Engineering. Springer Berlin Heidelberg, S. 85–102. isbn: 978-3-642-32719-3. DOI:10.1007/978-3-642-32720-9\_8.
- Lukovic, S., I. Kaitovic, M. Mura & U. Bondi (2010). „Virtual Power Plant As a Bridge between Distributed Energy Resources and Smart Grid“. In: *Hawaii International Conference on System Sciences* 0, S. 1–8. DOI:10.1109/HICSS.2010.437.
- Lünsdorf, O. (März 2012a). „Selbstorganisation virtueller Geräte für das Lastmanagement von Kleinverbrauchern“. Diss. Universität Oldenburg.
- Lünsdorf, O. (2012b). *Selbstorganisation virtueller Geräte für das Lastmanagement von Kleinverbrauchern*. Dissertation, IV, 181 S.
- Lünsdorf, O. & M. Sonnenschein (Okt. 2010). „A pooling based load shift strategy for household appliances“. In:
- Lust, T. & J. Teghem (2010). „The multiobjective multidimensional knapsack problem: a survey and a new approach“. In: *CoRR* abs/1007.4063. <http://dblp.uni-trier.de/db/journals/corr/corr1007.html#abs-1007-4063>.

- Macary, J.-P., K. Rechenberg, H. Rubenbauer, P. Eckert & A. Rasic (Nov. 2012). „Smart Power Applications and active influence of power quality in distribution networks with Energy Storage Solutions“. In: *VDE-Kongress 2012 - Intelligente Energieversorgung der Zukunft*. Stuttgart, Germany: VDE Verlag.
- Mahnke, W., S.-H. Leitner & M. Damm (2009). *OPC Unified Architecture*. 1st. Springer Publishing Company, Incorporated. isbn: 3540688986, 9783540688983.
- Mailler, R. & V. Lesser (2004). „Solving Distributed Constraint Optimization Problems Using Cooperative Mediation“. In: *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 1*. AAMAS '04. Washington, DC, USA: IEEE Computer Society, S. 438–445. isbn: 1-58113-864-4. DOI:[10.1109/AAMAS.2004.249](https://doi.org/10.1109/AAMAS.2004.249).
- Makkonen, S. & R. Lahdelma (2006). „Non-convex power plant modelling in energy optimisation“. In: *European Journal of Operational Research* 171.3. Feature Cluster: Heuristic and Stochastic Methods in Optimization Feature Cluster: New Opportunities for Operations Research, S. 1113–1126. issn: 0377-2217. DOI:<http://dx.doi.org/10.1016/j.ejor.2005.01.020>.
- Malinowski, J. & K. Kaderly (2004). „Peak shaving - a method to reduce utility costs“. In: *Annual Technical Conference, IEEE Region 5*. DOI:[10.1109/REG5.2004.1300158](https://doi.org/10.1109/REG5.2004.1300158).
- Manevitz, L. M. & M. Yousef (März 2002). „One-class Svms for Document Classification“. In: *J. Mach. Learn. Res.* 2, S. 139–154. issn: 1532-4435. <http://dl.acm.org/citation.cfm?id=944790.944808>.
- Mann, I. & L. Shapley (1960). *Values of Large Games, IV: Evaluating the Electoral College by Montecarlo Techniques*. Technical report. Santa Monica, CA: RAND Corporation. [http://www.rand.org/pubs/research/\\_memoranda/RM2651](http://www.rand.org/pubs/research/_memoranda/RM2651).
- Mao, Y. & M. Li (2008). „Optimal Reactive Power Planning Based on Simulated Annealing Particle Swarm Algorithm Considering Static Voltage Stability“. In: *Proceedings of the 2008 International Conference on Intelligent Computation Technology and Automation - Volume 01*. ICICTA '08. Washington, DC, USA: IEEE Computer Society, S. 106–110. isbn: 978-0-7695-3357-5. DOI:[10.1109/ICICTA.2008.427](https://doi.org/10.1109/ICICTA.2008.427).
- Martello, Silvano & P. Toth (1990). *Knapsack problems: algorithms and computer implementations*. New York, NY, USA: John Wiley & Sons, Inc. isbn: 0-471-92420-2.
- Mas-Colell, A., M. D. Whinston & J. R. Green (Juni 1995). *Microeconomic Theory*. Oxford University Press. isbn: 0195073401.
- Masterson, T. & D. Topiwala (Nov. 2008). *Multi-Agent Traffic Light Optimisation and Coordination*. TRT-UK Electronic release. White Paper.
- Mayer, C. & S. Rohjans (2010). „Smart Grids aus IT-Sicht und ihre Einbettung in nationale und internationale Frameworks“. In: *Handbuch Energiemanagement* 30. Hrsg. von H.-P. Beck, J. Buddenberg, E. Meller & C. Salander.
- McArthur, S., E. Davidson, V. Catterson, A. Dimeas, N. Hatziargyriou, F. Ponci & T. Funabashi (2007). „Multi-Agent Systems for Power Engineering Applications – Part I: Concepts,

- Approaches, and Technical Challenges“. In: *IEEE Transactions on Power Systems* 22.4, S. 1743–1752. issn: 0885-8950. DOI:10.1109/TPWRS.2007.908471.
- McMorran, A. W. (Jan. 2007). *An Introduction to IEC 61970-301 & 61968-11: The Common Information Model*. technical report. Glasgow, UK: University of Strathclyde.
- Meister, U. (Jan. 2012). *Ohne Markt kein Smart Grid*. Onlineartikel. Zürich, Genf: Avenir Suisse, Gerhard Schwarz (Hrsg.) <http://www.avenir-suisse.ch/13400/ohne-markt-kein-smart-grid/> (besucht am 02.04.2013).
- Mémoli, F. & G. Sapiro (2004). „Comparing Point Clouds“. In: *Proceedings of the 2004 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing (Nice, France)*. SGP '04. New York, NY, USA: ACM, S. 32–40. isbn: 3-905673-13-4. DOI:10.1145/1057432.1057436.
- Meng, Z., R. Shen & M. Jiang (Dez. 2011). „An Objective Penalty Functions Algorithm for Multiobjective Optimization Problem“. In: *American Journal of Operations Research* 1, S. 229–235. DOI:10.4236/ajor.2011.14026.
- Mercer, J. (1909). „Functions of Positive and Negative Type, and their Connection with the Theory of Integral Equations“. In: *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 209.441-458, S. 415–446. DOI:10.1098/rsta.1909.0016. eprint: <http://rsta.royalsocietypublishing.org/content/209/441-458/415.full.pdf+html>.
- Merkuryeva, G. & V. Bolshakovs (2011). „Benchmark Fitness Landscape Analysis“. In: *International Journal of Simulation Systems, Science & Technology (IJSSST)* 12.2. Hrsg. von D. Al-Dabass & V. Colla, S. 38–45.
- Merz, M. & M. Wüthrich (2013). *Mathematik für Wirtschaftswissenschaftler: Die Einführung mit vielen ökonomischen Beispielen*. Vahlen. isbn: 9783800644834. <http://books.google.de/books?id=DBD-SoT1oRYC>.
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller & E. Teller (1953). „Equation of State Calculations by Fast Computing Machines“. In: *The Journal of Chemical Physics* 21.6, S. 1087–1092. DOI:10.1063/1.1699114. <http://link.aip.org/link/?JCP/21/1087/1>.
- Meyerholt, D., J. Marx Gómez, A. Dada, J. Bremer & B. Rapp (2010). „Bringing sustainability to the daily business: The OEPI Project“. In: *Proceedings of the Workshop Environmental Information Systems and Services - Infrastructures and Platforms*. Hrsg. von A. J. Berre, D. Roman & P. Maué.
- Michalewicz, Z. (1995). „A survey of constraint handling techniques in evolutionary computation methods“. In: *Proceedings of the 4th Annual Conference on Evolutionary Programming*, S. 135–155. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.117.9084>.
- (1996). *Genetic algorithms + data structures = evolution programs (3rd ed.)* London, UK: Springer-Verlag. isbn: 3-540-60676-9.
- Michalewicz, Z. & D. B. Fogel (Dez. 2004). *How to solve it : modern heuristics*. Springer. isbn: 3540224947.

- Michalewicz, Z. & M. Schoenauer (März 1996). „Evolutionary algorithms for constrained parameter optimization problems“. In: *Evol. Comput.* 4 (1), S. 1–32. issn: 1063-6560.
- Mika, S., B. Schölkopf, A. Smola, K.-R. Müller, M. Scholz & G. Rätsch (1999a). „Kernel PCA and De-noising in Feature Spaces“. In: *Proceedings of the 1998 Conference on Advances in Neural Information Processing Systems II*. Cambridge, MA, USA: MIT Press, S. 536–542. isbn: 0-262-11245-0. <http://dl.acm.org/citation.cfm?id=340534.340729>.
- Mika, S., B. Schölkopf, A. Smola, K. R. Müller, M. Scholz & G. Rätsch (1999b). „Kernel PCA and de-noising in feature spaces“. In: *Proceedings of the 1998 conference on Advances in neural information processing systems II*. Cambridge, MA, USA: MIT Press, S. 536–542. isbn: 0-262-11245-0.
- Milano, F. (2010). *Power System Modelling and Scripting*. Power Systems. Springer. isbn: 9783642136696. <http://books.google.de/books?id=MQu7IqoLrfYC>.
- Minsky, M. & S. Papert (1969). *Perceptrons: An Introduction to Computational Geometry*. MIT Press. <http://books.google.de/books?id=36E0QgAACAAJ>.
- Mittelsdorf, M., A. Hüwel, T. Klingenberg & M. Sonnenschein (2013). „Submeter based Training of Multi-class Support Vector Machines for Appliance Recognition in Home Electricity Consumption Data.“ In: *SMARTGREENS*. Hrsg. von B. Donnellan, J. F. Martins, M. Helfert & K.-H. Krempels. SciTePress, S. 151–158. isbn: 978-989-8565-55-6. <http://dblp.uni-trier.de/db/conf/smartgreens/smartgreens2013.html#MittelsdorfHKS13>.
- Mizuno, K., Y. Fukui & S. Nishihara (Jan. 2008). „Urban Traffic Signal Control Based on Distributed Constraint Satisfaction“. In: *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, S. 65–65. DOI:10.1109/HICSS.2008.482.
- Moore, D. S. & G. P. McCabe (1993). *Introduction to the Practice of Statistics*. Second. New York: Freeman.
- Moya, M. M., M. W. Koch & L. D. Hostetler (1993). „One-class classifier networks for target recognition applications“. In: *NASA STI/Recon Technical Report N 93*, S. 24043.
- Mühlbacher, H. (Juni 2007). „Verbrauchsverhalten von Wärmepumpen bei dynamisch variierten Lasten und Übertragungskomponenten“. Dissertation. Lehrstuhl für Energie-wirtschaft und Anwendungstechnik der Technischen Universität München.
- Müller, K. (2013). „Design for Environment“. English. In: *Organizations' Environmental Performance Indicators*. Hrsg. von A. Dada, K. Stanoevska & J. Marx Gómez. Environmental Science and Engineering. Springer Berlin Heidelberg, S. 45–52. isbn: 978-3-642-32719-3. DOI:10.1007/978-3-642-32720-9\_4.
- Murty, K. G. & S. N. Kabadi (Nov. 1987). „Some NP-complete Problems in Quadratic and Nonlinear Programming“. In: *Math. Program.* 39.2, S. 117–129. issn: 0025-5610. DOI:10.1007/BF02592948.
- Myers, R. H. & D. C. Montgomery (1995). *Response Surface Methodology: Process and Product in Optimization Using Designed Experiments*. 1st. New York, NY, USA: John Wiley & Sons, Inc. isbn: 0471581003.
- Nabe, C., B. Hasche, M. Offermann, G. Papaefthymiou, F. Seefeldt, N. Thamling & H. Dziomba (Okt. 2011). *Potenziale der Wärmepumpe zum Lastmanagement im Strom und*

- zur Netzintegration erneuerbarer Energien. BMWi Vorhaben Nr. 50/10 Projektnummer: PSUPDE101686. ECOFYS Köln, Berlin: Ecofys Germany GmbH, Prognos AG.
- Nadaraya, E. (1964). „On estimating regression“. In: *Theory of Prob. and Appl.* 9, S. 141–142.
- Neddermeijer, H. G., G. J. van Oortmarssen, N. Piersma & R. Dekker (2000). „A framework for Response Surface Methodology for simulation optimization“. In: *Proceedings of the 32nd conference on Winter simulation*. WSC '00. Orlando, Florida, USA: Society for Computer Simulation International, S. 129–136. isbn: 0-7803-6582-8. <http://dl.acm.org/citation.cfm?id=510378.510401>.
- Negenborn, R. R., Z. Lukszo & H. Hellendoorn, (Hrsg.) (2010). *Intelligent Infrastructures*. Bd. 42. Intelligent Systems, Control and Automation: Science and Engineering. Springer. isbn: 978-90-481-3597-4. DOI:10.1007/978-90-481-3598-1.
- Nelder, J. A. & R. Mead (1965). „A Simplex Method for Function Minimization“. In: *Computer Journal* 7, S. 308–313.
- Nieße, A. & M. Sonnenschein (2013). „Using Grid Related Cluster Schedule Resemblance for Energy Rescheduling - Goals and Concepts for Rescheduling of Clusters in Decentralized Energy Systems.“ In: *SMARTGREENS*. Hrsg. von B. Donnellan, J. F. Martins, M. Helfert & K.-H. Krempels. SciTePress, S. 22–31. isbn: 978-989-8565-55-6. <http://dblp.uni-trier.de/db/conf/smartgreens/smartgreens2013.html#NiesseS13>.
- Nielsen, H. A., H. Madsen & T. S. Nielsen (2001). *Load scheduling for decentralized CHP plants*. Research Report. Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby: Informatics und Mathematical Modelling, Technical University of Denmark, DTU. <http://www.imm.dtu.dk/~han/pub/efp98akk.pdf>.
- Nieße, A., S. Lehnhoff, M. Tröschel, M. Uslar, C. Wissing, H.-J. Appelrath & M. Sonnenschein (Juni 2012a). „Market-Based Self-Organized Provision of Active Power and Ancillary Services“. In: IEEE.
- Nieße, A., S. Lehnhoff, M. Tröschel, M. Uslar, C. Wissing, H.-J. Appelrath & M. Sonnenschein (2012b). „Market-based self-organized provision of active power and ancillary services: An agent-based approach for Smart Distribution Grids“. In: *COMPENG*. IEEE, S. 1–5. isbn: 978-1-4673-1614-9.
- Nieße, A., M. Tröschel & M. Sonnenschein (2013). „Designing dependable and sustainable Smart Grids – How to apply Algorithm Engineering to distributed control in power systems“. In: *Environmental Modelling & Software*. DOI:<http://dx.doi.org/10.1016/j.envsoft.2013.12.003>.
- Nieße, A., S. Beer, J. Bremer, C. Hinrichs, O. Lünsdorf & M. Sonnenschein (Sep. 2014). „Conjoint dynamic aggregation and scheduling for dynamic virtual power plants“. In: *Federated Conference on Computer Science and Information Systems - FedCSIS 2014, Warsaw, Poland*. Hrsg. von M. Ganzha, L. A. Maciaszek & M. Paprzycki.
- Nikonowicz, Ł. B. & J. Milewski (2012). „Virtual Power Plants – general review: structure, application and optimization.“ In: *Journal of Power Technologies* 92.3. <http://papers.itc.pw.edu.pl/index.php/JPT/article/view/284/492>.

- Nischler, G., C. Gutsch, M. Beermann & H. Stigler (2011). „Auswirkungen von Elektromobilität auf das Energiesystem“. In: *Elektrotechnik und Informationstechnik* 128.1-2, S. 53–57.
- Nitsch, J., W. Krewitt, M. Nast, P. Viebahn, S. Gärtner, M. Pehnt, G. Reinhardt, R. Schmidt, A. Uihlein, C. Barthel & M. F. and Frank Merten (März 2004). *Ökologisch optimierter Ausbau der Nutzung erneuerbarer Energien in Deutschland*. Forschungsbericht FKZ 901 41 803. Stuttgart, Heidelberg, Wuppertal: Arbeitsgemeinschaft DLR/IFEU/WI. [http://ifeu.de/landwirtschaft/pdf/Oekologisch\\_optimierter\\_Ausbau\\_Langfassung.pdf](http://ifeu.de/landwirtschaft/pdf/Oekologisch_optimierter_Ausbau_Langfassung.pdf).
- Omran, M. G. & M. Mahdavi (2008). „Global-best harmony search.“ In: *Appl. Math. Comput.* 198.2, S. 643–656. DOI:10.1016/j.amc.2007.09.004.
- Ostermeier, A., A. Gawelczyk & N. Hansen (1994). „A Derandomized Approach to Self-Adaptation of Evolution Strategies“. In: *Evolutionary Computation* 2.4, S. 369–380.
- Osuna, E., R. Freund & F. Girosi (Sep. 1997). „An improved training algorithm for support vector machine“. In: *Neural Networks for Signal Processing [1997] VII. Proceedings of the 1997 IEEE Workshop*. Anelea Island: IEEE, S. 276–285.
- Page, B., A. G. Fleischer, J. Göbel & V. Wohlgemuth, (Hrsg.) (2013). *27th International Conference on Environmental Informatics for Environmental Protection, Sustainable Development and Risk Management, EnviroInfo 2013, Hamburg, Germany, September 2-4, 2013. Proceedings*. Berichte aus der Umweltinformatik. Shaker. isbn: 978-3-8440-1676-5.
- Palensky, P. (2001). *Distributed Reactive Energy Management*. Dissertation. Technischen Universität Wien.
- Palmer, C. C. & A. Kershenbaum (1994). „Representing Trees in Genetic Algorithms“. In: *Proceedings of the First IEEE Conference on Evolutionary Computation*. IEEE Press, S. 379–384.
- Pandiaraj, K., B. Fox, D. Morrow, S. Persaud & J. P. Martin (März 2002). „Centralised control of diesel gen-sets for peak shaving and system support“. In: *Generation, Transmission and Distribution, IEE Proceedings-* 149.2, S. 126–132. issn: 1350-2360. DOI:10.1049/ip-gtd:20020277.
- Pantoja, A., N. Quijano & K. Passino (Dez. 2011). „Dispatch of distributed generators using a local replicator equation“. In: *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, S. 7494–7499. DOI:10.1109/CDC.2011.6160627.
- Papadonikolakis, M. & C. Bouganis (Dez. 2010). „A novel FPGA-based SVM classifier“. In: *Field-Programmable Technology (FPT), 2010 International Conference on*, S. 283–286. DOI:10.1109/FPT.2010.5681485.
- Park, J., D. Kang, J. T. Kwok, S.-W. Lee, B.-W. Hwang & S.-W. Lee (2006). „Facial Image Reconstruction by SVDD-Based Pattern De-noising.“ In: *ICB*. Hrsg. von D. Zhang & A. K. Jain. Bd. 3832. Lecture Notes in Computer Science. Springer, S. 129–135. isbn: 3-540-31111-4. <http://dblp.uni-trier.de/db/conf/icb/icb2006.html#ParkKKLHL06>.

- Park, J., D. Kang, J. Kim, J. T. Kwok & I. W. Tsang (2007). „SVDD-Based Pattern Denoising“. In: *Neural Computing* 19.7, S. 1919–1938. issn: 0899-7667. DOI:<http://dx.doi.org/10.1162/neco.2007.19.7.1919>.
- Parzen, E. (1962). „On Estimation of a Probability Density Function and Mode“. English. In: *The Annals of Mathematical Statistics* 33.3, S. 1065–1076. issn: 00034851. <http://www.jstor.org/stable/2237880>.
- Peer, E. (2004). *A Serendipitous Software Framework for Facilitating Collaboration in Computational Intelligence*. University of Pretoria. <http://books.google.de/books?id=oZ2gHAAACAAJ>.
- Peters, W., L. Morkel, J. Köppel & J. Köller (März 2008). *Berücksichtigung von Auswirkungen auf die Meeresumwelt bei der Zulassung von Windparks in der Ausschließlichen Wirtschaftszone*. Endbericht eines Forschungsvorhabens gefördert aus Mitteln des Bundesumweltministeriums (FKZ 0329949). Unter Mitarbeit von K. Wippel, Z. Hagen und M. Treblin, mit einem Beitrag von Lothar Bach und Ulf Rahmel. Berlin: Institut für Landschaftsarchitektur und Umweltplanung Fachgebiet Landschaftsplanung, insb. Landschaftspflegerische Begleitplanung/ Umweltverträglichkeitsprüfung. [http://www.umweltpruefung.tu-berlin.de/v-menue/forschung/beendete\\_projekte/beruecksichtigung\\_von\\_auswirkungen\\_auf\\_die\\_meeresumwelt\\_bei\\_der\\_zulassung\\_von\\_windparks\\_in\\_der\\_ausschliesslichen\\_wirtschaftszone](http://www.umweltpruefung.tu-berlin.de/v-menue/forschung/beendete_projekte/beruecksichtigung_von_auswirkungen_auf_die_meeresumwelt_bei_der_zulassung_von_windparks_in_der_ausschliesslichen_wirtschaftszone).
- Pielke, M. & M. Kurrat (Apr. 2008). „Potentials for asset optimizations in low and medium voltage networks through grid oriented feeding of distributed generation“. In: *Transmission and Distribution Conference and Exposition, 2008. T & D. IEEE/PES*, S. 1–7. DOI:[10.1109/TDC.2008.4517231](https://doi.org/10.1109/TDC.2008.4517231).
- Platt, G. (2007). „The Decentralised Control of Electricity Networks – Intelligent and Self-Healing Systems“. In: *Grid-Interop Forum 2007*.
- Platt, J. (1999). „Fast training of support vector machines using sequential minimal optimization“. In: *Advances in Kernel Methods*. MIT press, S. 185–208.
- Platt, J. C. (Apr. 1998). *Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines*. technical report MSR-TR-98-14. Redmond: Microsoft Research.
- Powers, D. M. W. (2008). „Evaluation Evaluation“. In: *Proceedings of the 2008 Conference on ECAI 2008: 18th European Conference on Artificial Intelligence*. Amsterdam, The Netherlands, The Netherlands: IOS Press, S. 843–844. isbn: 978-1-58603-891-5. <http://dl.acm.org/citation.cfm?id=1567281.1567498>.
- Pozdnoukhov, A. & M. Kanevski (2007). *Multi-scale support vector regression for hotspot detection and modeling*. Research Report 06-007. University of Lausanne, Switzerland: Institute of Geomatics und Analysis of Risk.
- Psola, J.-H., W.-R. Canders & M. Henke (2013). „Modeling of a Redox Flow Battery Storage for Grid Applications“. In: *PCIM Asia, Shanghai, China*. Stuttgart: Mesago PCIM GmbH.
- Radulescu, A., M. López-Ibáñez & T. Stützle (2013). „Automatically Improving the Anytime Behaviour of Multiobjective Evolutionary Algorithms“. English. In: *Evolutionary Multi-Criterion Optimization*. Hrsg. von R. Purshouse, P. Fleming, C. Fonseca, S. Greco & J.

- Shaw. Bd. 7811. Lecture Notes in Computer Science. Springer Berlin Heidelberg, S. 825–840. isbn: 978-3-642-37139-4. DOI:[10.1007/978-3-642-37140-0\\_61](https://doi.org/10.1007/978-3-642-37140-0_61).
- Ramchurn, S. D., P. Vytelingum, A. Rogers & N. R. Jennings (Juli 2011). „Agent-based Homeostatic Control for Green Energy in the Smart Grid“. In: *ACM Trans. Intell. Syst. Technol.* 2.4, 35:1–35:28. issn: 2157-6904. DOI:[10.1145/1989734.1989739](https://doi.org/10.1145/1989734.1989739).
- (Apr. 2012). „Putting the ‘Smarts’ into the Smart Grid: A Grand Challenge for Artificial Intelligence“. In: *Commun. ACM* 55.4, S. 86–97. issn: 0001-0782. DOI:[10.1145/2133806.2133825](https://doi.org/10.1145/2133806.2133825). <http://doi.acm.org/10.1145/2133806.2133825>.
- Rapoport, A. (1970). *N-person game theory. Concepts and applications*. Ann Arbor science library. Univ. of Michigan Pr, 331 p. isbn: 9780486414553.
- Rapp, B. & J. Bremer (2012). „Design of an Event Engine for Next Generation CEMIS: A Use Case“. In: *EnvirolInfo 2012 – 26th International Conference on Informatics for Environmental Protection*. Hrsg. von W. P. E. Hans-Knud Arndt Gerlinde Knetsch. ISBN 978-3-8440-1248-4. Shaker Verlag, S. 753–760.
- (2013). „IT Solutions for EPI Management“. English. In: *Organizations’ Environmental Performance Indicators*. Hrsg. von A. Dada, K. Stanoevska & J. Marx Gómez. Environmental Science and Engineering. Springer Berlin Heidelberg, S. 19–31. isbn: 978-3-642-32719-3. DOI:[10.1007/978-3-642-32720-9\\_2](https://doi.org/10.1007/978-3-642-32720-9_2).
- Rapp, B., A. Solsbach, T. Mahmoud, A. Memari & J. Bremer (2011). „IT-for-Green: Next Generation CEMIS for Environmental, Energy and Resource Management“. In: *EnvirolInfo 2011 - Innovations in Sharing Environmental Observation and Information, Proceedings of the 25th EnvirolInfo Conference ‘Environmental Informatics’*. Hrsg. von W. Pillmann, S. Schade & P. Smits. ISBN 978-3-8440-0451-9. Shaker Verlag, S. 573 –581.
- Rau, M. (2003). „Nichtlineare modellbasierte prädiktive Regelung auf Basis lernfähiger Zustandsraummodelle“. Dissertation. München: Technische Universität München.
- Rautenstrauch, C. (1999). *Betriebliche Umweltinformationssysteme*. Berlin Heidelberg: Springer-Verlag.
- Ray, T., T. Kang & S. K. Chye (2000). „An Evolutionary Algorithm for Constrained Optimization.“ In: *GECCO*. Hrsg. von L. D. Whitley, D. E. Goldberg, E. Cantú-Paz, L. Spector, I. C. Parmee & H.-G. Beyer. Morgan Kaufmann, S. 771–777. isbn: 1-55860-708-0.
- Rechenberg, I. (1965). *Cybernetic solution path of an experimental problem*. Techn. Ber. Royal Air Force Establishment.
- Reiner, U., T. Leibfried, F. Allerding & H. Schmeck (Sep. 2009). „Potenzial rückspeisefähiger Elektrofahrzeuge und steuerbarer Verbraucher im Verteilnetz unter Verwendung eines dezentralen Energiemanagementsystems“. In: *Internationaler ETG-Kongress 2009 (ETG-FB 118)*. Hrsg. von ETG. Berlin-Offenbach: VDE, S. 329–334.
- Richardson, J. T., M. R. Palmer, G. E. Liepins & M. R. Hilliard (1989). „Some Guidelines for Genetic Algorithms with Penalty Functions“. In: *Proceedings of the 3rd International Conference on Genetic Algorithms*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., S. 191–197. isbn: 1-55860-066-3.

- Rohjans, S. & M. Specht (2013). „OPC UA: An Automation Standard for Future Smart Grids“. In: *Standardization in Smart Grids*. Power Systems. Springer Berlin Heidelberg, S. 191–209. isbn: 978-3-642-34915-7. DOI:[10.1007/978-3-642-34916-4\\_12](https://doi.org/10.1007/978-3-642-34916-4_12).
- Rohlfing, D. (2012). „Schöne neue Smart Energy Welt – Ansichten einer Technik im Übergang“. In: *Smart Energy*. Hrsg. von H.-G. Servatius, U. Schneidewind & D. Rohlfing. Springer Berlin Heidelberg, S. 63–77. isbn: 978-3-642-21820-0. [http://dx.doi.org/10.1007/978-3-642-21820-0\\_3](http://dx.doi.org/10.1007/978-3-642-21820-0_3).
- Rong, A. & R. Lahdelma (Juli 2004). *Efficient Algorithms for Optimizing Combined Heat and Power Production Under the Electricity Market*. technical report 615. Turku Centre for Computer Science, Finland.
- Rosenblatt, M. (1956). „Remarks on Some Nonparametric Estimates of a Density Function“. In: *The Annals of Mathematical Statistics* 27.3, pp. 832–837. issn: 00034851. <http://www.jstor.org/stable/2237390>.
- Rossi, F., P. Van Beek & T. Walsh (2006). *Handbook of Constraint Programming*. Foundations of artificial intelligence Bd. 35. Elsevier. isbn: 9780444527264. <http://books.google.de/books?id=VfRGngEACAAJ>.
- Ruiz-Llata, M., G. Guarnizo & M. Yébenes-Calvino (2010). „FPGA implementation of a support vector machine for classification and regression.“ In: *IJCNN*. IEEE, S. 1–5. isbn: 978-1-4244-6916-1.
- Ruiz, F. J., C. Angulo & N. Agell (2006). „Kernel Machines for Continuous and Discrete Variables“. In: *Knowledge-Based Intelligent Information and Engineering Systems, 10th International Conference*. Bournemouth, UK: KES 2006.
- Runarsson, T. P. & X. Yao (2000). „Stochastic ranking for constrained evolutionary optimization.“ In: *IEEE Trans. Evolutionary Computation* 4.3, S. 284–294.
- Sachenbacher, M. & B. C. Williams (Okt. 2005). „Solving soft constraints by separating optimization and satisfiability“. In: *In Proceedings of the Seventh International Workshop on Preferences and Soft Constraints*. Sitges, Spain.
- Salomon, R. (1995). „Reevaluating Genetic Algorithm Performance under Coordinate Rotation of Benchmark Functions - A survey of some theoretical and practical aspects of genetic algorithms“. In: *BioSystems* 39, S. 263–278.
- Sammon, J. (1969). „A Nonlinear Mapping for Data Structure Analysis“. In: *IEEE Transactions on Computers* 18.5, S. 401–409. issn: 0018-9340. DOI:<http://doi.ieeecomputersociety.org/10.1109/T-C.1969.222678>.
- Saravanan, B., S. Das, S. Sikri & D. Kothari (2013). „A solution to the unit commitment problem: a review“. English. In: *Frontiers in Energy* 7.2, S. 223–236. issn: 2095-1701. DOI:[10.1007/s11708-013-0240-3](https://doi.org/10.1007/s11708-013-0240-3).
- Schaback, R. & H. Wendland (Mai 2006). „Kernel techniques: From machine learning to meshless methods“. In: *Acta Numerica* 15, S. 543–639. issn: 1474-0508. DOI:[10.1017/S0962492906270016](https://doi.org/10.1017/S0962492906270016).

- Schild, K. (2014). *Dynamische Optimierung*. Vorlesungsskript. Universität Marburg. <https://www.uni-marburg.de/fb02/statistik/studium/vorl/dynopt> (besucht am 21.05.2014).
- Schlebusch, V., M. Wolff, D. Nestle, G. Gjardy, T. Borchard, A. Bukvic-Schäfer, T. Erge, M. Klobasa & M. Hollmann (Nov. 2006). „Energiemanagement in Verteilnetzen mit hohem Anteil an dezentralen Erzeugungsanlagen“. In: *Tagungsband Kasseler Symposium Energie-Systemtechnik*. Kassel, Germany.
- Schölkopf, B. (1997). „Support Vector Learning“. Dissertation. Oldenbourg Verlag, München: Fachbereich 13 Informatik der Technischen Universität Berlin.
- Schönberg, I. (2012). „Smart durch Kommunikation“. German. In: *Smart Energy*. Hrsg. von H.-G. Servatius, U. Schneidewind & D. Rohlfing. Springer Berlin Heidelberg, S. 379–392. isbn: 978-3-642-21819-4. DOI:10.1007/978-3-642-21820-0\_24.
- Schölkopf, B., S. Mika, C. Burges, P. Knirsch, K.-R. Müller, G. Rätsch & A. Smola (1999a). „Input space vs. feature space in kernel-based methods“. In: *IEEE Transactions on Neural Networks* 10(5), S. 1000–1017.
- Schölkopf, B., R. C. Williamson, A. J. Smola, J. Shawe-Taylor & J. C. Platt (1999b). „Support Vector Method for Novelty Detection“. In: *NIPS*. Hrsg. von S. A. Solla, T. K. Leen & K.-R. Müller. The MIT Press, S. 582–588. isbn: 0-262-19450-3.
- Schölkopf, B., J. Giesen & S. Spalinger (2004). *Kernel Methods for Implicit Surface Modeling*. Technical Report TR-125. Tübingen, Germany: Max Planck Institute for Biological Cybernetics.
- Schott, B., A. Püttner, T. Nieder, F. Maas genannt Bermpohl, M. Rohn & J. Mey (2013). *Entwicklung der Elektromobilität in Deutschland im internationalen Vergleich und Analysen zum Stromverbrauch*. Techn. Ber. Zentrum für Sonnenenergie- und Wasserstoff-Forschung Baden-Württemberg / DFKI GmbH.
- Schwaegerl, C. & L. Tao (2014). „Microgrids: Architectures and Control“. In: Hrsg. von N. Hatziargyriou. 1st. John Wiley und Sons Ltd. Kap. The Microgrids Concept, S. 1–24.
- Schwefel, H., (Hrsg.) (1995). *Evolution and optimum seeking. Evolution and optimum seeking / Schwefel, Hans-Paul. - New York [u.a.] : Wiley, 1995 Buch*. A Wiley-Interscience publication. New York [u.a.]: Wiley. XI, 444 s.
- Scott, D. W. (Sep. 1992). *Multivariate Density Estimation: Theory, Practice, and Visualization (Wiley Series in Probability and Statistics)*. 1. Aufl. Wiley. isbn: 0471547700. <http://www.worldcat.org/isbn/0471547700>.
- Seebach, D., C. Timpe & D. Bauknecht (Sep. 2009). *Costs and Benefits of Smart Appliances in Europe: A report prepared as part of the EIE project „Smart Domestic Appliances in Sustainable Energy Systems (Smart-A)“*. D 7.2 of WP 7 from the Smart-A project D 7.2. Freiburg: Öko-Institut e.V.
- Sen, S. K., M. Foskey, J. S. Marron & M. A. Styner (2008). „Support vector machine for data on manifolds: An application to image analysis.“ In: *ISBI*. IEEE, S. 1195–1198. <http://dblp.uni-trier.de/db/conf/isbi/isbi2008.html#SenFMS08>.

- Settles, B. (2012). *Active Learning*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers.
- Shah, A. (2012a). *On Coordinating Electricity Markets: Smart Power Scheduling for Demand Side Management and Economic Dispatch*. Harvard University.
- Shah, A. D. (2012b). „On Coordinating Electricity Markets: Smart Power Scheduling for Demand Side Management and Economic Dispatch“. PhD Theses. Cambridge, MA: Applied Mathematics, Harvard College. [http://www.eecs.harvard.edu/econcs/pubs/ashah\\_thesis.pdf](http://www.eecs.harvard.edu/econcs/pubs/ashah_thesis.pdf).
- Shahid, N., S. A. Aleem, I. H. Naqvi & N. Zaffar (2012). „Support Vector Machine based fault detection & classification in smart grids“. In: *GLOBECOM Workshops*. IEEE, S. 1526–1531. isbn: 978-1-4673-4942-0.
- Shannon, C. E. (1948). „A Mathematical Theory of Communication“. In: *The Bell System Technical Journal* 27, S. 379–423, 623–656. <http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf>.
- Shapley, L. S. (1953). „A value for n-person games“. In: *Contributions to the theory of games* 2. Hrsg. von Kuhn, S. 307–317.
- Shapley, L. (1967). *Utility Comparison and the Theory of Games*. Techn. Ber. Rand Corporation. <http://www.rand.org/pubs/papers/P3582.html>.
- Shawe-Taylor, J. & N. Cristianini (2004). *Kernel Methods for Pattern Analysis*. New York, NY, USA: Cambridge University Press. isbn: 0521813972.
- Shi, Y. & R. C. Eberhart (Mai 1998). „A Modified Particle Swarm Optimizer“. In: *Proceedings of IEEE International Conference on Evolutionary Computation*. Washington, DC, USA: IEEE Computer Society, S. 69–73. DOI:10.1109/ICEC.1998.699146.
- Silverman, B. W. (Apr. 1986). *Density Estimation for Statistics and Data Analysis (Chapman & Hall/CRC Monographs on Statistics & Applied Probability)*. 1. Aufl. Chapman und Hall/CRC. isbn: 0412246201. <http://www.worldcat.org/isbn/0412246201>.
- Simon, D. (2013). *Evolutionary Optimization Algorithms*. Wiley. isbn: 9781118659502.
- Smith, A. & D. Coit (1997). „Handbook of Evolutionary Computation“. In: Department of Industrial Engineering, University of Pittsburgh, USA: Oxford University Press und IOP Publishing. Kap. Penalty Functions, Section C5.2.
- Smola, A. J. & B. Schölkopf (Aug. 2004). „A Tutorial on Support Vector Regression“. In: *Statistics and Computing* 14.3, S. 199–222. issn: 0960-3174. DOI:10.1023/B:STC0.0000035301.49549.88.
- Soares, J. P., T. M. Sousa, H. Morais, Z. Vale, B. Canizes & A. Silva (2013). „Application-Specific Modified Particle Swarm Optimization for energy resource scheduling considering vehicle-to-grid.“ In: *Appl. Soft Comput.* 13.11, S. 4264–4280. <http://dblp.uni-trier.de/db/journals/asc/asc13.html#SoaresSMVCS13>.
- Solsbach, A., R. Isenmann & J. Marx Gómez (2011). „iSTORM – Idea and Reference Architecture approaching Inter-organisational sustainability Reporting“. In: *Innovations*

- in Sharing Environmental Observations and Information (Part 2)*. Hrsg. von P. S. E. W. Pillmann S. Schade. ISBN 978-3-8440-0451-9. Shaker Verlag, S. 639–646.
- Soltani, M., M. Raoofat & M. A. Rostami (Mai 2012). „Optimal reliable strategy of virtual power plant in energy and frequency control markets“. In: *Electrical Power Distribution Networks (EPDC), 2012 Proceedings of 17th Conference on*, S. 1–6.
- Somorjai, R. L., B. Dolenko, A. Demko, M. Mandelzweig, A. E. Nikulin, R. Baumgartner & N. J. Pizzi (2004). „Mapping high-dimensional data onto a relative distance plane: an exact method for visualizing and characterizing high-dimensional patterns“. In: *J. of Biomedical Informatics* 37.5, S. 366–379. issn: 1532-0464. DOI:<http://dx.doi.org/10.1016/j.jbi.2004.07.005>.
- Sonnenschein, M., B. Rapp & J. Bremer (2010). „Demand Side Management und Demand Response“. In: *Handbuch Energiemanagement* 30. Hrsg. von H.-P. Beck, J. Buddenberg, E. Meller & C. Salander.
- (Dez. 2011). „Demand Side Management und Demand Response (Neufassung)“. In: *Handbuch Energiemanagement* 31. Hrsg. von H.-P. Beck, J. Buddenberg, E. Meller & C. Salander.
- Sonnenschein, M., M. Tröschel & O. Lünsdorf (2013). „Smart Grids for Optimised Utilisation of Renewable Energy Supply“. In: *EnviroInfo*. Hrsg. von B. Page, A. G. Fleischer, J. Göbel & V. Wohlgemuth. Berichte aus der Umwelthinformatik. Shaker, S. 178–187. isbn: 978-3-8440-1676-5.
- Sonnenschein, M., O. Lünsdorf, J. Bremer & M. Tröschel (2014). „Decentralized Control of Units in Smart Grids for the Support of Renewable Energy Supply“. In: *Environmental Impact Assessment Review* 0. in press, S. –. issn: 0195-9255.
- Sörensen, K. (2013). „Metaheuristics – the metaphor exposed“. In: *International Transactions in Operational Research*. issn: 1475-3995. DOI:[10.1111/itor.12001](https://doi.org/10.1111/itor.12001).
- Squire, W. (1975). „Computer implementation of the Schwarz-Christoffel transformation“. In: *Journal of the Franklin Institute* 299.5, S. 315–322.
- Stadler, M., W. Krause, M. Sonnenschein & U. Vogel (Jan. 2007). „The Adaptive Fridge – Comparing different control schemes for enhancing load shifting of electricity demand.“ In: *21st Conference on Informatics for Environmental Protection – Enviroinfo Warsaw 2007*. Shaker Verlag, S. 199–206. isbn: 978-3-8322-6397-3.
- Stadler, M. (Nov. 2003). „The relevance of demand-side-measures and elastic demand curves to increase market performance in liberalized electricity markets: The case of Austria“. Dissertation. TU Wien: Institut für Elektrische Anlagen und Energiewirtschaft.
- Stadler, P. D.-I. I. (Okt. 2005). „Demand Response: Nichtelektrische Speicher für Elektrizitätsversorgungssysteme mit hohem Anteil erneuerbarer Energien“. Habilitation. Fachbereich Elektrotechnik, Universität Kassel.
- Steck, M. H. E. (2013). „Entwicklung und Bewertung von Algorithmen zur Einsatzplanerstellung virtueller Kraftwerke“. Dissertation. München: Technische Universität München.
- Steinke, F., B. Schölkopf & V. Blanz (2005). „Support vector machines for 3D shape processing. Eurographics“. In: *Computer Graphics forum*. Bd. 24. 3.

- Steinwart, I. & A. Christmann (2008). *Support Vector Machines*. 1st. Springer Publishing Company, Incorporated. isbn: 0387772413.
- Strese, J. (2006). „Neue Chancen in Regelenergiemärkten mit virtuellen Kraftwerken“. In: *Tagungsband der VWEW Fachtagung Smart Grids - der Beitrag virtueller Kraftwerke zur nachhaltigen Energieversorgung*. VWEW06. Fulda.
- Stromnetz Hamburg (2014). *Standardlastprofile für 2014*. <http://www.stromnetz-hamburg.de/de/stromversorger.htm> (besucht am 04.03.2014).
- Sudeikat, J. O. & O. Heitmann (2013). „Towards Modular Assembling of Virtual Power Plant Control Systems – The Smart Power Hamburg Platform.“ In: *EnviroInfo*. Hrsg. von B. Page, A. G. Fleischer, J. Göbel & V. Wohlgemuth. Berichte aus der Umweltinformatik. Shaker, S. 188–197. isbn: 978-3-8440-1676-5. <http://dblp.uni-trier.de/db/conf/enviroinfo/enviroinfo2013.html#SudeikatH13>.
- Surry, P. D. & N. J. Radcliffe (1997). „The COMOGA method: Constrained optimisation by multiobjective genetic algorithms“. In: *Control and Cybernetics* 26.3, S. 391–412.
- Szegö, G. (1978). „Towards global optimisation 2“. In: Hrsg. von L. Dixon & G. Szegö. *Towards global optimisation / eds L. C. W. Dixon and G. P. Szegö*. North-Holland Pub. Co. Kap. The global optimization problem: an introduction, S. 1–15. isbn: 9780444851710. <http://books.google.co.uk/books?id=SUSrAAAAIAAJ>.
- Tavakkoli, A., M. Nicolescu, G. Bebis & M. N. Nicolescu (2008a). „Efficient background modeling through incremental Support Vector Data Description.“ In: *ICPR*. IEEE. isbn: 978-1-4244-2175-6. <http://dblp.uni-trier.de/db/conf/icpr/icpr2008.html#TavakkoliNBN08>.
- Tavakkoli, A., M. Nicolescu, M. Nicolescu & G. Bebis (2008b). „Incremental SVDD Training: Improving Efficiency of Background Modeling in Videos“. In: *Signal and Image Processing*. Hrsg. von P. Cristea. Calgary, Canada: Acta Press.
- Tax, D. M. J. & R. P. W. Duin (1999). „Data Domain Description using Support Vectors“. In: *European Symposium on Artificial Neural Networks – ESANN*, S. 251–256.
- (2004). „Support Vector Data Description“. In: *Mach. Learn.* 54.1, S. 45–66. issn: 0885-6125. DOI:<http://dx.doi.org/10.1023/B:MACH.0000008084.60811.49>.
- Teuteberg, F. & J. Marx Gómez (2010). „Green Computing & Sustainability - Status quo und Herausforderungen für BUIS der nächsten Generation“. In: *Praxis der Wirtschaftsinformatik* 47.274, S. 6–17.
- Thoma, M. (2007). „Optimierte Betriebsführung von Niederspannungsnetzen mit einem hohen Anteil an dezentraler Erzeugung“. Dissertation. Zürich, Schweiz: Eidgenössische Technische Hochschule.
- Thomas, B. (2007). *Mini-Blockheizkraftwerke: Grundlagen, Gerätetechnik, Betriebsdaten*. Vogel Buchverlag.
- Tikhonov, A. N. & V. Y. Arsenin (1977). *Solutions of ill-posed problems*. Translated from the Russian, Preface by translation editor Fritz John, Scripta Series in Mathematics. Washington, D.C.: John Wiley & Sons, New York: V. H. Winston & Sons, S. xiii+258.

- Toffler, A. (1970). *Future shock*. Future Shock Bd. 644. New York: Random House. isbn: 9780394425863.
- Treur, J. (2013). „Dynamic Allocation of a Domestic Heating Task to Gas-Based and Heatpump-Based Heating Agents“. In: *Multiagent System Technologies*. Hrsg. von M. Klusch, M. Thimm & M. Paprzycki. Bd. 8076. Lecture Notes in Computer Science. Springer Berlin Heidelberg, S. 386–399. isbn: 978-3-642-40775-8. DOI:[10.1007/978-3-642-40776-5\\_32](https://doi.org/10.1007/978-3-642-40776-5_32).
- Tröschel, M. (2010). „Aktive Einsatzplanung in holonischen Virtuellen Kraftwerken“. Dissertation. Oldenburg: Carl von Ossietzky Universität.
- Tröschel, M. & H.-J. Appelrath (Sep. 2009). „Towards Reactive Scheduling for Large-Scale Virtual Power Plants.“ In: *MATES*. Hrsg. von L. Braubach, W. van der Hoek, P. Petta & A. Pokahr. Bd. 5774. Lecture Notes in Computer Science. Springer, S. 141–152. isbn: 978-3-642-04142-6.
- Tröschel, M., S. Scherfke, S. Schütte, A. Nieße & M. Sonnenschein (Okt. 2011). „Using Electric Vehicle Charging Strategies to Maximize PV-Integration in the Low Voltage Grid“. In: *Proceeding of: Internationale Konferenz und Ausstellung zur Speicherung Erneuerbarer Energien (IRES 2011)*. Berlin.
- Uslar, M., T. Schmedes, A. Lucks, T. Luhmann, L. Winkels & H.-J. Appelrath (Sep. 2005). „Interaction of EMS related systems by using the CIM standard“. In: *Proceedings of the ITEE 2005 Magdeburg*. Shaker Verlag, Aachen, S. 596–610.
- Vapnik, V. N. & A. Y. Chervonenkis (1974). *Theory of Pattern Recognition [in Russian]*. USSR: Nauka.
- Vapnik, V. & A. Lerner (1963). „Pattern Recognition using Generalized Portrait Method“. In: *Automation and Remote Control* 24.
- Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag New York, Inc. isbn: 0-387-94559-8.
- Vapnik, V. (1982). *Estimation of Dependences Based on Empirical Data: Springer Series in Statistics (Springer Series in Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc. isbn: 0387907335.
- Vassilev, V. K., T. C. Fogarty & J. F. Miller (März 2000). „Information Characteristics and the Structure of Landscapes“. In: *Evol. Comput.* 8.1, S. 31–60. issn: 1063-6560. DOI:[10.1162/106365600568095](https://doi.org/10.1162/106365600568095).
- Vytelingum, P., T. D. Voice, S. D. Ramchurn, A. Rogers & N. R. Jennings (2010). „Agent-based Micro-storage Management for the Smart Grid“. In: *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1 (Toronto, Canada)*. AAMAS '10. Richland, SC: International Foundation for Autonomous Agents und Multiagent Systems, S. 39–46. isbn: 978-0-9826571-1-9. <http://dl.acm.org/citation.cfm?id=1838206.1838212>.
- Wang, J., Z. Liu & P. Lu (2008). „Electricity Load Forecasting Using Rough Set Attribute Reduction Algorithm Based on Immune Genetic Algorithm and Support Vector Machines“. In: *Proceedings of the 2008 International Conference on Risk Management & Engineering*

- Management*. ICRMEM '08. Washington, DC, USA: IEEE Computer Society, S. 239–244. isbn: 978-0-7695-3402-2. DOI:10.1109/ICRMEM.2008.85.
- Watts, D. & S. Strogatz (1998). „Collective dynamics of 'small-world' networks“. In: *Nature* 393, S. 440–442.
- Wedde, H. F., S. Lehnhoff, E. Handschin & O. Krause (2007). „Dezentrale vernetzte Energiebewirtschaftung (DEZENT) im Netz der Zukunft“. In: *Wirtschaftsinformatik* 49.5, S. 361–369.
- Weidelt, T., A. Nieße, H.-J. Appelrath, M. Korte & W. Nebel (Dez. 2008). „Design and Implementation of a Standard-Compliant CHP Control Unit“. In: *Proceedings of the 3rd International Conference on Integration of Renewable and Distributed Energy Resources*.
- Weinberger, E. (1990). „Correlated and uncorrelated fitness landscapes and how to tell the difference“. English. In: *Biological Cybernetics* 63.5, S. 325–336. issn: 0340-1200. DOI:10.1007/BF00202749.
- Wissing, C. & H.-J. Appelrath (2013). „Market-Based Redispatch in Distribution Grids - Incentivizing Flexible Behavior of Distributed Energy Ressource“. In: *EnviroInfo*. Hrsg. von B. Page, A. G. Fleischer, J. Göbel & V. Wohlgemuth. Berichte aus der Umweltinformatik. Shaker, S. 244–250. isbn: 978-3-8440-1676-5.
- Witten, I. H., E. Frank & M. A. Hall (2011). *Data Mining: Practical Machine Learning Tools and Techniques*. 3. Aufl. Amsterdam: Morgan Kaufmann. isbn: 978-0-12-374856-0. <http://www.sciencedirect.com/science/book/9780123748560>.
- Woldt, T. (2008). *Beitrag zur Energiesystemintegration dezentraler Energiewandlungsanlagen mit besonderer Berücksichtigung kleiner Kraft-Wärme-Kopplungsanlagen in kommunalen Versorgungsstrukturen*. BTU-Forschungshefte Energie. dissertation.de. isbn: 9783866243620. <http://books.google.de/books?id=Lu5hngEACAAJ>.
- Wolpert, D. H. & W. G. Macready (Apr. 1997). „No Free Lunch Theorems for Optimization“. In: *Trans. Evol. Comp* 1.1, S. 67–82. issn: 1089-778X. DOI:10.1109/4235.585893.
- Wu, F., K. Moslehi & A. Bose (2005). „Power System Control Centers: Past, Present, and Future“. In: *Proceedings of the IEEE* 93.11, S. 1890–1908. issn: 0018-9219. DOI:10.1109/JPROC.2005.857499.
- Xiong, H., M. Swamy & M. Ahmad (März 2005). „Optimizing the kernel in the empirical feature space“. In: *Neural Networks, IEEE Transactions on* 16.2, S. 460–474. issn: 1045-9227. DOI:10.1109/TNN.2004.841784.
- Xiong, W., M.-j. Li & Y.-l. Cheng (2008). „An Improved Particle Swarm Optimization Algorithm for Unit Commitment“. In: *Proceedings of the 2008 International Conference on Intelligent Computation Technology and Automation - Volume 01*. ICICTA '08. Washington, DC, USA: IEEE Computer Society, S. 21–25. isbn: 978-0-7695-3357-5. DOI:10.1109/ICICTA.2008.363.
- Yan-bin, C. (2009). „The Optimization Model Based on Support Vector Machine for the Energy-Conserving Generation Dispatch“. In: *ICNC (1)*. Hrsg. von H. Wang, K. S. Low, K. Wei & J. Sun. IEEE Computer Society, S. 569–573. isbn: 978-0-7695-3736-8.

- 
- Yao, X., Y. Liu & G. Lin (1999). „Evolutionary programming made faster“. In: *Evolutionary Computation, IEEE Transactions on* 3.2, S. 82–102.
- Ygge, F. (1998). *Market-Oriented Programming and its Application to Power Load Management*. Ph.D. Thesis. <http://www.enersearch.se/knowledgebase/publications/conference-journals/thesis/thesis.html>. Lund University.
- Ygge, F., H. Akkermans, A. Andersson, M. Krejic & E. Boertjes (1999). „The HOMEBOTS System and Field Test: A Multi-Commodity Market for Predictive Power Load Management“. In: *Proceedings of the Fourth International Conference and Exhibition on The Practical Application of Intelligent Agents and Multi-Agents (PAAM99)*. <http://www.enersearch.se/knowledgebase/publications/conference-journals/PAAM99/paam99.pdf>. London, UK.
- ZVEI (1970). *ZVEI-Kennzahlensystem : ein Instrument zur Unternehmenssteuerung*. Betriebswirtschaftliche Schriftenreihe des ZVEI. Zentralverband der Elektrotechnischen Industrie. <http://ci.nii.ac.jp/ncid/BA11518445>.
- Zheng, E.-H., M. Yang, P. Li & Z.-H. Song (2006). „Fuzzy Support Vector Clustering.“ In: *ISNN (1)*. Hrsg. von J. Wang, Z. Yi, J. M. Zurada, B.-L. Lu & H. Yin. Bd. 3971. Lecture Notes in Computer Science. Springer, S. 1050–1056. isbn: 3-540-34439-X. <http://dblp.uni-trier.de/db/conf/isnn/isnn2006-1.html#ZhengYLS06>.