# Optimierungsmethoden für die Energiewirtschaft: Stand und Entwicklungstendenzen*

*Univ.-Prof. Dr. sc. nat.* **W. Römisch**, *Berlin*

**Zusammenfassung.** An einem repräsentativen Beispiel eines energiewirtschaftlichen Optimierungsproblems werden zunächst moderne Entwicklungen bei Optimierungsverfahren und zugehöriger Software besprochen für den Fall, daß die Daten fixierbar bzw. vorhersagbar sind. Genauer wird dabei auf Fortschritte im Zusammenhang mit der klassischen Methode der Lagrange Relaxation eingegangen. Im Hauptteil der Arbeit werden Gründe für den Übergang zu stochastischen Optimierungsmodellen, strukturelle Eigenschaften solcher Modelle, Dekompositionsansätze und Lösungsverfahren diskutiert. Speziell wird auf die Konstruktion von Szenariobäumen zur Modellierung der Stochastik der Datenprozesse und auf die geeignete Modellierung und Einbeziehung des Risikos von Entscheidungen eingegangen. Besonderer Wert wird auf gemischt-ganzzahlige Modelle und auf die Diskussion aktueller Literatur gelegt.

## 1 Einleitung

Die Liberalisierung der Elektrizitätswirtschaft stellt die Energieversorgungsunternehmen vor neue Herausforderungen. Die klassische Aufgabe des kostenminimalen Einsatzes der Erzeugungsanlagen wandelt sich zur simultanen Planung von Stromerzeugung und -handel. Dies bedeutet, daß die Minimierung der Erzeugungskosten und die Maximierung der Erträge aus dem Portfolio von Liefer- und Bezugsverträgen gleichzeitig, d.h., in einem gemeinsamen Modell, erfolgen muß (vgl. z.B. [35], [5], [29]). Einerseits führt dies zu einer Vergrößerung der Modelldimensionen und andererseits zur Zunahme des Einflusses der Stochastik von Daten.

Gegenstand dieser Arbeit sind einige Bemerkungen zu neueren Ansätzen und Entwicklungen bei Modellen und Lösungsverfahren für typische Optimierungsprobleme der Energiewirtschaft. Bei einer solchen Zielstellung muß man sich in mindestens zweierlei Hinsicht einschränken. Erstens ist dies bei der Wahl der zu besprechenden Algorithmen notwendig. Nicht berücksichtigt werden in dieser Arbeit Algorithmen, die weder Problemstrukturen noch andere Informationen als Funktionswerte berücksichtigen, wie z.B. genetische Algorithmen, simulated annealing, Heuristiken verschiedenster Art (verwiesen sei z.B. auf den Sammelband [49]). Damit im Zusammenhang steht zweitens die Wahl des zugrunde gelegten Optimierungsmodells. In dieser Arbeit wird ein Optimierungsproblem betrachtet, daß in einer Mischung sowohl hinreichende Allgemeinheit als auch typische Strukturen von energiewirtschaftlichen Modellen repräsentiert.

Wir nehmen an, daß dem Modell ein Zeithorizont von $T$ Intervallen und ein zeitlicher Prozeß $\xi = \{\xi_t\}_{t=1}^T$ von veränderlichen Daten zugrunde liegt. Die Entscheidungsgrößen werden mit $\{x_i\}_{i=1}^I$ bezeichnet, wobei $x_i = x_i^T$, $x_i^t = (x_{i1}, \ldots, x_{it})$. Sowohl die Daten $\xi_t$ als auch die Entscheidungen $x_{it}$ sind dabei i.a. endlichdimensionale Größen. Das Ziel besteht in der Maximierung des Gesamtertrages bzw. in der Minimierung der Gesamtkosten im betrachteten Zeithorizont. Hier wird ein Minimumproblem betrachtet, daß in kompakter Form die folgende Gestalt annimmt.

$$\min\{\sum_{t=1}^{T}\sum_{i=1}^{I} c_{it}(\xi_t, x_i^t) : x_{it} \in X_{it}, \sum_{i=1}^{I} g_{it}(\xi_t, x_{it}) \geq d_t(\xi_t), t = 1, \ldots, T, \tag{1.1}$$

$$a_{it}(\xi_t)x_{it} + a_{i,t-1}(\xi_t)x_{i,t-1} \geq b_{it}(\xi_t), t = 2, \ldots, T, i = 1, \ldots, I\} \tag{1.2}$$

Die Funktionen $c_{it}(\xi_t, \cdot)$ und $g_{it}(\xi_t, \cdot)$ sind dabei im allgemeinen nichtlinear von $x_i^t$ bzw. $x_{it}$ abhängig. Dies erlaubt z.B. die Modellierung nichtlinearer Kostenverläufe von in Betrieb befindlichen Erzeugereinheiten bzw. von Stromverträgen und von zeitabhängigen Anfahrkosten

* Gefördert im Rahmen des Schwerpunktprogramms "Echtzeit-Optimierung großer Systeme" der Deutschen Forschungsgemeinschaft.

thermischer Einheiten. Die Restriktionen (1.1) ermöglichen z.B. die Modellierung von Schaltvariablen und von Kapazitätsschranken, und von Bedingungen an Gruppen von Einheiten, wie z.B. die Lastdeckungsrestriktion. Die Restriktion (1.2) modelliert dynamische Vorgänge, wie z.B. Mindeststillstands- und Mindestlaufzeiten thermischer Einheiten oder die Dynamik der Füllstände von Reservoirs hydraulischer Kraftwerke. Die Einführung des Datenprozesses $\xi$ hat an dieser Stelle die Funktion, die Eingangsgrößen zu identifizieren, die zwar im Zeitpunkt $t = 1$ bekannt sind, aber im zukünftigen Zeithorizont den Charakter eines stochastischen Prozesses besitzen. Dazu gehören z.B. Preise auf Strom- und Primärenergiemärkten, die elektrische Last, Zuflüsse zu hydraulischen Kraftwerken und der durch Windenergie erzeugte Strom.
Der Umgang mit dieser Daten-Stochastik, ihre Abbildung im Modell, unterscheidet nun die *klassischen* determinierten von *stochastischen* Optimierungsmodellen. In ersteren Modellen werden die meist mit einigem Aufwand erhaltenen Vorhersagen für (die Komponenten von) $\xi$ in das Modell (1.1)-(1.2) eingesetzt. Sie haben natürlich den Vorteil, daß sie im Bereich der klassischen mathematischen Optimierung behandelbar sind. Ihr häufig beobachteter Nachteil besteht im ungünstigen Verhalten von Lösungen bei kleinen Störungen der Daten. Demgegenüber verhalten sich Lösungen stochastischer Modelle robust(er) bzgl. solcher Störungen. Bei solchen Modellen geht $\xi$ als stochastischer Prozeß in Zielfunktion und Restriktionen des Modells ein. Nachgedacht werden muß aber über die geeignete Formulierung solcher Modelle, insbesondere der Zielfunktion, über ihre Eigenschaften und über Konsequenzen für Lösungsverfahren.
In Kapitel 2 werden einige neuere Entwicklungen bei Algorithmen für determinierte (lineare, gemischt-ganzzahlige, nichtlineare) Optimierungsmodelle kommentiert. Insbesondere wird auch auf neuere Ergebnisse und Weiterentwicklungen bei klassischen Optimierungsmethoden für Modelle der Energiewirtschaft eingegangen. In Kapitel 3 werden verschiedene Aspekte stochastischer Optimierungsmodelle, wie Modellierung, Risiko-Formulierung, Dekompositionsstrukturen und -methoden, und Lösungsverfahren diskutiert. Eine wichtige Rolle spielen hierbei die Darstellung bzw. Approximation des stochastischen Datenprozesses durch Szenariobäume und die Auswirkung der Einbeziehung von Risikomaßen in die Optimierung.

# 2 Optimierungsalgorithmen für determinierte Modelle

## 2.1 Optimierungstools

In den letzten $10 - 20$ Jahren hat eine enorme Entwicklung bei Lösungsverfahren für Optimierungsprobleme stattgefunden. Dies hat sich in mehreren Ebenen vollzogen: Einerseits wurden neue Lösungsverfahren entwickelt und existierende wesentlich verbessert; andererseits wurde aber auch die Implementierung von klassischen Algorithmen (wie z.B. der Simplexmethode der linearen Optimierung als Basis für verschiedene Algorithmen) entscheidend effektiviert. Verwiesen sei auf die neueren Monographien [36] (für lineare oder nichtlineare Programme ohne Ganzzahligkeitsbedingungen) und [55] (für ganzzahlige Modelle). Die moderne Software für Grundalgorithmen der Optimierung hat inzwischen eine derartigen Perfektionsstand erreicht, daß eigene Implementierungen kaum effizient sein können, und, daß die Aktualisierung der zur Verfügung stehenden kommerziellen Software-Versionen wesentlich ist. Solche Software-Produkte sind z.B.:
- CPLEX (Version 7.1) (LP, GGLP, QP) (http://www.ilog.com);
- OSL (Version 3) (LP, GGLP, QP, SLP) (http://www6.software.ibm.com/sos/osl);
- MINOS (Version 5.5) (große LP und NLP) [34];
- SNOPT (Version 5.3) (große LP und NLP) [15].

Dabei wurden die üblichen Abkürzungen LP (für lineare Programme), GGLP (gemischt-ganzzahlige lineare Programme), QP (Programme mit quadratischer Zielfunktion und linearen Restriktionen), NLP (differenzierbare nichtlineare Programme) und SLP (stochastische lineare Programme) verwendet. Die gegenwärtig favorisierten Algorithmen sind für LPs Varianten des Simplex-Verfahrens und der Barriere-Verfahren; für GGLPs sind es LP-basierte Branch-and-Cut Verfahren (Branch-and-bound mit Schnitten). Für NLPs sind es SQP-Verfahren, die sukzessive die Ziel- bzw. Lagrange-Funktion quadratisch approximieren (mit Quasi-Newton-Typ Aufdatierungen) und die Restriktionen linearisieren (vgl. [36]). MINOS ist zu favorisieren, wenn die Restriktionen (fast) linear sind und nur relativ wenige Variable nichtlinear in der Zielfunktion auftauchen. SNOPT ist etwas zuverlässiger im Fall hoch nichtlinearer Programme. Die genannte Software ist auch im Rahmen der (algebraischen) Modellierungssysteme AMPL (http://www.ampl.com) und GAMS (http://www.gams.com) für Optimierungsprobleme aller Art erhältlich. Solche Modellierungsumgebungen und -sprachen erlauben die Erzeugung großer komplexer Modelle und ihre schnelle Adaption an neue Situationen. Sie stellen deshalb eine große Hilfe für Anwender von mathematischen Optimierungsmodellen und -methoden dar.

## 2.2 Lagrange Relaxation verkoppelnder Restriktionen

Im Unterschied zu Standard-Optimierungstools (vgl. Kap. 2.1) nutzt die *Lagrange Relaxation* strukturelle Eigenschaften des Optimierungsproblems (1.1)–(1.2) aus. Sie gehört zu den wirklich klassischen Techniken zur Lösung energiewirtschaftlicher Optimierungsmodelle und wird seit Anfang der 80er Jahre verwendet (vgl. [48], [56]). Interessant ist, daß diese Methodik über die Jahre ständig weiterentwickelt wurde und daß diese Entwicklung sicher andauern wird. Ursache dafür sind mathematische Fortschritte in den verschiedenen Komponenten dieser Technik. Dies soll im folgenden kurz kommentiert werden. Lange Zeit war klar, was mit der Lagrange Relaxation gemeint war, nämlich die *direkte* Relaxation verkoppelnder Restriktionen. Für das Modell (1.1)–(1.2) bedeutet dies, die Lagrange-Funktion

$$L(x,\lambda) := \sum_{t=1}^{T}\sum_{i=1}^{I}(c_{it}(\xi_t, x_i^t) - \lambda_t g_{it}(\xi_t, x_{it})) + \sum_{t=1}^{T}\lambda_t d_t(\xi_t)$$

zu betrachten. Die zugehörige Dualisierung war gut untersucht und man wußte um die kleine (relative) Dualitätslücke, wenn $I$ groß ist. Motiviert durch [22] tauchte dann Anfang der 90er Jahre eine *Split*-Dekompositionsvariante auf, bei der die Variable $x$ künstlich um eine neue $y$ verdoppelt wird und beide Variablen in jeweils einem Teil der Restriktionen auftauchen. Die Gleichung $x = y$ wird dann relaxiert mit einem Lagrange Multiplikator $\mu$, die Lagrange-Funktion hat die Gestalt

$$L(x,y,\mu) := \sum_{t=1}^{T}\sum_{i=1}^{I}(c_{it}(\xi_t, x_i^t) + \mu_{it}(x_{it} - y_{it}))$$

und wird minimiert bzgl. der Restriktionen

$$\begin{aligned}
&x_{it} \in X_{it}, \sum_{i=1}^{I} g_{it}(\xi_t, y_{it}) \geq d_t(\xi_t), t = 1,\ldots,T, i = 1,\ldots,I,\\
&a_{it}(\xi_t)x_{it} + a_{i,t-1}(\xi_t)x_{i,t-1} \geq b_{it}(\xi_t), t = 2,\ldots,T, i = 1,\ldots,I.
\end{aligned}$$

Dadurch entstehen andere Subprobleme als beim vorhergehenden Ansatz, die Dimension des dualen Problems ist unterschiedlich (in manchen Fällen auch kleiner) und die Ergebnisse von [22] versprachen zumindest im linearen Fall kleinere Dualitätslücken. In der neueren Arbeit [30] wurde jedoch gezeigt, daß die Split-Dekomposition i.a. zu größeren und im linearen Fall zu denselben Dualitätslücken führt. Bei Modellierung von Nichtlinearitäten in der Restriktion (1.1) kann also der direkte Ansatz sogar besser sein. Deshalb sollte man in der Regel bei der klassischen direkten Lagrange Relaxation bleiben.

Wesentliche Fortschritte gab es bei den Lösungsverfahren für die dualen (konkaven, nichtdifferenzierbaren) Maximierungsprobleme. Anfang der 80er Jahre kamen zunächst klassische Subgradienten-Verfahren zum Einsatz (siehe [48]). Mit den Weiterentwicklungen bei Lösungsverfahren für nichtdifferenzierbare Optimierungsprobleme, insbesondere der sogenannten Subgradienten-Bündel Verfahren (vgl. [24]), gab es Anfang/Mitte der 90er Jahre erste Versuche, Bündel-Methoden zur Lösung dualer Blockauswahlprobleme einzusetzen. Die Vorteile dieser Algorithmen bestehen in den sehr guten, robusten Konvergenzeigenschaften und in der deutlichen Verringerung der Anzahl der zu lösenden Subprobleme (vgl. [8], [12],[16,17],[30] und die dort zitierte Literatur). Besonders günstig ist die Situation im Fall, daß die duale Funktion stückweise linear ist. Dies ist der Fall, wenn die Funktionen $g_{it}$ und $c_{it}$ linear bzw. stückweise linear konvex in $x_{it}$ bzw $x_i^t$ sind. In diesem Fall wird von einigen Verfahren die Lösung bereits in endlich vielen Schritten erreicht (vgl. die Diskussion in Kap. 3.4 von [18]). Obwohl sich die Bündel-Methoden inzwischen wohl etabliert haben, wird es bei Umsetzung moderner Entwicklungen (z.B. [12], [31], [32]) zu weiteren Verbesserungen kommen. In engem Zusammenhang mit der Entwicklung der dualen Lösungsalgorithmen standen und stehen Lagrange-Heuristiken zur Bestimmung guter primaler Lösungen. Wurden bisher sehr ausgefeilte reine Heuristiken (wie z.B. [57] im thermischen Fall) bevorzugt, so deuten sich mit den Überlegungen in [12] eine enge Verbindung mit dem dualen Löser und in [51] eine abgesicherte Kopplung mit Branch-and-bound Techniken an. Abschließend sei auch auf die Ausbaubarkeit des Konzeptes der Lagrange Relaxation auf stochastische Optimierungsmodelle und auf die Bemerkungen in Kap. 3.3 verwiesen.

# 3 Stochastische Optimierungsmodelle und -methoden

## 3.1 Szenario-basierte stochastische Programme

Ist der stochastische Charakter des Datenprozesses $\xi$ nicht zu vernachlässigen, wie etwa bei mittleren oder langen Zeithorizonten oder schwer überschaubarer multivariater Daten-Stochastik, so sollten stochastische Optimierungsmodelle verwendet werden. Heute ist bereits ein bemerkenswerter Erfahrungsschatz für die Entwicklung und Lösung solcher Modelle, insbesondere in der Energiewirtschaft, vorhanden (vgl. [4], [54]).

Es sei nun $\xi = \{\xi_t\}_{t=1}^T$ ein stochastischer Prozeß und wir nehmen an, daß er $S$ Realisierungen oder *Szenarios* $\xi^s = \{\xi_t^s\}_{t=1}^T$ mit den Wahrscheinlichkeiten $p_s$, $s = 1,\ldots,S$, besitzt. Nach Voraussetzung soll der Prozeß in $t = 1$ determiniert sein, d.h., es muß $\xi_1^1 = \ldots = \xi_1^S$ gelten. Ansonsten besitzen die Szenarios *Baumstruktur*, d.h., die Anzahl der Szenarios wächst mit wachsendem $t$, das Anwachsen erfolgt durch Verzweigung in bestimmten *Knoten* des Baums und in $t = T$ ist die Anzahl gleich $S$. Diese Baumstruktur induziert eine Strukturierung der Information des Entscheidungsprozesses $x = \{x_i\}_{i=1}^I$, wobei $x_i$ die Szenarien $\{x_{it}^s\}_{t=1}^T$, $s = 1,\ldots,S$, besitzt. Der Entscheidungsprozeß $x$ ist nämlich *nichtantizipativ* in dem Sinn, daß die Entscheidung $x_t$ zum Zeitpunkt $t$ nicht von der Zukunft abhängt, also nicht vorgreifend (oder antizipativ) ist. Diese Eigenschaft kann wie folgt mathematisch formalisiert werden. Es bezeichne $\mathcal{E}_t$ eine Zerlegung der Menge $\{1,\ldots,S\}$ in durchschnittsfremde Teilmengen, so daß die Anzahl der Elemente von $\mathcal{E}_t$ gerade der Anzahl von Szenarien des Datenprozesses $\xi$ zum Zeitpunkt $t$ und die Verfeinerung der Zerlegung beim Übergang von $t$ zu $t+1$, d.h., von $\mathcal{E}_t$ zu $\mathcal{E}_{t+1}$, gerade der Verzweigungsstruktur von $\xi$ entspricht. Dabei gilt $\mathcal{E}_1 = \{1,\ldots,S\}$ und $\mathcal{E}_T = \{\{1\},\ldots,\{S\}\}$. Dann kann die Nichtantizipativität von $x$ gerade in Form linearer Gleichungsbedingungen an die Szenarien von $x$ in folgender Art und Weise formuliert werden (vgl. die Herleitung in Kap. 2.4 von [46]).

$$x_{it}^\sigma = \sum_{\substack{E\in\mathcal{E}_t\\ \sigma\in E}} (\sum_{\substack{s=1\\ s\in E}}^S p_s)^{-1} \sum_{\substack{s=1\\ s\in E}}^S p_s x_{it}^s\,,\ \sigma = 1,\ldots,S,\ t = 1,\ldots,T-1, i = 1,\ldots,I. \tag{3.1}$$

Für $t = 1$ bedeutet (3.1) gerade $x_{i1}^{\sigma} = \sum_{s=1}^{S} p_s x_{i1}^{s}$, $\sigma = 1, \ldots, S$, d.h., $x_{i1}^{1} = \ldots = x_{i1}^{S}$. Betrachtet man nun wieder das Optimierungsmodell (1.1)–(1.2), so ist es natürlich, zu verlangen, daß die dortigen Restriktionen für alle Szenarien des Entscheidungsprozesses erfüllt sein sollen. Überdies müssen die (jetzt) stochastischen Kosten $\sum_{t=1}^{T} \sum_{i=1}^{I} c_{it}(\xi_t, x_i^t)$ von (1.1) durch eine neue Zielfunktion ersetzt werden. Hier soll dies zunächst durch die erwarteten Kosten $\sum_{s=1}^{S} \sum_{t=1}^{T} \sum_{i=1}^{I} p_s c_{it}(\xi_t^s, (x_i^s)^t)$ (mit $(x_i^s)^t := (x_{i1}^s, \ldots, x_{it}^s)$) als Zielfunktion erfolgen. Wir kommen jedoch in Kap. 3.4 auf diese Frage zurück. Insgesamt entsteht damit ein stochastisches Optimierungsproblem in *Szenario-Formulierung* von folgender Gestalt:

$$\min\{\sum_{s=1}^{S} \sum_{t=1}^{T} \sum_{i=1}^{I} p_s c_{it}(\xi_t^s, (x_i^s)^t) : x \text{ erfüllt die Bedingungen (3.1)}, \tag{3.2}$$

$$x_{it}^s \in X_{it}, \sum_{i=1}^{I} g_{it}(\xi_t^s, x_{it}^s) \geq d_t(\xi_t^s), t = 1, \ldots, T, \tag{3.3}$$

$$a_{it}(\xi_t^s) x_{it}^s + a_{i,t-1}(\xi_t^s) x_{i,t-1}^s \geq b_{it}(\xi_t^s), t = 2, \ldots, T, i = 1, \ldots, I, s = 1, \ldots, S\} \tag{3.4}$$

Wir kehren nun zum Bild von der Baumstruktur zurück. Passend dazu nennt man den Knoten $t = 1$ *Wurzelknoten* und die Szenarioenden in $t = T$ *Blätter*. Abbildung 1 zeigt einen sogenannten *binären* Szenariobaum, wobei die $t_k$ die Zeitpunkte bezeichnen, an denen eine Verzweigung in

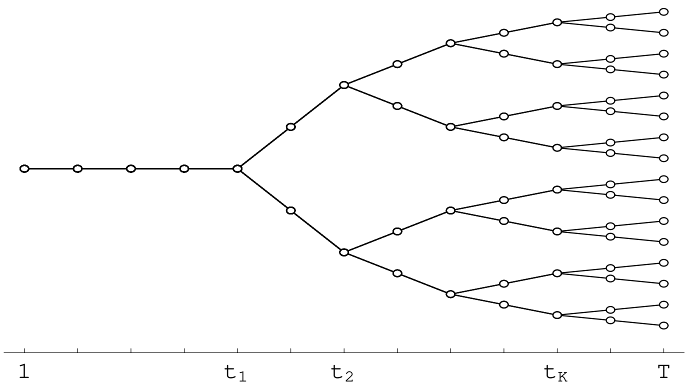



Abbildung 1: Beispiel eines binären Szenariobaums

zwei Äste erfolgt. Wir numerieren nun mit einer Menge $\mathcal{N}$ die Knoten des Baums, wobei $n = 1$ für den Wurzelknoten steht. Jeder andere Knoten $n$ besitzt einen eindeutig bestimmten *Vorgänger* $n_-$ und eine Übergangswahrscheinlichkeit $\pi_{n/n_-} > 0$. Die Wahrscheinlichkeit $\pi_n$ eines Knotens $n$ ist rekursiv durch $\pi_1 := 1$, $\pi_n := \pi_{n/n_-} \pi_{n_-}$, $n > 1$, definiert. Mit $\mathcal{N}_+(n)$ wird die Menge der *Nachfolger* des Knotens $n$ und mit $\text{path}(n)$ die Menge der Knoten des Pfades von der Wurzel bis zum Knoten $n$ bezeichnet. $t(n)$ sei die Anzahl der Elemente von $\text{path}(n)$. Für $t = 1, \ldots, T$ bezeichnet $\mathcal{N}_t$ die Menge $\{n \in \mathcal{N} : t(n) = t\}$; es gilt $\sum_{n \in \mathcal{N}_t} \pi_n = 1$. Ein Szenario entspricht dann einem Pfad $\text{path}(n)$ von der Wurzel zu einem Blatt $n \in \mathcal{N}_T$, und es gilt $\{\pi_n\}_{n \in \mathcal{N}_T} = \{p_s\}_{s=1}^{S}$. Schließlich bezeichnen wir mit $\{\xi^n\}_{n \in \mathcal{N}}$ die Szenariobaum-Darstellung von $\xi$ und mit $\{x_i^n\}_{n \in \mathcal{N}}$ die des Entscheidungsprozesses $x_i$ für $i = 1, \ldots, I$. Mit diesen Vorbereitungen läßt sich analog zu (3.2)–(3.4) die *Szenariobaum-Formulierung* dieses stochastischen Optimierungsmodells wie folgt formulieren:

$$\min\{\sum_{n \in \mathcal{N}} \sum_{i=1}^{I} \pi_n c_{it(n)}(\xi^n, x_i^{\text{path}(n)}) : x_i^n \in X_{it(n)}, \sum_{i=1}^{I} g_{it(n)}(\xi^n, x_i^n) \geq d_{t(n)}(\xi^n), n \in \mathcal{N}, \tag{3.5}$$

$$a_{it(n)}(\xi^n) x_i^n + a_{it(n_-)}(\xi^n) x_i^{n_-} \geq b_{it(n)}(\xi^n), n \in \mathcal{N} \setminus \{1\}, i = 1, \ldots, I\} \tag{3.6}$$

Hierbei bezeichnen wir analog zur früheren Notation $x_i^t$ mit $x_i^{\text{path}(n)}$ für $n \in \mathcal{N}$ den Pfad von $x_i^1$ zu $x_i^n$. Die Nichtantizipativitätsbedingung tritt in der Szenariobaum-Formulierung des Modells nicht mehr explizit auf. Sie ist in die Baumkonstruktion integriert worden. Die Dimensionen beider Modelle sind stark unterschiedlich. So ist zum Beispiel die Anzahl der Entscheidungsvariablen in der Szenario-Formulierung (3.2)–(3.4) gleich $ITS$ und in der Baum-Formulierung (3.5)–(3.6) gleich $IN$, wobei $N$ die Anzahl der Knoten ist. Wegen $N << TS$ ist die Baum-Formulierung von der Dimension her deutlich günstiger.

## 3.2 Konstruktion von Szenariobäumen

In jeder praxisnahen Anwendung der stochastischen Optimierung stellt die (approximative) Darstellung des stochastischen Datenprozesses eine zentrale Fragestellung dar. Dies trifft natürlich auch auf die Energiewirtschaft zu, wo solche Darstellungen zwar bereits für einzelne Komponenten (wie die Last in [18,20,41] oder Spotmarktpreise und Zuflüsse in [13,10]) entwickelt wurden, aber insbesondere die Modellierung multivariater Datenprozesse noch weitgehend am Anfang steht. Universelle Methodiken zur Modellierung von Datenprozessen in Form von Szenariobäumen sind bisher nicht bekannt. Der gegenwärtige Stand der Literatur in diesem sich schnell entwickelnden Teilgebiet wird in der Übersichtsarbeit [10] sehr gut dargestellt. Alle aus der Literatur bekannten methodischen Ansätze bestehen aus folgenden Bestandteilen:
(a) Erzeugung von (ausreichend vielen) Szenarien des Datenprozesses $\xi$;
(b) Konstruktion von Szenariobäumen aus Datenszenarien oder Verteilungsinformationen;
(c) (optionale) Nachbehandlung des Szenariobaums.
Alle diese Bestandteile sollen nachfolgend kommentiert werden.
(a) Ausgangspunkt sind historische statistische Daten des zu modellierenden Prozesses. Der wesentlichste Zugang besteht darin, ein stochastisches Modell zu entwickeln und dessen Parameter an die vorhandenen Daten anzupassen. Dies können z.B. (multivariate) Zeitreihenmodelle (z.B. für die elektrische Last wie in [19,20]) oder Modelle vom Diffusionstyp (für die Modellierung von Marktpreisen) sein. Sind solche Modelle ausreichend gerechtfertigt und kalibriert, lassen sich im Prinzip beliebig viele Realisierungen bzw. Szenarien generieren. Diese Modelle beruhen stets auf einer Annahme über die Wahrscheinlichkeitsverteilung des Datenprozesses. Es existieren jedoch auch nichtparametrische (oder verteilungsfreie) Ansätze. Der einfachste Ansatz dieser Art besteht darin, historische Daten, die unter vergleichbaren Umständen erhalten wurden, direkt als Szenarien zu verwenden und ihnen gleiche Wahrscheinlichkeiten zuzuordnen. Ergänzend könnten Expertenszenarien hinzugefügt werden, bei denen dann aber die zugehörigen Wahrscheinlichkeiten angepaßt werden müssen.
(b) Zunächst sollen drei allgemeinere Methoden zur Konstruktion von Szenariobäumen erwähnt werden: (i) die baryzentrischen Szenariobaumkonstruktionen [14], (ii) die Konstruktion über EVPI-basierte importance sampling Methoden (vgl. Kap. 3.1 in [10]) und (iii) Konstruktion der Bäume so, daß sie vorgegebene Momenten-Bedingungen erfüllen (vgl. Kap. 3.2 in [10]). Eine weitere allgemeine Methodik ist die Konstruktion von Szenariobäumen mittels Cluster-Analyse. Man geht dabei von $S$ Datenszenarien mit identischem Startpunkt $\xi_1$ aus. Diese Szenarien bilden einen Fächer, d.h., einen Baum einfachster Struktur. Auf der Grundlage eines Ähnlichkeitsmaßes für Szenarien werden dann die Szenarien in jedem $t > 1$ auf Ähnlichkeit, d.h., Zugehörigkeit zu einem Cluster, untersucht. Zugehörigkeit zu einem Cluster bedeutet dann Zusammenfassung der entsprechenden Szenarien zu einem Repräsentanten (z.B. den Mittelwert) mit der Wahrscheinlichkeit, die gleich der Summe der Wahrscheinlichkeiten der einzelnen zugehörigen Szenarien ist. Die Bildung neuer Cluster bedeutet Verzweigung von den Vorgängern. Bei dieser Methode muß die Struktur des Baumes nicht vorgegeben werden. Allerdings ist die Wahl des Ähnlichkeitsmaßes entscheidend für das Ergebnis. In diese Klasse von Cluster-basierten Methoden fällt auch

die Methodik in [21]. Dabei wird auf der Grundlage einer Wahrscheinlichkeitsmetrik, d.h., eines Abstandes von multivariaten Wahrscheinlichkeitsverteilungen, beginnend bei $t = T$ entschieden, wo die Szenarien zusammengefaßt werden. Bei einfacheren Vorgehensweisen gibt man die Verzweigungspunkte (z.B. Zeitpunkte für neue Beobachtungen bzw. Messungen) und evtl. auch den Verzweigungsgrad (z.B. binär oder dreifaltig) fest vor und bestimmt "lediglich" die Werte der Szenarien in allen Knoten. Letzteres erfolgt dann aus Verteilungsinformationen wie z.B. in (iii) und bei den Konstruktionen für die Last-Szenariobäume in [18] und [11].
(c) Der in (b) erhaltene Szenariobaum kann zu groß für die nachfolgende Optimierung sein. "Zu groß" bedeutet dabei, daß er zu viele Knoten besitzt und die Variablenanzahl des Optimierungsproblems folglich zu riesig ist, um es in vernünftiger Zeit zu lösen. Orientiert man sich z.B. in (b) zu sehr daran, den Datenprozeß gut zu approximieren, so entstehen zu viele Szenarien und damit zu dichte Bäume. Dies macht eine *Reduktion* des Szenariobaums, anschaulich ein Ausdünnen des Baumes, erforderlich. Neben anderen Zugängen (z.B. auf Sensitivitätsanalyse beruhenden, vgl. die Literatur in [10], [11]) sei hier der Zugang aus [11] erwähnt, der darin besteht, den reduzierten Szenariobaum (näherungsweise) optimal bzgl. einer gewissen Wahrscheinlichkeitsmetrik zu wählen. Diese Wahrscheinlichkeitsmetrik wird derart an das zu lösende stochastische Optimierungsproblem angepaßt, daß sich sein Optimalwert und seine Lösungen stabil bzgl. dieser Metrik verhalten. Es zeigt sich in [11], daß die optimale Reduktion näherungsweise durch einfach zu implementierende (heuristische) Algorithmen realisierbar ist. In [23] wurden die Reduktionsalgorithmen weiterentwickelt und umfangreich getestet.

### 3.3 Dekompositionszugänge

Die in Abschnitt 3.1 entwickelten Modellformulierungen stellen bereits für kleine und mittlere Anzahlen von Szenarien bzw. Knoten aus mathematischer Sicht hochdimensionale Optimierungsprobleme mit spezieller Struktur dar. Selbst im Fall, daß die Probleme keine ganzzahligen Entscheidungen enthalten, scheitern die in Kap. 2.1 besprochenen Standardlöser oftmals an den gewaltigen Dimensionen. Jedoch sind für solche Modelle eine ganze Reihe von strukturausnutzenden *Dekompositionstechniken* bekannt (vgl. die Übersichtsarbeiten [3] und [47] sowie die dort zitierte Literatur). Für lineare Modelle sind insbesondere als primaler Zugang die nested Benders Dekomposition (vgl. [33]) und als primal-duale Zugänge augmented Lagrange-Methoden ([47]) und der progressive hedging Algorithmus ([45]) zu nennen.
Im Fall gemischt-ganzzahliger linearer Modelle stellen sich die meisten Fragen neu (vgl. die Übersichtsarbeit [28]). Für solche stochastischen Optimierungsprobleme sind, wenn nicht sehr spezielle Strukturen vorliegen, *Dekompositionstechniken* die einzige praktikable Alternative. Primale Techniken leiden unter dem Verlust an Konvexität und sind gegenwärtig nur für kleinere Modelle praktikabel (vgl. die Literaturhinweise in [46]). Mit *dualen* Techniken gelingt wie im determinierten Fall die Berechnung unterer Schranken für den Optimalwert und die dualen Lösungen können wieder als Ausgangspunkt für globale Optimierungstechniken dienen.
Die beiden für gemischt-ganzzahlige stochastische Optimierungsmodelle wesentlichen Dekompositionsansätze beruhen auf dem klassischen Prinzip der Lagrange Relaxation (vgl. [48] und Kap. 2.2). Es handelt sich dabei um die (i) *Szenario-Dekomposition*, d.h., die Relaxation der Nichtantizipativitätsbedingungen, und (wie in Kap. 2.2) die (ii) Relaxation von Restriktionen, die Entscheidungskomponenten *verkoppeln*. Folgt man bei der Dualisierung den klassischen Ansätzen von Rockafellar-Wets [44], so werden die zu relaxierenden Restriktionen, die ja in beiden Fällen durch stochastische Größen gegeben sind, mit Lagrange-Multiplikatoren multipliziert und die Erwartungswerte dieser Terme in die Lagrange-Funktion aufgenommen. Beim Ansatz für die

Szenario-Dekomposition führt diese Idee ausgehend von (3.2) zu folgender Lagrange-Funktion

$$L_1(x,\lambda) \;=\; \sum_{s=1}^{S}\sum_{t=1}^{T}\sum_{i=1}^{I} p_s(c_{it}(\xi_t^s,(x_i^s)^t) - \lambda_{it}^s x_{it}^s).$$

Die zugehörige duale Funktion hat dann die Gestalt

$$\begin{aligned} D_1(\lambda) &:= \inf\{L_1(x,\lambda)\,:\,x \text{ erfüllt die Bedingungen (3.3) und (3.4)}\} \\ &= \sum_{s=1}^{S} p_s D_{1s}(\lambda)\,, \qquad \text{wobei} \\ D_{1s}(\lambda) &:= \inf\{\sum_{t=1}^{T}\sum_{i=1}^{I}(c_{it}(\xi_t^s,(x_i^s)^t) - \lambda_{it}^s x_{it}^s)\,:\,x_{it}^s \in X_{it}, \sum_{i=1}^{I} g_{it}(\xi_t^s,x_{it}^s) \geq d_t(\xi_t^s), \\ &\qquad t=1,\ldots,T\,,\, a_{it}(\xi_t^s)x_{it}^s + a_{i,t-1}(\xi_t^s)x_{i,t-1}^s \geq b_{it}(\xi_t^s), t=2,\ldots,T, i=1,\ldots,I\}. \end{aligned} \tag{3.7}$$

Die Berechnung von Funktionswerten der dualen Funktion erfolgt also über die Lösung von $S$ determinierten Subproblemen (3.7) der gleichen Struktur wie das zugrunde liegende Modell (1.1)–(1.2) mit dem Szenario $\xi^s$, $s=1,\ldots,S$, anstelle des stochastischen Datenprozesses. Die Dualisierung führt also zu einer Szenario-Dekomposition. Das duale Problem hat die Form

$$\max\{D_1(\lambda)\,:\,\lambda \in \Lambda_1\}, \tag{3.8}$$

wobei $\Lambda_1$ ein endlichdimensionaler Raum ist, der komplementär zu dem durch die Gleichungen (3.1) definiertem Raum ist. Mit den Bezeichnungen aus Kap. 3.1 wird $\Lambda_1$ wie folgt beschrieben:

$$\Lambda_1 = \{\lambda\,:\, \sum_{\substack{s=1\\ s\in E}}^{S} p_s\lambda_{it}^s = 0\,,\, E \in \mathcal{E}_t\,,\, t=1,\ldots,T\,,\, i=1,\ldots,I\}$$

Aus der Struktur von $\Lambda_1$ geht hervor, daß außer $\lambda = 0$ kein Element aus $\Lambda_1$ nichtantizipativ ist, also kein anderes eine Szenariobaum-Struktur wie $\xi$ oder $x$ besitzt. Wegen $\lambda_{iT}^s = 0$ für $i=1,\ldots,I\,,\, s=1,\ldots,S$, ist deshalb die Dimension von $\Lambda_1$ gleich $IS(T-1)$ und kann i.a. auch nicht verringert werden.
Bei der Lagrange Relaxation verkoppelnder Restriktionen ist es günstiger von der Szenariobaum-Formulierung (3.5)–(3.6) des stochastischen Optimierungsproblems auszugehen. Es seien also $\{\xi^n\}_{n\in\mathcal{N}}$, $\{x_i^n\}_{n\in\mathcal{N}}$ und $\{\pi_n\}_{n\in\mathcal{N}}$ die Szenariobäume des Datenprozesses $\xi$ bzw. des Enscheidungsprozesses $x_i$ und die Menge der Knotenwahrscheinlichkeiten. Analog bezeichne $\{\lambda^n\}_{n\in\mathcal{N}}$ den entsprechenden Szenariobaum des stochastischen Lagrange Multiplikators $\lambda$ (der die Rolle stochastischer Preise spielt) zur Dualisierung der Restriktionen $\sum_{i=1}^{I} g_{it(n)}(\xi^n,x_i^n) \geq d_{t(n)}(\xi^n)$ für $n \in \mathcal{N}$. Dann haben die Lagrange-Funktion $L_2$ und die duale Funktion $D_2$ die Gestalt

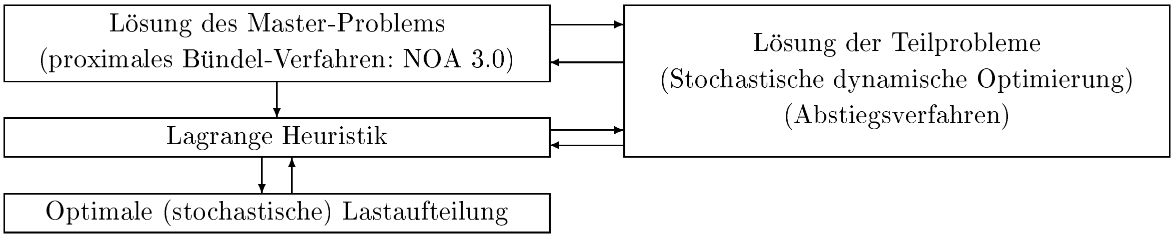
$$L_2(x,\lambda) \;=\; \sum_{n\in\mathcal{N}}\sum_{i=1}^{I}\pi_n(c_{it(n)}(\xi^n,x_i^{\mathrm{path}(n)}) - \lambda^n g_{it(n)}(\xi^n,x_i^n)) + \sum_{n\in\mathcal{N}}\pi_n\lambda^n d_{t(n)}(\xi^n)\,.$$

$$\begin{aligned} D_2(\lambda) &:= \inf\{L_2(x,\lambda)\,:\,x_i^n \in X_{it(n)}, n\in\mathcal{N}, a_{it(n)}(\xi^n)x_i^n + a_{it(n_-)}(\xi^n)x_i^{n-} \geq b_{it(n)}(\xi^n), \\ &\qquad\qquad n \in \mathcal{N}\setminus\{1\}, i=1,\ldots,I\} \\ &= \sum_{i=1}^{I} D_{2i}(\lambda) + \sum_{n\in\mathcal{N}}\pi_n\lambda^n d_{t(n)}(\xi^n)\,, \qquad \text{wobei} \\ D_{2i}(\lambda) &:= \inf\{\sum_{n\in\mathcal{N}}\pi_n(c_{it(n)}(\xi^n,x_i^{\mathrm{path}(n)}) - \lambda^n g_{it(n)}(\xi^n,x_i^n))\,:\,x_i^n \in X_{it(n)}, n\in\mathcal{N}, \\ &\qquad a_{it(n)}(\xi^n)x_i^n + a_{it(n_-)}(\xi^n)x_i^{n-} \geq b_{it(n)}(\xi^n), n\in\mathcal{N}\setminus\{1\}\}, \end{aligned} \tag{3.9}$$

wobei die Subprobleme (3.9) hier stochastische Programme kleinerer Dimension darstellen. Analog zu (3.8) hat das duale Problem wieder die Form

$$\max\{D_2(\lambda)\ :\ \lambda \in \Lambda_2\}, \tag{3.10}$$

wobei die Menge $\Lambda_2$ in diesem Fall einfach durch Nichtnegativitätsbedingungen gegeben ist, d.h., $\Lambda_2 := \{\lambda = \{\lambda^n\}_{n\in\mathcal{N}} : \lambda^n \geq 0\,,\, n \in \mathcal{N}\}$.
Die beiden dualen Probleme (3.8) und (3.10) fungieren als Master-Probleme. Ihrer iterativen Lösung entspricht nun eine sukzessive Dekomposition des Originalproblems (in Szenario- oder Szenariobaum-Formulierung). Beide duale Maximierungsprobleme sind konkav und nichtdifferenzierbar (da die dualen Funktionen selbst Ergebnis von Minimierungsproblemen sind). Deshalb erfordert ihre Lösung den Einsatz von Verfahren, die Subgradienten (anstelle) von Gradienten verwenden und mittels eines *Bündels* solcher Subgradienten stückweise lineare Funktionen erzeugen, die die duale Funktion von oben (immer besser) approximieren. Moderne Subgradienten-Bündel Verfahren sind in [26,32,31] zu finden.
Nach der näherungsweisen Lösung der dualen Probleme ist mit dem dualen Optimalwert $\max_\lambda D_i(\lambda)$ $(i = 1, 2)$ jedoch nur eine untere Schranke für den Optimalwert des Originalproblems gefunden und es existiert eine *Dualitätslücke*. Offene Fragen hierbei sind vergleichende Abschätzungen für die Dualitätslücke bei beiden Strategien, die die Abschätzung in [52] im Fall der Relaxierung der Nichtantizipativitätsbedingungen verwenden bzw. weiterentwickeln.
Beide Dekompositionsansätze erfordern abschließende globale Optimierungstechniken bzw. (Lagrange) Heuristiken, um aus den nach Lösung des Master-Problems erhaltenen Informationen (näherungsweise) Lösungen des Originalproblems zu berechnen. Jedoch gibt es in der neueren Literatur zu stochastischen Optimierungsmodellen der Energiewirtschaft auch Versuche, Methodiken, die eigentlich im gemischt-ganzzahligen Fall theoretisch nicht gerechtfertigt sind, direkt anzuwenden (z.B. progressive hedging in [50,52] und augmented Lagrange Methoden in [7]).
In der Arbeit [6] ist eine Variante der Szenario-Dekomposition einschließlich der Anbindung eines Branch-and-bound Verfahrens für zweistufige Modelle beschrieben und implementiert worden. Diese Vorgehensweise erscheint insbesondere dann als günstig, wenn Ähnlichkeiten der Szenario-Subprobleme (3.7) effektiv ausgenutzt werden können. Die Dekomposition verkoppelnder Restriktionen ist dann vorteilhaft, wenn die stochastischen Subprobleme (3.9) schnell gelöst werden können (insbesondere durch Ausnutzung der Szenariobaum-Struktur). Sie ist inzwischen in zahlreichen Arbeiten auf stochastische energiewirtschaftliche Modelle angewendet worden (z.B. [2], [8], [9,18,19,37],[29], [53]). Die Methodiken der einzelnen Arbeiten unterscheiden sich dabei vor allem durch die verwendeten Algorithmen zur Lösung des Master-Problems, der Teilprobleme und für die Lagrange Heuristik.
Im folgenden gehen wir auf den Zugang aus [18,19,37] näher ein und erläutern dessen Möglichkeiten am Beispiel der Lösung eines stochastischen Modells zur Optimierung des wöchentlichen Kraftwerkseinsatzes bei stochastischer Last im hydro-thermischen System der VEAG. Hierbei wird das proximale Bündel Verfahren von Kiwiel (vgl. [26,27]) zur Lösung des Master-Problems (3.10) eingesetzt (vgl. Kap. 3.4 in [18]). Dieses Verfahren besitzt gute Konvergenzeigenschaften, insbesondere, wenn die duale Funktion stückweise linear ist (vgl. Kap. 2.2). Überdies ist es für große Probleme geeignet. Die Subprobleme entsprechen stochastischen Modellen für die einzelnen Erzeugereinheiten. Im Fall der thermischen Modelle werden diese mit stochastischer dynamischer Optimierung mit in den Zustandsbaum integriertem Szenariobaum gelöst. Für die hydraulischen Modelle wurde ein spezielles Abstiegsverfahren (vom Netzwerkfluß-Typ) verwendet, das die Szenariobaumstruktur ausnutzt ([18,37]). Schließlich wird mittels einer Lagrange Heuristik zunächst eine zulässige Lösung des Originalproblems bestimmt und danach mit einer optimalen stochastischen Lastaufteilung verbessert. Die nachfolgende Abbildung veranschaulicht die Struktur des gesamten *Optimierungsverfahrens*.




Dieses Verfahren wurde am hydro-thermischen Kraftwerkssystem der VEAG mit $T = 168$ (eine Woche im Stundentakt), $I = 32$ (25 thermisch und 7 hydraulisch), stückweise linearen Brennstoffkosten, stückweise konstanten Anfahrkosten und stochastischer Last getestet. Dabei wurden zu Testzwecken Last-Szenariobäume mit unterschiedlichen Anzahlen von Szenarien und Knoten zufällig erzeugt. Die Testrechnungen wurden auf einer HP 9000 Workstation (770/J280) (180 MHz, 768 MByte, HP-UX 10.20) durchgeführt. Die Rechenzeiten (in Sekunden) betrugen für 1 Szenario (d.h. $S = 1$) etwa 15s, für $S = 20$ etwa 90s und für $S = 100$ (9000 Knoten) etwa 1100s (vgl. [18,19,37] für eine detailliertere Darstellung der Ergebnisse).

## 3.4 Risiko-Modellierung

In neueren Arbeiten zur Stochastik (insbesondere zur Finanzmathematik und zum Operations Research) werden geeignete Eigenschaften und Ansätze von Risikomaßen diskutiert. Die Diskussion entwickelt sich dabei entlang verschiedener Linien: (i) eine durch [1] initiierte axiomatische Richtung, die sich an wünschenswerten Eigenschaften von Risikomaßen orientiert und im Begriff eines *kohärenten Risikomaßes* mündet (vgl. auch [25]), (ii) eine durch die *stochastische Dominanz*, eine Ordnungsrelation zwischen (reellen) Zufallsgrößen, geprägte Richtung (vgl. [38,39]), und (iii) eine an für die Optimierung wesentlichen Eigenschaften ausgerichtete Linie (vgl. [42,43]). Interessanterweise entstammen allen drei Entwicklungslinien ähnliche Vorschläge, die im Kontext der Optimierung energiewirtschaftlicher Modelle interessant erscheinen.

Gemäß [1] ist ein Risikomaß $\rho$ ein Funktional auf einem linearen Raum von (reellen) Zufallsgrößen, die im folgenden mit $z$ bzw. $\bar{z}$ bezeichnet werden sollen. Ein Risikomaß $\rho$ heißt *kohärent*, falls für $\lambda \geq 0$ und $c \in I\!R$ die Bedingungen

$$\rho(z + \bar{z}) \leq \rho(z) + \rho(\bar{z}) \quad \text{und} \quad \rho(\lambda z) = \lambda \rho(z)\,, \tag{3.11}$$

$$\rho(z) = -c \quad \text{falls} \quad z \equiv c\,, \qquad \rho(z) \geq \rho(\bar{z}) \quad \text{falls} \quad z \leq \bar{z}\,, \tag{3.12}$$

erfüllt sind. Wegen Bedingung (3.11) sind kohärente Risikomaße insbesondere *konvex*, was sie für die Optimierung interessant macht. Einige populäre Ansätze für Risikomaße wie z.B. die Varianz bzw. die Standardabweichung oder der Value-at-Risk (3.16) sind nicht kohärent. Kohärent sind aber folgende Risikomaße

$$\rho(z) \;:=\; -I\!E[z] + \alpha I\!E[\max\{0, I\!E[z] - z\}], \tag{3.13}$$

$$\rho(z) \;:=\; -I\!E[z] + \alpha \left( I\!E[(\max\{0, I\!E[z] - z\})^2] \right)^{\frac{1}{2}}, \tag{3.14}$$

falls $\alpha \in (0, 1]$ (vgl. Kap. 2 in [39]). Diese Risikomaße enthalten die absolute Semiabweichung bzw. die Standard-Semiabweichung. Hierbei und später bezeichnen $I\!E[\cdot]$ den Erwartungswert und $I\!P(\cdot)$ die Wahrscheinlichkeit, und $\rho$ ist wohl-definiert auf Räumen von Zufallsgrößen mit endlichen ersten Momenten im Fall (3.13) bzw. mit endlichen zweiten Momenten im Fall (3.14). Die günstigen Eigenschaften der in (3.13) bzw. (3.14) verwendeten Semiabweichungen wurden im Zusammenhang mit der Untersuchung der Relation der stochastischen Dominanz zwischen Zufallsgrößen in [38,39] entdeckt. Die Überlegungen in [39] führten auch zur Betrachtung des folgenden Risikomaßes, dem $\alpha$-*Conditional* (oder *Tail*) *Value-at-Risk*, das in gewissem Sinn dual zu (3.13) ist (vgl Kap. 4 in [39]).

$$\rho_\alpha(z) := CVaR_\alpha(z) := \min\{r + \frac{1}{1-\alpha}\mathbb{E}[\max\{0, z-r\}] \,:\, r \in \mathbb{R}\} \quad (\alpha \in (0,1)) \tag{3.15}$$

In [43] wird gezeigt, daß für jede Zufallsgröße $z$ die Menge der Minimalelemente in (3.15) ein abgeschlossenes, beschränktes Intervall ist, von dem der $\alpha$-*Value-at-Risk* von $z$

$$VaR_\alpha(z) := \min\{r \in \mathbb{R} : \mathbb{P}(z \leq r) \geq \alpha\} \tag{3.16}$$

der linke Randpunkt ist. Überdies gilt, daß das Risikomaß $CVaR_\alpha$ die Bedingung (3.11) und eine modifizierte Bedingung (3.12) erfüllt (vgl. [40]), und, daß $CVaR_\alpha$ stetig von $\alpha$ abhängt, während dies für $VaR_\alpha$ nicht der Fall ist. Insgesamt wird damit die spezielle Stellung von $CVaR_\alpha$ unter den Risikomaßen ersichtlich. Die Bezeichnung $\alpha$-Conditional Value-at-Risk wird dadurch motiviert, daß im Fall einer stetigen Verteilung von $z$ die Gleichheit $CVaR_\alpha(z) = \mathbb{E}[z|z \geq VaR_\alpha(z)]$ mit der bedingten Erwartung von $z$ bzgl. des Ereignisses $\{z \geq VaR_\alpha(z)\}$ gilt.
Noch deutlicher wird die spezielle Rolle von $CVaR$, wenn man Optimierungsmodelle betrachtet, die $CVaR$ in der Zielfunktion (d.h. zur Minimierung des Risikos von Entscheidungen) oder in Restriktionen (d.h. zur unbedingten Einhaltung bestimmter Risikoschranken) enthalten. Dies soll abschließend an einem stochastischen Optimierungsproblem der Gestalt wie in Kap. 3.1 bei Einbeziehung der Minimierung des Risikos in Form von $CVaR_\alpha$ demonstriert werden. Mit den Bezeichnungen aus Kap. 3.1 seien $f(x_1, \xi)$ die (stochastischen) Kosten der Entscheidung $x_1$ in $t = 1$ bzw. im Wurzelknoten $n = 1$. Diese erhält man rekursiv durch stochastische dynamische Optimierung aus den Modellen (3.2)–(3.4) bzw. (3.5)–(3.6) (vgl. auch Kap. 2.2 in [46] für eine kompakte Darstellung). In der Menge $X_1$ seien alle Restriktionen von $x_1$ zusammengefaßt. Für $\alpha \in (0,1)$ und $\lambda \in [0,1]$ betrachten wir nun das Optimierungsproblem

$$\min\{\lambda\mathbb{E}[f(x_1,\xi)] + (1-\lambda)\rho_\alpha(f(x_1,\xi)) : x_1 \in X_1\}. \tag{3.17}$$

Die Zielfunktion in (3.17) entspricht also einem gewichteten Mittel der erwarteten Kosten, die in den Modellformulierungen in Kap. 3.1 ausschließlich verwendet wurden, und des Risikoterms $\rho_\alpha(f(x_1,\xi))$, der das Risiko der Entscheidung $x_1$ bewertet. Entsprechend der Definition (3.15) von $CVaR_\alpha$ ist es sinnvoll, die Funktion $F_\alpha$ und das folgende modifizierte Problem zu (3.17) zu betrachten:

$$F_\alpha(x_1, r) := r + \frac{1}{1-\alpha}\mathbb{E}[\max\{0, f(x_1,\xi) - r\}] \quad (x_1 \in X_1, r \in \mathbb{R}),$$

$$\min\{\lambda\mathbb{E}[f(x_1,\xi)] + (1-\lambda)F_\alpha(x_1,r) : x_1 \in X_1, r \in \mathbb{R}\}. \tag{3.18}$$

Aus den Ergebnissen von Kap. 4 in [43] ergibt sich, daß für jede Lösung $(x_1^*, r^*)$ von (3.18) $x_1^*$ eine Lösung von (3.17) ist und daß die Abschätzung $VaR_\alpha(f(x_1^*,\xi)) \leq \rho_\alpha(f(x_1^*,\xi)) = F_\alpha(x_1^*, r^*)$ gilt. Man erhält also sogar eine Abschätzung für den $\alpha$-Value-at-Risk der Kosten in der optimalen Lösung. Überdies nimmt das modifizierte Problem (3.18) im Fall diskreter Wahrscheinlichkeitsverteilungen eine besonders einfache Gestalt an. Sind nämlich $\xi^s$, $s = 1, \ldots, S$, die endlich vielen Szenarien von $\xi$ mit den Wahrscheinlichkeiten $p_s$, $s = 1, \ldots, S$, so führt man neue Variable $y_s$, $s = 1, \ldots, S$, ein und betrachtet das im Vergleich zu (3.18) erweiterte Problem

$$\min\{\lambda\sum_{s=1}^{S} p_s f(x_1,\xi^s) + (1-\lambda)(r + \frac{1}{1-\alpha}\sum_{s=1}^{S} p_s y_s) \quad : \quad x_1 \in X_1\,,\, r \in \mathbb{R}\,,\, y_s \geq 0\,, \tag{3.19}$$

$$f(x_1,\xi^s) - r - y_s \leq 0\}. \tag{3.20}$$

Für jede Lösung $(x_1^*, r^*, y^*)$ von (3.19)–(3.20) ist dann $(x_1^*, r^*)$ wieder Lösung von (3.18) mit allen dort diskutierten Schlußfolgerungen. Hier läßt sich aber der $\alpha$-Value-at-Risk von $f(x_1^*, \xi)$ explizit berechnen. Wenn wir nämlich annehmen, daß die Szenarios von $\xi$ so numeriert sind, daß die optimalen Szenariokosten in der Form $f(x_1^*, \xi^1) < \ldots < f(x_1^*, \xi^S)$ geordnet sind, so gilt

$$VaR_\alpha(f(x_1^*, \xi)) = f(x_1^*, \xi^{s_*(\alpha)}) \quad \text{wobei} \sum_{s=1}^{s_*(\alpha)} p_s \geq \alpha > \sum_{s=1}^{s_*(\alpha)-1} p_s \,.$$

Die Hilfsvariable $y$ taucht im erweiterten Problem nur in linearen Strukturen auf. Der Risikoterm zur Modellierung von $CVaR$ kann also im Rahmen linearer Optimierungsstrukturen (mit)behandelt werden. Nur die Eigenschaften der ursprünglichen (reduzierten) stochastischen Kostenfunktion $f(x_1, \xi)$ sind für die Struktur von (3.19)–(3.20) von Bedeutung. Ist z.B. $f(x_1, \xi)$ als Funktion von $x_1$ konvex, so ist das Optimierungsproblem (3.19)–(3.20) ebenfalls konvex. Weitere Entwicklungen bei Optimierungsmodellen mit Risikomaßen, insbesondere im dynamischen Fall, sind zu erwarten. Andere Fragestellungen, wie z.B. die Erhaltung von Dekompositionsstrukturen im Fall von Risiko-Zielfunktionen, stehen zur Untersuchung an.

## Literatur


[1] Artzner, P., Delbaen, F.; Eber, J.-M.; Heath, D.: Coherent measures of risk. *Mathematical Finance* 9 (1999), 203–228.

[2] Bacaud, L., Lemaréchal, C., Renaud, A., Sagastizábal, C.: Bundle methods in stochastic optimal power management: A disaggregated approach using preconditioners. *Computational Optimization and Applications* 20 (2001), 227–244.

[3] Birge, J.R.: Stochastic programming computation and applications. *INFORMS Journal Computing* 9 (1997), 111–133.

[4] Birge, J.R., Louveaux, F.: *Introduction to Stochastic Programming.* Springer, New York, 1997.

[5] Bogensperger, J.H.: Wochenplanung in Stromhandel und Erzeugung. In: *Optimierung in der Energieversorgung III*, VDI-Berichte 1508, VDI-Verlag, Düsseldorf 1999, 183–189.

[6] Carøe, C.C.; Schultz, R.: Dual decomposition in stochastic integer programming. *Operations Research Letters* 24 (1999), 37–45.

[7] Carpentier, P.; Cohen, G.; Culioli, J.-C.; Renaud, A.: Stochastic optimization of unit commitment: a new decomposition framework. *IEEE Transactions on Power Systems* 11 (1996), 1067–1073.

[8] Conejo, A.J.; Arroyo, J.M.; Jiménez Redondo, N.; Prieto, F.J.: Lagrangian relaxation applications to electric power operations and planning problems. In [49], 173–203.

[9] Dentcheva, D.; Römisch, W.: Optimal power generation under uncertainty via stochastic programming. *Stochastic Programming Methods and Technical Applications* (K. Marti and P. Kall Eds.), Lecture Notes in Economics and Mathematical Systems Vol. 458, Springer, Berlin 1998, 22–56.

[10] Dupačová, J.; Consigli, G.; Wallace, S.W.: Scenarios for multistage stochastic programs. *Annals of Operations Research* 100 (2000), 25–53.

[11] Dupačová, J.; Gröwe-Kuska, N.; Römisch, W.: Scenario reduction in stochastic programming: An approach using probability metrics. Preprint 00-9, Institut für Mathematik, Humboldt-Universität Berlin, 2000 und erscheint in *Mathematical Programming.*

[12] Feltenmark, S.; Kiwiel, K.C.: Dual applications of proximal bundle methods, including Lagrangian relaxation of nonconvex problems. *SIAM Journal on Optimization* 10 (2000), 697–721.

[13] Fleten, S.-E.; Wallace, S.W.; Ziemba, W.T.: Hedging electricity portfolios via stochastic programming. *Stochastic Programming E-Print Series* 5 (1999) (http://dochost.rz.hu-berlin.de/speps).

[14] Frauendorfer, K.: Barycentric scenario trees in convex multistage stochastic programming. *Mathematical Programming* 75 (1996), 277–294.

[15] Gill, P.E.; Murray, W.; Saunders, M.A.: User's Guide for SNOPT 5.3: A Fortran package for large-scale nonlinear programming. Report NA97-5, Department of Mathematics, University of California, San Diego, 1997.

[16] Gollmer, R.; Möller, A.; Nowak, M.P.; Römisch, W.; Schultz, R.: Primal and dual methods for unit commitment in a hydro-thermal power system. In: Proceedings 13th Power Systems Computation Conference (Trondheim 1999), Vol. 2, 724–730.

[17] Gollmer, R.; Nowak, M.P.; Römisch, W.; Schultz, R.: Unit commitment in power generation – A basic model and some extensions. *Annals Operations Research* 96 (2000), 167–189.

[18] Gröwe-Kuska, N., Kiwiel, K.C., Nowak, M.P., Römisch, W., Wegner, I.: Power management in a hydro-thermal system under uncertainty by Lagrangian relaxation. Preprint 99-19, Institut für Mathematik, Humboldt-Universität Berlin, 1999 und erscheint in IMA Volume *Decision Making under Uncertainty: Energy and Environmental Models*, Springer-Verlag.

[19] Gröwe-Kuska, N.; Nowak, M.P.; Römisch, W.; Wegner, I.: Optimierung eines hydro-thermischen Kraftwerkssystems unter Ungewißheit. In: *Optimierung in der Energieversorgung III*, VDI-Berichte 1508, VDI-Verlag, Düsseldorf 1999, 147–157.

[20] Gröwe-Kuska, N.; Nowak, M.P.; Wegner, I.: Modeling of uncertainty for the real-time management of power systems. In: *Online Optimization of Large Scale Systems* (Grötschel, M., Krumke, S., Rambau, J. eds.), Springer-Verlag, Berlin 2001, 621–645.

[21] Gröwe-Kuska, N.; Heitsch, H.; Römisch, W.: Modellierung stochastischer Datenprozesse für Optimierungsmodelle der Energiewirtschaft. In: *IT-Lösungen für die Energiewirtschaft* (erscheint).

[22] Guignard, M.; Kim, S.: Lagrangean decomposition: a model yielding stronger lagrangean bounds. *Mathematical Programming* 39 (1987), 215–228.

[23] Heitsch, H.; Römisch, W.: Scenario reduction algorithms in stochastic programming. Preprint 01-8, Institut für Mathematik, Humboldt-Universität Berlin, 2001.

[24] Hiriart-Urruty, J.-B.; Lemaréchal, C.: *Convex Analysis and Minimization Algorithms I and II*, Springer-Verlag, Berlin 1993.

[25] Jaschke, S.; Küchler, U.: Coherent risk measures and good-deal bounds. *Finance and Stochastics* 5 (2001), 181–200.

[26] Kiwiel, K.C.: Proximity control in bundle methods for convex nondifferentiable minimization. *Mathematical Programming* 46 (1990), 105–122.

[27] Kiwiel, K.C.: User's Guide for NOA 2.0/3.0: A Fortran Package for Convex Nondifferentiable Optimization. Polish Academy of Science, System Research Institute, Warsaw, 1993/1994.

[28] Klein Haneveld, W.K.; van der Vlerk, M.H.: Stochastic integer programming: general models and algorithms. In [54], 39–57.

[29] Krasenbrink, B.; Nießen, S.; Haubrich, H.-J.: Risikomanagement in Stromerzeugung und -handel. In: *Optimierung in der Energieversorgung III*, VDI-Berichte 1508, VDI-Verlag, Düsseldorf 1999, 53–62.

[30] Lemaréchal, C., Renaud, A.: A geometric study of duality gaps, with applications. *Mathematical Programming* 90 (2001), 399–427.

[31] Lemaréchal, C., Sagastizábal, C.: Variable metric bundle methods: from conceptual to implementable forms. *Mathematical Programming* 76 (1997), 393–410.

[32] Mifflin, R.: A quasi-second-order proximal bundle algorithm. *Mathematical Programming* 73 (1996), 51–72.

[33] Morton, D.P.: An enhanced decomposition algorithm for multistage stochastic hydroelectric scheduling. *Annals Operations Research* 64 (1996), 211–235.

[34] Murtagh, B.A.; Saunders, M.A.: MINOS 5.5 User's Guide, Report SOL 83-20R, Department of Operations Research, Stanford University, Stanford, CA, revised 1998.

[35] Nießen, S.: Kraftwerkseinsatz- und Handelsplanung im liberalisierten Strommarkt. *Aachener Beiträge zur Energieversorgung*, Band 54, Aachen 1998.

[36] Nocedal, J.; Wright, S.J.: *Numerical Optimization*. Springer-Verlag, New York 1999.

[37] Nowak, M.P.; Römisch, W.: Stochastic Lagrangian relaxation applied to power scheduling in a hydro-thermal system under uncertainty. *Annals Operations Research* 100 (2000), 251–272.

[38] Ogryczak, W.; Ruszczyński, A.: From stochastic dominance to mean-risk models: semideviations as risk measures. *European Journal Operational Research* 116 (1999), 33–50.

[39] Ogryczak, W.; Ruszczyński, A.: Dual stochastic dominance and related mean-risk models. *Stochastic Programming E-Print Series* 2 (2001) (http://dochost.rz.hu-berlin.de/speps) und erscheint in *SIAM Journal Optimization*.

[40] Pflug, G.Ch.: Some remarks on the value-at-risk and the conditional value-at-risk. In: *Probabilistic Constrained Optimization: Methodology and Applications* (Uryasev, S. ed.), Kluwer, 2000, 272–281.

[41] Philpott, A.B.; Craddock, M.; Waterer, H.: Hydro-electric unit commitment subject to uncertain demand. *European Journal Operational Research* 125 (2000), 410–424.

[42] Rockafellar, R.T.; Uryasev, S.: Optimization of conditional value-at-risk. *Journal of Risk* 2 (2000), 21–41.

[43] Rockafellar, R.T.; Uryasev, S.: Conditional value-at-risk for general loss distributions. Research Report #2001-5, Center for Applied Optimization, University of Florida, Gainesville, 2001.

[44] Rockafellar, R.T.; Wets, R.J-B: The optimal recourse problem in discrete time: $L^1$-multipliers for inequality constraints. *SIAM Journal Control Optimization* 16 (1978), 16–36.

[45] Rockafellar, R.T.; Wets, R.J-B: Scenarios and policy aggregation in optimization under uncertainty. *Mathematics Operations Research* 16 (1991), 119–147.

[46] Römisch, W.; Schultz, R.: Multistage stochastic integer programs: An introduction. In: *Online Optimization of Large Scale Systems* (Grötschel, M.; Krumke, S.; Rambau, J. eds.), Springer-Verlag, Berlin 2001, 579–598.

[47] Ruszczyński, A.: Decomposition methods in stochastic programming. *Mathematical Programming* 79 (1997), 333–353.

[48] Sheble, G.B.; Fahd, G.N.: Unit commitment literature synopsis. *IEEE Transactions on Power Systems* 9 (1994), 128–135.

[49] Song, Y.H. (ed.): *Modern Optimisation Techniques in Power Systems*. Kluwer, Dordrecht 1999.

[50] Takriti, S.; Birge, J.R.; Long, E.: A stochastic model for the unit commitment problem. *IEEE Transactions Power Systems* 11 (1996), 1497–1508.

[51] Takriti, S.; Birge, J.R.: Using integer programming to refine Lagrangian-based unit commitment solutions. *IEEE Transaction Power Systems* 15 (2000), 151–156.

[52] Takriti, S.; Birge, J.R.: Lagrangian solution techniques and bounds for loosely coupled mixed-integer stochastic programs. *Operations Research* 48 (2000), 91–98.

[53] Takriti, S.; Krasenbrink, B.; Wu, L.S.-Y.: Incorporating fuel constraints and electricity spot prices into the stochastic unit commitment problem. *Operations Research* 48 (2000), 268–280.

[54] Wets, R.J-B, Ziemba, W.T.: Stochastic programming. State of the Art, 1998. *Annals Operations Research* 85 (1999).

[55] Wolsey, L.A.: *Integer Programming*, Wiley, New York 1998.

[56] Wood, A.J.; Wollenberg, B.F.: *Power Generation, Operation, and Control*. Wiley, New York, 1996.

[57] Zhuang, F.; Galiana, F.D.: Towards a more rigorous and practical unit commitment by Lagrangian relaxation. *IEEE Transactions Power Systems* 3 (1988), 763–773.