Stochastic Lagrangian Relaxation applied to Power Scheduling in a Hydro-Thermal System under Uncertainty

Matthias P. Nowak and Werner Römisch

Humboldt University Berlin
Institute of Mathematics
10099 Berlin, Germany

A dynamic (multi-stage) stochastic programming model for the weekly cost-optimal generation of electric power in a hydro-thermal generation system under uncertain demand (or load) is developed. The model involves a large number of mixed-integer (stochastic) decision variables and constraints linking time periods and operating power units. A stochastic Lagrangian relaxation scheme is designed by assigning (stochastic) multipliers to all constraints coupling power units. It is assumed that the stochastic load process is given (or approximated) by a finite number of realizations (scenarios) in scenario tree form. Solving the dual by a bundle subgradient method leads to a successive decomposition into stochastic single (thermal or hydro) unit subproblems. The stochastic thermal and hydro subproblems are solved by a stochastic dynamic programming technique and by a specific descent algorithm, respectively. A Lagrangian heuristics that provides approximate solutions for the first stage (primal) decisions starting from the optimal (stochastic) multipliers is developed. Numerical results are presented for realistic data from a German power utility and for numbers of scenarios ranging from 5 to 100 and a time horizon of 168 hours. The sizes of the corresponding optimization problems go up to 200.000 binary and 350.000 continuous variables, and more than 500.000 constraints.

1. Introduction

Mathematical models for the efficient operation of electric power generation systems often lead to rather complex optimization problems. In particular, they are characterized by combinations of challenges like mixed-integer decisions, nonlinear costs, large dimensions and data uncertainty. The latter aspect mostly concerns uncertainties of electrical load forecasts, of generator failures, of flows to hydro reservoirs or plants, and of fuel or electricity prices (cf. [12,13,15,18,29] for earlier relevant work). The present paper aims at treating power optimization in a hydro-thermal system under uncertain electrical load. More precisely, a generation system comprising thermal units and pumped hydro storage plants as encountered at the German utility VEAG Vereinigte Energiewerke AG Berlin is considered. The relevant mathematical optimization model contains a large number of binary and continuous variables, constraints and uncertainty appearing in the load constraints. The time horizon is about 7 days as it is needed for the efficient weekly operation of hydro-thermal systems involving weekly load and pumping cycles.

The machinery of stochastic programming offers modelling and solution techniques for such optimization problems under uncertainty. In the present paper, a multi-stage stochastic programming model in which the expected production costs are minimized and stages refer to the availability of further observations of the load is developed. In particular, the first stage refers to the time period for which a reliable load forecast is available. The attention is focused on the (deterministic) first-stage scheduling decisions (on/off and outputs), which are obtained by minimizing the total expected generation costs and, hence, hedge against uncertainty. Since the stochastic programming model contains mixed-integer decisions in all stages and is large-scale, new questions on the design of solution algorithms are raised.

Nowadays, solution methods are well developed for linear multi-stage stochastic programs without integrality constraints (cf. the monographs [3,17,16,38] and the state-of-the-art surveys [2,34]). Recently, progress has been made for mixed-integer stochastic programming models and applications to power optimization. The following algorithmic approaches for mixed-integer multi-stage models appear in the literature: (a) Stochastic branch and bound methods ([26]), (b) scenario decomposition by splitting methods combined with suitable heuristics ([32], [25], [36,37]), (c) scenario decomposition combined with branch and bound ([6,7]), (d) stochastic (augmented) Lagrangian relaxation of coupling constraints ([8,30], [10,33]). The approaches in (b) and (c) are based on a successive decomposition of the stochastic program into finitely many deterministic (or scenario) programs, which may be solved by available conventional techniques. The idea of (d) is a successive decomposition into finitely many smaller stochastic subproblems for which (efficient) solution techniques have to be developed eventually. Due to the nonconvexity of the underlying stochastic program, the successive decompositions in (b)-(d) have to be combined with certain global optimization techniques (branch-and-bound, heuristics etc.).

The approach followed in the present paper consists in a *stochastic version* of the classical Lagrangian relaxation idea ([23]), which is very popular in power optimization ([1,11,14,24,35,39,40]). Since the corresponding coupling constraints contain random variables, stochastic multipliers are needed for the dualization, and the dual problem represents a nondifferentiable stochastic program. Subsequently, the approach is based on the same, but *stochastic*, ingredients as in the classical case: a solver for the nondifferentiable dual, subproblem solvers, and a Lagrangian heuristics. It turns out that, with a state-of-the-art bundle method for solving the dual, efficient stochastic subproblem solvers based on a specific

descent algorithm and stochastic dynamic programming, respectively, and a specific Lagrangian heuristics for determining a nearly optimal first-stage solution, this *stochastic Lagrangian relaxation* algorithm becomes efficient.
The paper is organized as follows. In Section 2 a detailed description of the hydro-thermal generation system is given and the stochastic programming model is developed. Section 3 describes the stochastic Lagrangian relaxation approach together with its components: algorithms for solving the stochastic dual, single-unit and economic dispatch problems, and the Lagrangian heuristics. Numerical experience is provided for all (sub)algorithms. Finally, numerical results for the stochastic Lagrangian relaxation based algorithm are reported in Section 4 for realistic data of the VEAG system.

2. Model

We consider a power generation system comprising (coal-fired and gas-burning) thermal units, pumped hydro storage plants and delivery contracts, and describe a model for its weekly cost-optimal generation under uncertainty on the electrical load (cf. [10], [28]). Let T denote the number of time intervals obtained from a uniform discretization of the operation horizon. Let I and J denote the number of thermal and pumped hydro storage units in the system, respectively. Delivery contracts are regarded as particular thermal units. The decision variables in the model correspond to the outputs of units, i. e., the electric power generated or consumed by each unit of the system. They are denoted by $\mathbf{u}_i^t$, $\mathbf{p}_i^t$, $i=1,\ldots,I$, and $\mathbf{s}_j^t, \mathbf{w}_j^t$, $j=1,\ldots,J$, $t=1,\ldots,T$, where $\mathbf{u}_i^t \in \{0,1\}$ and $\mathbf{p}_i^t$ are the on/off decisions and the production levels of the thermal unit i during the time period t. Thus, $\mathbf{u}_i^t=0$ and $\mathbf{u}_i^t=1$ mean that the unit i is off-line and on-line during period t, respectively. $\mathbf{s}_j^t$, $\mathbf{w}_j^t$ are the generation and pumping levels of the pumped hydro storage plant j during the period t, respectively. Further, by $\mathbf{l}_j^t$ we denote the storage level (or volume) in the upper reservoir of plant j at the end of the interval t. All variables mentioned above have finite upper and lower bounds representing unit limits and reservoir capacities of the generation system:

$$
\begin{array}{l}
p_i^{min}\mathbf{u}_i^t \le p_i^t \le p_i^{max}\mathbf{u}_i^t,\ \mathbf{u}_i^t \in \{0,1\},\ i=1,\ldots,I,\ t=1,\ldots,T,\\
0 \le \mathbf{s}_j^t \le s_j^{max},\ 0 \le \mathbf{w}_j^t \le w_j^{max},\\
0 \le \mathbf{l}_j^t \le l_j^{max},\ j=1,\ldots,J,\ t=1,\ldots,T.
\end{array}
\tag{1}
$$

The constants $p_i^{min}, p_i^{max}, s_j^{max}, w_j^{max}$, and l_j^{max} denote the minimal/maximal outputs of the units and the maximal storage levels in the upper reservoirs, respectively. The dynamics of the storage level, which is measured in electrical energy, is modelled by the equations:

$$
\begin{array}{l}
\mathbf{l}_j^t = \mathbf{l}_j^{t-1} - \mathbf{s}_j^t + \eta_j \mathbf{w}_j^t\,, t=1,\ldots,T,\\
\mathbf{l}_j^0 = l_j^{in},\quad \mathbf{l}_j^T = l_j^{end}\quad, j=1,\ldots,J.
\end{array}
\tag{2}
$$

Here, l_j^{in} and l_j^{end} denote the initial and final levels in the upper reservoir, respectively, and η_j is the cycle (or pumping) efficiency of plant j. The cycle efficiency is defined as the quotient of the generation and of the pumping load that correspond to the same amount of water. The equalities (2) show, in particular, that there occur no in- or outflows in the upper reservoirs and, hence, that the storage plants of the system operate with a constant amount of water. Together with the upper and lower bounds for $\mathbf{l}_j^t$ the equations (2) mean that certain reservoir constraints have to be maintained for all storage plants during the whole time horizon. Further single-unit constraints are minimum up- and down-times and possible must-on/off constraints for each thermal unit. Minimum up- and down-time constraints are imposed to prevent thermal stress and high maintenance costs due to excessive unit cycling. Denoting by τ_i the minimum down-time of unit i, the corresponding constraints are described by the inequalities:

$$\mathbf{u}_i^{t-1} - \mathbf{u}_i^t \leq 1 - \mathbf{u}_i^\tau, \quad \tau = t+1, \ldots, \min\{t + \tau_i - 1, T\},\ t = 1, \ldots, T, \tag{3}$$

Analogous constraints can be formulated describing minimum up-times. The next constraints are coupling across power units: the load and reserve constraints. The first constraints are essential for the operation of the power system and express that the sum of the output powers is greater than or equal to the load demand in each time period. Denoting by $\mathbf{d}^t$ the electrical load (or demand) during period t, the load constraints are described by the inequalities:

$$\sum_{i=1}^{I} \mathbf{p}_i^t + \sum_{j=1}^{J} (\mathbf{s}_j^t - \mathbf{w}_j^t) \geq \mathbf{d}^t,\ t = 1, \ldots, T. \tag{4}$$

In order to compensate unexpected events (e.g. sudden load increases or decreases, outages of units) within a specified short time period, a spinning reserve describing the total amount of generation available from all units synchronized on the system minus the present load is prescribed. The corresponding constraints are given by the following inequalities:

$$\sum_{i=1}^{I} (p_i^{max} \mathbf{u}_i^t - \mathbf{p}_i^t) \geq \mathbf{r}^t,\ t = 1, \ldots, T, \tag{5}$$

where $\mathbf{r}^t > 0$ is the spinning reserve in period t, which is assumed to be proportional to $\mathbf{d}^t$. The objective function is given by the total costs for operating the thermal units. These costs consist of the sum of the costs of each individual unit over the whole time horizon, i. e.,

$$\sum_{i=1}^{I} \sum_{t=1}^{T} [C_i(\mathbf{p}_i^t,\ \mathbf{u}_i^t) + S_i^t(\mathbf{u}_i)], \tag{6}$$

where C_i are the fuel costs for the operation of the thermal unit i during period t and S_i^t are the start-up costs for getting the unit on-line in this period. We assume that each C_i is piecewise linear convex, strictly monotonically increasing and of the form

$$C_i(\mathbf{p},\,\mathbf{u}) = \max_{l=1,\ldots,L} \{a_{il}\,\mathbf{p} + b_{il}\,\mathbf{u}\}\ , \tag{7}$$

where a_{il} and b_{il} are fixed cost coefficients. The start-up costs $S_i^t(\mathbf{u}_i)$ may vary from a maximum cold-start value to a much smaller value when the unit i is still relatively close to its operation temperature. The following description of start-up costs reflects this dependence on the down-time:

$$S_i^t(\mathbf{u}_i) = \max_{\tau=0,\ldots,\tau_i^c} c_i^\tau\,(\mathbf{u}_i^t - \sum_{\kappa=1}^{\tau} \mathbf{u}_i^{t-\kappa}),$$

where $c_i^0 = 0$ and c_i^τ, $\tau = 0,\ldots,\tau_i^c$ are fixed increasing cost coefficients, τ_i^c is the time the unit i needs to cool down, and $c_i^{\tau_i^c}$ its maximum cold-start costs. Altogether, minimizing the objective function (6) subject to the constraints (1)-(5) leads to a cost-optimal schedule for all units of the power system during the specified time horizon. It is worth mentioning that a cost-optimal schedule has the following two interesting properties, which are both a consequence of the strict monotonicity of the fuel costs. If a schedule $(\mathbf{u},\,\mathbf{p},\,\mathbf{s},\,\mathbf{w})$ is optimal, then the load constraints (4) are typically satisfied with equality and we have $\mathbf{s}_j^t\mathbf{w}_j^t = 0$ for all $j = 1,\ldots,J$, $t = 1,\ldots,T$, i. e., generation and pumping do not occur simultaneously (cf. [15]).

The minimization problem (1)-(6) represents a mixed-integer program with linear constraints, and IT binary and $(I + 2J)T$ continuous decision variables, respectively. For a typical configuration of the VEAG-owned generation system with $I = 25$ (thermal), $J = 7$ (hydro) and $T = 168$ (i. e., 7 days with hourly discretization), the dimension of the model is shown in the first row of Fig. 1.

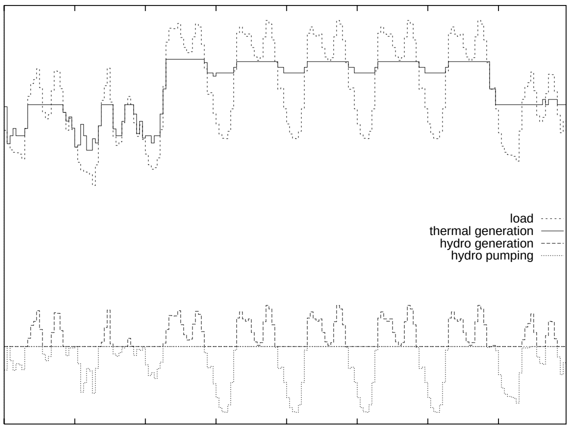

Figure 1: Load curve and hydro-thermal schedule

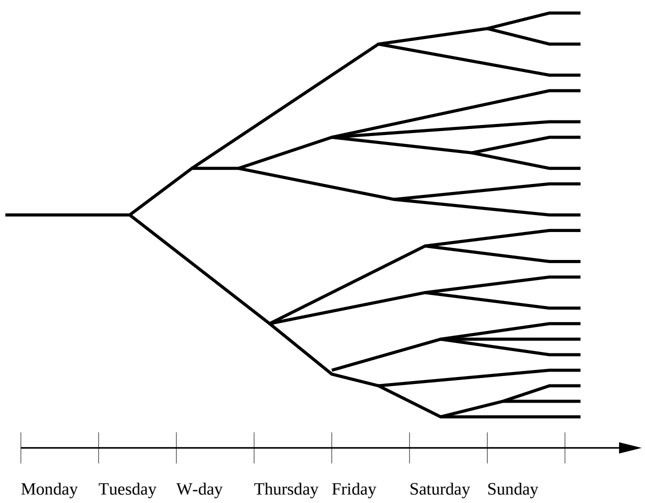

Figure 2: Example of a scenario tree

Fig. 1 shows a typical load curve of a peak load week and a corresponding cost-optimal hydro-thermal schedule. The load curve in Fig. 1 exhibits two overlapping cycles: a daily and weekly cycle. Pumped hydro storage plants are designed to exploit these two cycles by saving fuel costs when serving the peak load with hydro-energy and pumping to refill the reservoir during off-peak periods, i. e., during the nights and weekends. The hydro schedule in Fig. 1 reflects this typical operation of pumped hydro storage plants. The remaining load, i. e., the difference between the original system load and the hydro schedule, shows a more uniform structure than the original load. This portion of the load is covered by the total output of thermal units. So far we have tacitly assumed that the electrical load is given and deterministic over the whole time horizon. In electric utilities, schedulers forecast the electrical load for each time period of the day or week in advance. But, clearly, the actual electrical load may deviate from the predicted load at any time period due to various unforeseeable (random) influences (temperature, daylight, switch off of local consumers etc.). This gives rise to a stochastic model of the electrical load $\{\mathbf{d}^t : t = 1, \ldots, T\}$ as a (discrete-time) stochastic process on some probability space $(\Omega, \mathcal{A}, I\!\!P)$ reflecting that the information on the load is complete for $t = 1$, and that the uncertainty increases with growing t. Let $\{\mathcal{A}_t\}_{t=1}^T$ be the filtration generated by the load process, i.e., $\mathcal{A}_t$ is the σ-field generated by the random vector $(\mathbf{d}^1, \ldots, \mathbf{d}^t)$. Hence, we have $\{\emptyset, \Omega\} = \mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \ldots \subseteq \mathcal{A}_t \subseteq \ldots \subseteq \mathcal{A}_T \subseteq \mathcal{A}$. The sequence of scheduling decisions $\{(\mathbf{u}^t, \mathbf{p}^t, \mathbf{s}^t, \mathbf{w}^t) : t = 1, \ldots, T\}$ also forms a stochastic process on $(\Omega, \mathcal{A}, I\!\!P)$, which is assumed to be adapted to the filtration of σ-fields, i.e., *non-anticipative*. The latter condition means that the decision $(\mathbf{u}^t, \mathbf{p}^t, \mathbf{s}^t, \mathbf{w}^t)$ depends only on the data history $(\mathbf{d}^1, \ldots, \mathbf{d}^t)$ or, equivalently, that $(\mathbf{u}^t, \mathbf{p}^t, \mathbf{s}^t, \mathbf{w}^t)$ is $\mathcal{A}_t$-measurable. Since all decision variables are uniformly bounded, we may restrict our attention to decisions $(\mathbf{u}, \mathbf{p}, \mathbf{s}, \mathbf{w})$ belonging to $L^\infty(\Omega, \mathcal{A}, I\!\!P; I\!\!R^m)$, where $m := 2(I+J)T$. Then the non-anticipativity condition can be formulated equivalently as

$$(\mathbf{u}, \mathbf{p}, \mathbf{s}, \mathbf{w}) \in \mathop{\times}_{t=1}^{T} L^\infty\left(\Omega, \mathcal{A}_t, I\!\!P; I\!\!R^{2(I+J)}\right), \tag{8}$$

and the (stochastic) optimization problem consists in minimizing the expected costs

$$I\!\!E\{\sum_{i=1}^{I} \sum_{t=1}^{T} [C_i(\mathbf{p}_i^t, \mathbf{u}_i^t) + S_i^t(\mathbf{u}_i)]\} \tag{9}$$

over all decisions $(\mathbf{u}, \mathbf{p}, \mathbf{s}, \mathbf{w})$ satisfying the non-anticipativity constraint (8), and $I\!\!P$-almost surely, the constraints (1)-(5). Among the constraints (1)-(5), (2) and (3) reflect the dynamics of the model and (4), (5) couple power units. Altogether, the stochastic program involves $2(I+J)T$ stochastic decision variables. It is a discrete-time dynamic or multi-stage stochastic recourse problem, where the *stages* correspond to steps in the decision process at which new observations of the stochastic load are taken into account. For the numerical solution of the dynamic recourse model we now assume that an (approximate) *discrete* multivariate probability distribution of the stochastic load vector $\mathbf{d} = (\mathbf{d}^1, \ldots, \mathbf{d}^T)$ is given, such that its support consists of finitely many atoms or *scenarios* and that the non-anticipativity constraint (8) is satisfied. This approximation of the load can be represented in the form of a *scenario tree.* Each path of the tree from the root to a leaf corresponds to one scenario; each node of the tree corresponds to a component of the decision $(\mathbf{u}, \mathbf{p}, \mathbf{s}, \mathbf{w})$. Fig. 2 shows an example of a load scenario tree over a weekly time horizon, where new observations of the electrical load lead to a number of additional daily scenarios. Since the decision variable $(\mathbf{u}, \mathbf{p}, \mathbf{s}, \mathbf{w})$ exhibits the same tree structure as the load, the model may easily become extremely large if the number of nodes in the scenario tree increases.

Scenarios	Nodes	Variables		Constraints	Nonzeros
		binary	continuous		
1	168	4200	6652	13441	19657
10	756	18900	29484	60490	88462
20	1176	29400	45864	94100	137612
30	1663	41575	64857	133070	194601
50	2478	61950	96642	198290	289976
100	4200	105000	163800	336100	491500

Table 1
Dimension of the mixed-integer LP depending on the numbers of nodes and scenarios with T=168, I=25 and J=7

Table 1 shows how the dimension of the model (1)-(5), (8), (9) increases with the number of nodes and scenarios for a scenario tree with equidistant binary branches (without taking into account the constraints of type (3) and the objective function).

3. Stochastic Lagrangian Relaxation

The huge size of the model, described in the previous section, prevents the application of state-of-the-art mixed-integer LP solvers. However, decomposition techniques may provide a practicable alternative. Here, we make use of the fact that the model is loosely coupled with respect to the operation of different units. Associating stochastic Lagrange multipliers with the coupling constraints (4) and (5) leads to the Lagrangian L and the dual function D:

$$L(\mathbf{u},\mathbf{p},\mathbf{s},\mathbf{w};\boldsymbol{\lambda}) = \mathbb{E}\sum_{t=1}^{T}\{\sum_{i=1}^{I}[C_i(\mathbf{p}_i^t,\mathbf{u}_i^t)+S_i^t(\mathbf{u}_i)] \tag{10}$$

$$+\boldsymbol{\lambda}_1^t(\mathbf{d}^t-\sum_{i=1}^{I}\mathbf{p}_i^t-\sum_{j=1}^{J}(\mathbf{s}_j^t-\mathbf{w}_j^t))+\boldsymbol{\lambda}_2^t(\mathbf{r}^t-\sum_{i=1}^{I}(\mathbf{u}_i^t p_i^{max}-\mathbf{p}_i^t)\}$$

$$D(\boldsymbol{\lambda}) = \min_{(\mathbf{u},\mathbf{p},\mathbf{s},\mathbf{w})} L(\mathbf{u},\mathbf{p},\mathbf{s},\mathbf{w};\boldsymbol{\lambda}) \ , \tag{11}$$

where the minimization in (11) is subject to the remaining single unit constraints (1), (2), (3) and (8). Justified by general duality results for convex multi-stage stochastic programs (see [31] and Sect. 4 of [10]) we consider the dual problem

$$\max\{D(\boldsymbol{\lambda}) : \boldsymbol{\lambda} = (\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) \in \mathop{\times}_{t=1}^{T} L^1(\Omega,\,\mathcal{A}_t,\,\mathbb{P}\,;\mathbb{R}_+^2)\}\,. \tag{12}$$

In particular, this means that the stochastic multiplier process $\boldsymbol{\lambda}$ is nonnegative $\mathbb{P}$-almost surely and adapted to the filtration $\{\mathcal{A}_t\}_{t=1}^T$ generated by the load process. Hence, they exhibit the same tree structure as $\mathbf{d}$. Furthermore, the dimension of the dual problem (12) is twice the number N of nodes in the scenario tree. The optimal value of the dual problem (12) provides a lower bound for the optimal costs of the nonconvex (primal) model. For a discussion of the (relative) duality gap in our context of power optimization, the reader is referred to [1], [24] and Sect. 4 of [10]. Due to the relaxation of the coupling constraints (4) and (5), the minimization in (11) decomposes into stochastic single unit subproblems and the dual function takes the form

$$D(\boldsymbol{\lambda}) = \sum_{i=1}^{I} D_i(\boldsymbol{\lambda}) + \sum_{j=1}^{J}\hat{D}_j(\boldsymbol{\lambda}_1) + \mathbb{E}\sum_{t=1}^{T}[\boldsymbol{\lambda}_1^t\mathbf{d}^t+\boldsymbol{\lambda}_2^t\mathbf{r}^t] \ , \tag{13}$$

where $D_i(\boldsymbol{\lambda})$ and $\hat{D}_j(\boldsymbol{\lambda}_1)$ refer to the optimal values of the thermal and hydro subproblems, respectively. They have the following form:

$$D_i(\boldsymbol{\lambda}) = \min_{\mathbf{u}_i} \Big\{ \mathbb{E} \sum_{t=1}^{T} (\min_{\mathbf{p}_i^t} [C_i(\mathbf{p}_i^t, \mathbf{u}_i^t) - (\boldsymbol{\lambda}_1^t - \boldsymbol{\lambda}_2^t)\, \mathbf{p}_i^t] - \boldsymbol{\lambda}_2^t \mathbf{u}_i^t p_i^{max} + S_i^t(\mathbf{u}_i)) : \tag{14}$$

$$\mathbf{u}_i^t p_i^{min} \le \mathbf{p}_i^t \le \mathbf{u}_i^t p_i^{max},\ \mathbf{u}_i^t \in \{0,1\},\ (\mathbf{u}_i, \mathbf{p}_i) \in \mathop{\times}_{t=1}^{T} L^\infty(\Omega,\, \mathcal{A}_t,\, I\!\!P; I\!\!R^2),$$

$$\mathbf{u}_i^{t-1} - \mathbf{u}_i^t \le 1 - \mathbf{u}_i^\tau,\ \tau = t+1, \ldots, \min\{t + \tau_i - 1,\, T\},\ t = 1, \ldots, T \Big\}$$

$$\hat{D}_j(\boldsymbol{\lambda}_1) = \min_{(\mathbf{s}_j, \mathbf{w}_j)} \Big\{ \mathbb{E} \sum_{t=1}^{T} [\boldsymbol{\lambda}_1^t (\mathbf{w}_j^t - \mathbf{s}_j^t)] : (\mathbf{s}_j, \mathbf{w}_j) \in \mathop{\times}_{t=1}^{T} L^\infty(\Omega,\, \mathcal{A}_t,\, I\!\!P; I\!\!R^2)\,, \tag{15}$$

$$0 \le \mathbf{s}_j^t \le s_j^{max},\ 0 \le \mathbf{w}_j^t \le w_j^{max},\ 0 \le \mathbf{l}_j^t \le l_j^{max},\ t = 1, \ldots, T, \tag{16}$$

$$\mathbf{l}_j^t = \mathbf{l}_j^{t-1} - \mathbf{s}_j^t + \eta_j \mathbf{w}_j^t,\ t = 1, \ldots, T,\ \mathbf{l}_j^0 = l_j^{in},\ \mathbf{l}_j^T = l_j^{end} \Big\}\,. \tag{17}$$

The thermal subproblem (14) for unit i is a mixed-integer multi-stage stochastic program. But, it reduces to a combinatorial multi-stage stochastic program, since the inner minimization with respect to the one-dimensional continuous variable $\mathbf{p}_i^t$ can be carried out explicitly by examining the kinks of the fuel costs C_i. The hydro subproblem (15) for plant j is a linear multi-stage stochastic program. Altogether, the dual function D is concave and nondifferentiable on $I\!\!R^{2N}$, and polyhedral due to (7).

Similar to the deterministic case, the *stochastic Lagrangian relaxation algorithm* for solving the model in Section 2 consists of the following ingredients:

(a) Maximization of the dual function D by a proximal bundle method using function and subgradient information (Sect. 3.1);

(b) Efficient solvers for the stochastic single unit subproblems: stochastic dynamic programming (Sect. 3.2) and a specific descent algorithm (Sect. 3.3);

(c) Lagrangian heuristics for finding a feasible first-stage decision (Sect. 3.4);

(d) Economic dispatch for determining a nearly optimal first-stage decision (Sect. 3.5).

In the remaining part of this section we provide a description of these ingredients.

3.1. Proximal Bundle Method

We consider the maximization of the dual concave function D on the set $I\!\!R_+^{2N}$, and assume that the set of maximizers is nonempty. Function values $D(\boldsymbol{\lambda})$ are evaluated according to (13) and a corresponding subgradient $g(\boldsymbol{\lambda}) \in \partial D(\boldsymbol{\lambda})$ is given by $(g_1(\boldsymbol{\lambda}), \ldots, g_n(\boldsymbol{\lambda}), g_{n+1}(\boldsymbol{\lambda}), \ldots, g_{2N}(\boldsymbol{\lambda}))$, where $g_n(\boldsymbol{\lambda})$ for $n = 1, \ldots, N$ is equal to the value of the stochastic process $\{\mathbf{d}^t - \sum_{i=1}^{I} \mathbf{p}_i^t(\boldsymbol{\lambda}) -$

$\sum_{j=1}^{J}(\mathbf{s}_j^t(\boldsymbol{\lambda})) - \mathbf{w}_j^t(\boldsymbol{\lambda}))\}_{t=1}^T$ at node n and $g_{N+n}(\boldsymbol{\lambda})$ for $n = 1, \ldots, N$ is equal to the value of the stochastic process $\{\mathbf{r}^t - \sum_{i=1}^{I}(\mathbf{u}_i^t(\boldsymbol{\lambda})p_i^{max} - \mathbf{p}_i^t(\boldsymbol{\lambda}))\}_{t=1}^T$ at node n. Here, $(\mathbf{u}(\boldsymbol{\lambda}), \mathbf{p}(\boldsymbol{\lambda}), \mathbf{s}(\boldsymbol{\lambda}), \mathbf{w}((\boldsymbol{\lambda}))$ is a Lagrangian solution, i.e, it belongs to $\arg\min_{(\mathbf{u},\mathbf{p},\mathbf{s},\mathbf{w})} L(\mathbf{u}, \mathbf{p}, \mathbf{s}, \mathbf{w}; \boldsymbol{\lambda})$.
The proximal bundle method [19,21] generates a sequence $(\boldsymbol{\lambda}^k)$ in $\mathbb{R}_+^{2N}$ converging to some maximizer, and trial points $\bar{\boldsymbol{\lambda}}^k \in \mathbb{R}_+^{2N}$ starting with $\bar{\boldsymbol{\lambda}}^1 = \boldsymbol{\lambda}^1$ for evaluating subgradients $g(\bar{\boldsymbol{\lambda}}^k)$ of D and its polyhedral upper approximation

$$\tilde{D}_k(\boldsymbol{\lambda}) = \min_{j \in J^k} \left\{ D(\bar{\boldsymbol{\lambda}}^j) + < g(\bar{\boldsymbol{\lambda}}^j), \boldsymbol{\lambda} - \bar{\boldsymbol{\lambda}}^j > \right\} , \tag{18}$$

where J^k is a subset of $\{1, ..., k\}$. In iteration k the next trial point $\bar{\boldsymbol{\lambda}}^{k+1}$ is selected by

$$\bar{\boldsymbol{\lambda}}^{k+1} \in \arg\max\{\tilde{D}_k(\boldsymbol{\lambda}) - \frac{1}{2}u_k\|\boldsymbol{\lambda} - \boldsymbol{\lambda}^k\|^2 : \boldsymbol{\lambda} \in \mathbb{R}_+^{2N}\} \tag{19}$$

where u_k is a proximity weight. A descent step to $\boldsymbol{\lambda}^{k+1} = \bar{\boldsymbol{\lambda}}^{k+1}$ occurs if $D(\bar{\boldsymbol{\lambda}}^{k+1}) \geq D(\boldsymbol{\lambda}^k) + \kappa\delta_k$, where $\kappa \in (0,1)$ is fixed and $\delta_k = \tilde{D}_k(\bar{\boldsymbol{\lambda}}^{k+1}) - D(\boldsymbol{\lambda}^k) \geq 0$. If $\delta_k = 0$, then $\boldsymbol{\lambda}^k$ is optimal. Otherwise, a null step $\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k$ improves the next polyhedral function $\tilde{D}_{k+1}$. Strategies for updating u_k and choosing J^{k+1} are discussed in [19,21]. The method is implemented such that the cardinality of J^k is bounded (by some natural number NGRAD) and that it terminates if δ_k is less than a given (relative) optimality tolerance opt.tol.
Our computational experience with the proximal bundle code NOA 3.0 ([20]) for solving (12) is very encouraging (cf. Section 4). In our test runs, for instance, NOA 3.0 applied to solving (12) performed in 300 iterations as good as a standard subgradient method (with step lengths $\frac{1}{k}$) in 10.000 iterations.

3.2. Stochastic Dynamic Programming

In order to solve the thermal subproblem (14) by dynamic programming, the state space is extended by including the recent history such that minimum up/down-times and start-up costs depend just on the current and the previous state. Fig. 3 shows a part of the state transition graph of a thermal unit having a minimum up-time of 6 hours, a minimum down-time of 5 hours, and a cooling down time of 8 hours. It shows possible and feasible transitions on some fixed arc of the scenario tree, where the arrows refer to feasible transitions. Let $\alpha_i^t(s)$ denote the node weight at time t and state s and $\widehat{S_i}(s, \tilde{s})$ the arc weight for the arc from state s to state $\tilde{s}$ in the state transition graph. The node weights $\alpha_i^t(s)$ are equal to 0 for off-line states s and it holds

$$\alpha_i^t(s) = \min_{p_i}\{C_i(p_i, 1) - (\boldsymbol{\lambda}_1^t - \boldsymbol{\lambda}_2^t)p_i - \boldsymbol{\lambda}_2^t p_i^{max} \ : \ p_i^{min} \leq p_i \leq p_i^{max}\} \ . \tag{20}$$

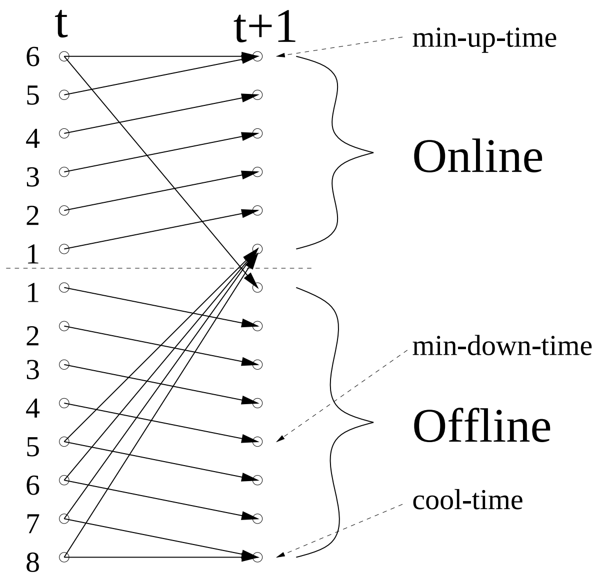

Figure 3. Transition graph for 2 time periods

for on-line states s. The arc weights $\widehat{S_i}(s,\tilde{s})$ describe start-up costs for the thermal unit. They are independent of $\boldsymbol{\lambda}$, and are non-zero only for arcs leading from off-line states to on-line states. The cost-to-go functions are given by

$$\gamma_i^t(s) = \alpha_i^t(s) + \mathbb{E}\left(\min_{\tilde{s}}\{\widehat{S_i}(s,\tilde{s}) + \gamma_i^{t+1}(\tilde{s})\}|\mathcal{A}_t\right), \tag{21}$$

where $\mathbb{E}(\cdot|\mathcal{A}_t)$ denotes the conditional expectation w.r.t. the σ-field $\mathcal{A}_t$. Now, the dynamic programming algorithm works as follows. First the cost-to-go functions are computed for all states s via the backward recursion (21). Then the optimal decisions are obtained by a forward computation. Since each node of the scenario tree is considered only twice, the algorithm is reasonably fast. For one thermal unit, one load scenario, and one week with an hourly discretization the algorithm needs just 40 milliseconds running time on an HP-workstation.

3.3. Descent Algorithm for Stochastic Storage Problems

The hydro subproblem (15) for unit j is solved by a descent method that generates a finite sequence of feasible decisions $(\mathbf{s}_j, \mathbf{w}_j)$ with decreasing objective values

$$\mathbb{E}\sum_{t=1}^{T} \boldsymbol{\lambda}_1^t(\mathbf{w}_j^t - \mathbf{s}_j^t)$$

and terminates with an optimal solution. The method begins by finding a hydro decision that is feasible, i.e., it satisfies (16) and (17).

For the description of a descent step we consider for simplicity the case $\eta_j = 1$ and set $\mathbf{x}_j = \mathbf{w}_j - \mathbf{s}_j$. Since the descent algorithm employs the underlying scenario tree structure, we need some additional notation. By $\mathcal{N}$ we denote the set of nodes of the tree, by $n = 1$ its root node, by path(n) the path from the root to node n, by $\mathcal{N}_+(n)$ the set of successors to node n, and by π_n the probability of node n. Furthermore, $\mathcal{N}_L$ denotes the set of leaves, i.e., the set of all nodes n with $\mathcal{N}_+(n) = \emptyset$. Denoting by x_j^n and λ_1^n the value of $\mathbf{x}_j$, and $\boldsymbol{\lambda}_1$ at node n, the objective function takes the form

$$\sum_{n \in \mathcal{N}} \pi_n \lambda_1^n x_j^n \tag{22}$$

The next feasible iterate $\tilde{\mathbf{x}}_j$ is chosen such that the objective (22) decreases and that $\tilde{x}_j^n = x_j^n$ holds for each node $n \in \mathcal{N} \setminus (\{n_G\} \cup G_L)$, i.e.,

$$\sum_{n \in \mathcal{N}} \pi_n \lambda_1^n (\tilde{x}_j^n - x_j^n) = \sum_{n \in \{n_G\} \cup G_L} \pi_n \lambda_1^n (\tilde{x}_j^n - x_j^n) < 0\,, \tag{23}$$

where n_G and G_L denote the root node and the set of leaves of a subtree with a set G of nodes contained in $\mathcal{N}$, i.e., it holds that $\{n_G\} \cup G_L \subseteq G$, $n_G \in \mathrm{path}(n)$ for each $n \in G$, $\mathcal{N}_+(n) \cap G = \emptyset$ for each $n \in G_L$ and $\mathcal{N}_+(n) \subseteq G$ for each $n \in G \setminus G_L$. A subtree having these properties is called *d-subtree.* It is shown in [27] that for each nonoptimal feasible hydro decision $\mathbf{x}_j$, a d-subtree and a feasible decision $\tilde{\mathbf{x}}_j$ exist such that (23) is satisfied. Moreover, the conditions on a node n to form a root node of a d-subtree are as follows:

$$\begin{array}{ll} \text{Case of increasing the level } l_j^n: & \min\{x_j^{max} - x_j^n, d_n^{up}\}\{\lambda_1^n \pi_n + r_n^{up}\} \le 0\,, \\ \text{Case of decreasing the level } l_j^n: & \min\{x_j^n - x_j^{min}, d_n^{down}\}\{\lambda_1^n \pi_n + r_n^{down}\} \le 0\,, \end{array}$$

where l_j^n denotes the value of $\mathbf{l}_j$ at node n, $x_j^{max} = w_j^{max}$, $x_j^{min} = -s_j^{max}$ and d_n^{up}, d_n^{down}, r_n^{up} and r_n^{down} are for each $n \in \mathcal{N}$ defined by

$$d_n^{up} = \begin{cases} x_j^n - x_j^{min} & \text{if } b_n^{up} = 1 \\ \min\{l_j^{max} - l_j^n, \min\limits_{n_+ \in \mathcal{N}_+(n)} d_{n_+}^{up}\} & \text{if } b_n^{up} = 0 \end{cases}$$

$$d_n^{down} = \begin{cases} x_j^{max} - x_j^n & \text{if } b_n^{down} = 1 \\ \min\{l_j^n, \min\limits_{n_+ \in \mathcal{N}_+(n)} d_{n_+}^{down}\} & \text{if } b_n^{down} = 0 \end{cases}$$

$$r_n^{up} = \begin{cases} \lambda_1^n \pi_n & \text{if } b_n^{up} = 1 \\ \sum\limits_{n_+ \in \mathcal{N}_+(n)} r_{n_+}^{up} & \text{if } b_n^{up} = 0 \end{cases}$$

$$r_n^{down} = \begin{cases} \lambda_1^n \pi_n & \text{if } b_n^{down} = 1 \\ \sum\limits_{n_+ \in \mathcal{N}_+(n)} r_{n_+}^{down} & \text{if } b_n^{down} = 0. \end{cases}$$

Here, b_n^{up} and b_n^{down} are binary decisions at node $n \in \mathcal{N}$. All superscripts *up*/*down* refer to cases of an increased/decreased level l_j^n for $n \in G \setminus G_L$. In case of an increased level, the correspondence of binary decisions to the tree G is determined by $n \in G_L \Leftrightarrow b_n^{up} = 1$ and $n \in G \setminus G_L \Leftrightarrow b_n^{up} = 0$. The decision to reduce the storage is denoted by $b_n^{up} = 1$, while $b_n^{up} = 0$ refers to the decision to keep the additional amount. Similarly, the notations $b_n^{down} = 1$ and $b_n^{down} = 0$ are used. Now, the descent algorithm EXCHA works as follows:

Step 1: Input and initialization;

Step 2: Determine a feasible point;

Step 3: Compute d_n^{up}, d_n^{down}, r_n^{up}, and r_n^{down} at all nodes $n \in \mathcal{N}$;

Step 4: Find the root node of the d-subtree having steepest descent; else **STOP**;

Step 5: Update x_j^n and l_j^n at all nodes $n \in \mathcal{N}$;

Step 6: Goto Step 3.

The extension of the descent algorithm to the case $\eta_j \in (0, 1)$ and implementation issues are given in [27]. The algorithm EXCHA was implemented and tested for

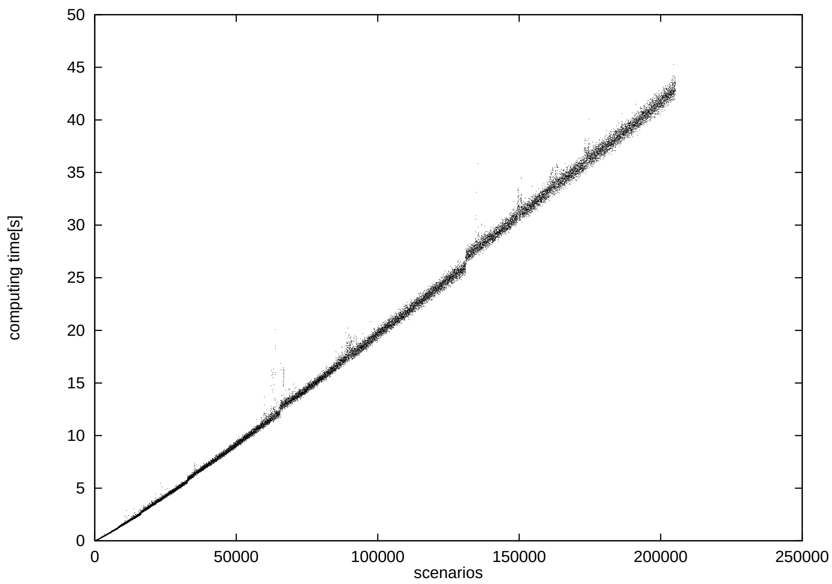

Figure 4. Computing times [s] of *EXCHA*

solving stochastic hydro storage subproblems (15). Fig. 4 shows the computing times (in seconds) of EXCHA on an HP-workstation for subproblems with $T \leq 18$ and binary trees branching at all time periods with numbers of scenarios ranging up to 200.000. The efficiency of the algorithm is due to the fact that in each step only a few elementary computations are needed and some variable attains its upper or lower bound.

3.4. Lagrangian Heuristics

When the bundle method delivers a solution $\boldsymbol{\lambda}^*$ of the dual stochastic program (12), the optimal value $D(\boldsymbol{\lambda}^*)$ provides a lower bound for the optimal cost of the primal model. In general, however, the "dual optimal" scheduling decisions $(\mathbf{u}(\boldsymbol{\lambda}^*), \mathbf{p}(\boldsymbol{\lambda}^*), \mathbf{s}(\boldsymbol{\lambda}^*), \mathbf{w}(\boldsymbol{\lambda}^*))$ violate the load and reserve constraints (4) and (5). In the following, we describe a Lagrangian heuristics that determines a nearly optimal primal first-stage decision starting from the optimal multiplier $\boldsymbol{\lambda}^*$. In a *first* step, the mean value functions of the (discrete-time) stochastic processes $\mathbf{d}$ (load), $\mathbf{r}$ (reserve), $\mathbf{l}$ (storage levels) and $\boldsymbol{\lambda}^*$ are computed. Clearly, they coincide with their realizations (scenarios) during all time periods belonging to the first stage. Next, generation and pumping decisions, s_j and w_j, are determined from the constraints (2), where $\mathbf{l}_j$ is replaced by its expectation $\mathbb{E}[\mathbf{l}_j]$. Furthermore, binary decisions u_i are computed by dynamic programming as solutions of the thermal subproblems (14), where the stochastic multiplier $\boldsymbol{\lambda}^*$ is replaced by its expectation $\lambda^* = \mathbb{E}[\boldsymbol{\lambda}^*]$. For one of the test runs explained in Section 4, Fig. 5 shows the results after the first step of the heuristics: the mean load and reserve curves $\mathbb{E}[\mathbf{d}^t]$ and $\mathbb{E}[\mathbf{r}^t]$, the hydro generation and pumping curves $\sum_{j=1}^J s_j^t$ and $\sum_{j=1}^J w_j^t$, and the reduced mean load curve $\mathbb{E}[\mathbf{d}^t] - \sum_{j=1}^J (s_j^t - w_j^t)$ for $t = 1, \ldots, T$. Furthermore, it shows that the reserve constraint (25) is violated

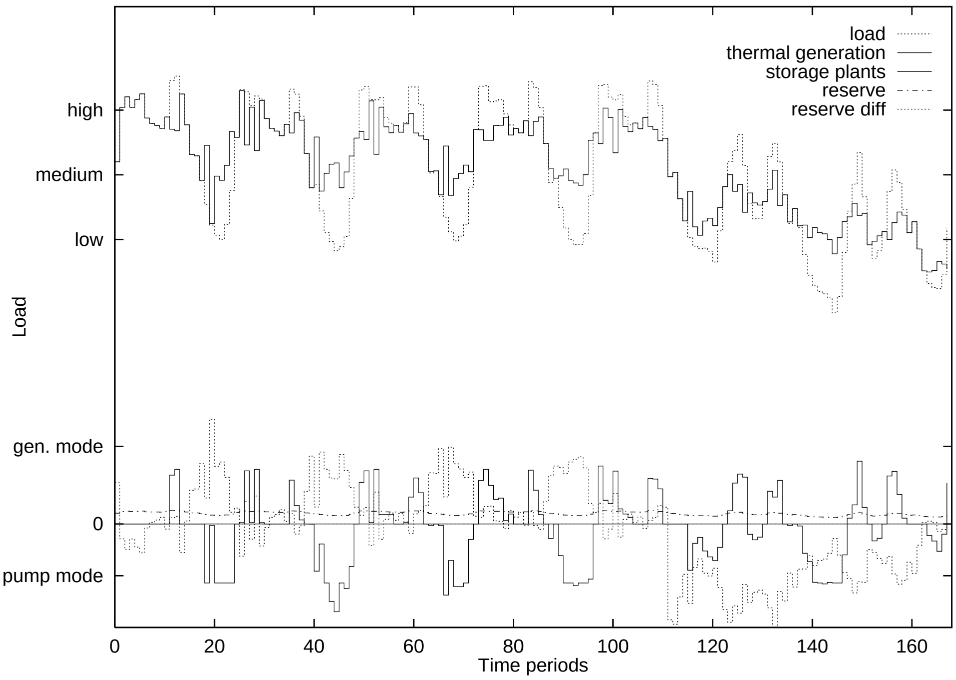

Figure 5. Schedules after averaging

e.g. during $1 \le t \le 12$ and $110 \le t \le 168$.
In order to find scheduling decisions (u, p, s, w) that are feasible for the reserve constraint (25), the schedules of the hydro and the thermal units,respectively, are modified during the next two steps. The *second* step consists in applying a *water rescheduling* procedure, which is taken from [9]. Its idea is to reduce the value

$$\mathbb{E}[\mathbf{d}^t] + \mathbb{E}[\mathbf{r}^t] + \sum_{j=1}^{J}(w_j^t - s_j^t) \tag{24}$$

by modifying the schedule of the hydro units if the (modified) reserve constraint

$$\sum_{i=1}^{I} u_i^t p_i^{max} \ge \mathbb{E}[\mathbf{d}^t] + \mathbb{E}[\mathbf{r}^t] + \sum_{j=1}^{J}(w_j^t - s_j^t) \tag{25}$$

is violated at time t and the value (24) is largest in a certain set of neighbouring time periods. In the *third* step, the hydro schedules are kept fixed and binary variables u_i^t satisfying the reserve constraint (25) are determined by the thermal heuristics described in [40]. Its main idea consists in determining the time t, where the constraint (25) is most violated, and in increasing $\boldsymbol{\lambda}_2^{*t}$ as much as necessary to switch on (by dynamic programming) just as many thermal units as needed to satisfy (25) at t. This is repeated until the reserve constraint (25) is satisfied in all time periods. Since this technique does not distinguish between identical units that appear quite often in real-life power systems, the start-up costs of such units are slightly modified. In our numerical experiments this modification led to improved results (cf. Section 4).

3.5. Economic Dispatch

The Lagrangian heuristics ends with a binary schedule u_i^t for the thermal units such that a feasible schedule (u, p, s, w) exists for the primal model in Section 2 when replacing the stochastic load $\mathbf{d}$ and reserve $\mathbf{r}$ by their expected values. In a final step, a cost-optimal schedule (p, s, w) is determined for fixed u by solving the corresponding primal model (with fixed start-up costs). The aim of this section is to develop an algorithmic approach for solving this economic dispatch problem. The approach also applies to multi-stage stochastic power scheduling models with fixed stochastic binary decisions $\mathbf{u}$. Since this may be of independent interest, we consider the model:

$$\min_{(\mathbf{p},\mathbf{s},\mathbf{w})} \left\{ \mathbb{E} \sum_{i=1}^{I}\sum_{t=1}^{T} C_i(\mathbf{p}_i^t, \mathbf{u}_i^t) \;:\; (\mathbf{p}, \mathbf{s}, \mathbf{w}) \in \mathop{\times}_{t=1}^{T} L^\infty(\Omega, \mathcal{A}_t, \mathbb{P}; \mathbb{R}^{I+2J}) \right., \tag{26}$$

$$\mathbf{u}_i^t p_i^{min} \le \mathbf{p}_i^t \le \mathbf{u}_i^t p_i^{max},\ t = 1, \ldots, T,\ i = 1, \ldots, I, \tag{27}$$

$$0 \le \mathbf{s}_j^t \le s_j^{max},\ 0 \le \mathbf{w}_j^t \le w_j^{max},\ 0 \le \mathbf{l}_j^t \le l_j^{max}, \tag{28}$$

$$\mathbf{l}_j^t = \mathbf{l}_j^{t-1} - \mathbf{s}_j^t + \eta_j \mathbf{w}_j^t,\ t = 1,\dots,T,\ \mathbf{l}_j^0 = l_j^{in},\ \mathbf{l}_j^T = l_j^{end},\ j = 1,\dots,J, \tag{29}$$

$$\sum_{i=1}^{I} \mathbf{p}_i^t + \sum_{j=1}^{J} (\mathbf{s}_j^t - \mathbf{w}_j^t) \geq \mathbf{d}^t\ ,\ t = 1,\dots,T, \tag{30}$$

$$\left. \sum_{i=1}^{I} (\mathbf{u}_i^t p_i^{max} - \mathbf{p}_i^t) \geq \mathbf{r}^t\ ,\ t = 1,\dots,T \right\} . \tag{31}$$

The structure of the stochastic program (26)-(31) is partly similar to (15) excepting the thermal units. This motivates the idea to apply the same technique as in Section 3.3. Thermal and hydro units are coupled by the constraints (30). Moving the sum $\sum_{j=1}^{J} (\mathbf{s}_j^t - \mathbf{w}_j^t)$ to the right-hand side in (30) and taking the right-hand side as a parameter, the optimization problem (26), (27), (30) decomposes into parametric programs for each time period t and scenario ω. Denoting the parameter by θ, the parametric programs and their optimal value functions $\phi_{t,\omega}(\cdot)$ have the form:

$$\phi_{t,\omega}(\theta) = \min_{p^t} \left\{ \sum_{i=1}^{I} C_i(p_i^t, \mathbf{u}_i^t(\omega))\ :\ \mathbf{u}_i^t(\omega) p_i^{min} \leq p_i^t \leq \mathbf{u}_i^t(\omega) p_i^{max}\ , i = 1,\dots,I, \right.$$
$$\left. \mathbf{d}^t(\omega) - \theta \leq \sum_{i=1}^{I} p_i^t \leq \sum_{i=1}^{I} \mathbf{u}_i^t(\omega) p_i^{max} - \mathbf{r}^t(\omega) \right\} .$$

Such optimal value functions may be evaluated by efficient algorithms (see e.g. [4,5,22] for the case of (piecewise) linear and quadratic costs). Now, the economic dispatch problem (26) - (31) can be reformulated as

$$\min_{(\mathbf{s},\mathbf{w})} \left\{ \mathbb{E} \sum_{t=1}^{T} \phi_{t,\cdot} (\sum_{j=1}^{J} (\mathbf{s}_j^t - \mathbf{w}_j^t))\ :\ (28)\ ,\ (29) \right\} \tag{32}$$

This reformulation allows to study how the objective function varies when altering the operation of the hydro units. If the functions $\phi_{t,\omega}$ were differentiable, the linearization of the model (32) takes the form

$$\sum_{j=1}^{J} \min_{(\mathbf{s}_j,\mathbf{w}_j)} \mathbb{E} \sum_{t=1}^{T} \frac{\mathrm{d}\phi_{t,\cdot}}{\mathrm{d}\theta} (\sum_{j=1}^{J} (\mathbf{s}_j^t - \mathbf{w}_j^t))(\mathbf{s}_j^t - \mathbf{w}_j^t)\ , \tag{33}$$

which is very similar to (15). Hence, successive linearizations combined with the descent technique described in Section 3.3 could be used to solve (32). This suggests replacing each piecewise linear function $\phi_{t,\omega}$ by a differentiable function that is obtained from $\phi_{t,\omega}$ by smoothing its kinks with quadratic functions on small intervals that are reduced progressively. This descent method was implemented, tested and compared with CPLEX 4.0. Test runs of our code ECDISP

CPLEX-function	Pricing strategy primal/dual					
	-1	0	1	2	3	4
Simplex/primal	1232.4	1188.4	1918.1	2664.1	2440.7	1696.9
Simplex/dual		1086.1	946.2	1103.4	1466.5	1083.8
baropt	94.78					
hybbaropt/primal	114.7	114.3	114.3	486.5	114.4	114.3
hybbaropt/dual		115.0	114.6	693.0	1424.8	114.8
hybnetopt/primal	957.6	910.3	1298.0	2252.8	1960.9	1162.6
hybnetopt/dual		1393.8	1253.7	1412.0	1833.9	1392.3

Table 2
Computing times [s] for different CPLEX-functions and options

were performed for the VEAG system with 25 thermal units and 7 pumped hydro storage plants. Table 2 contains results for a test example with one load scenario and 192 time periods, which is equivalent to an LP with 14200 columns, 17856 rows, and 46256 non-zeros. The table shows computing times of CPLEX 4.0 on a SPARCstation IPX (4/50) with 64 MB main memory and 40 MHz, which have to be compared with the ECDISP computing time of 50.95 seconds. Since the barrier method performs significantly better than the simplex method, and even

scen's	nodes	columns	rows	non-zeros	ECDISP[s]	CPLEX[s]	adv'
3	336	24840	31248	80944	18.69	97.61	5.22
5	462	34148	42966	111294	29.48	162.47	5.51
7	588	43456	54684	141644	47.93	206.00	4.30
9	687	50766	63891	165487	43.09	305.43	7.09
11	792	58520	73656	190776	67.17	500.30	7.45
13	930	68716	86490	224018	86.73	461.54	5.32
15	1035	76470	96255	249307	98.04	569.18	5.81
17	1036	76528	96348	249532	117.42	620.65	5.29
19	1120	82728	104160	269760	91.63	1720.33	18.7
21	1232	91000	114576	296736	131.94	243.27	1.84
22	1260	93064	117180	303476	128.18	794.93	6.20

Table 3
Comparison of ECDISP with CPLEX

better than the network simplex method, further comparisons were made with the barrier method only. Table 3 contains results for test problems with $T = 192$ and up to 22 scenarios. CPLEX 4.0 ran out of memory for problems with a higher number of scenarios. The advantage of using ECDISP ranges from 1.8 up to 18.7, and in average ECDISP is 5-6 times faster. Figs. 6 and 7 show that the number of steps and the computing times of ECDISP grow almost linearly with respect to the number of scenarios.

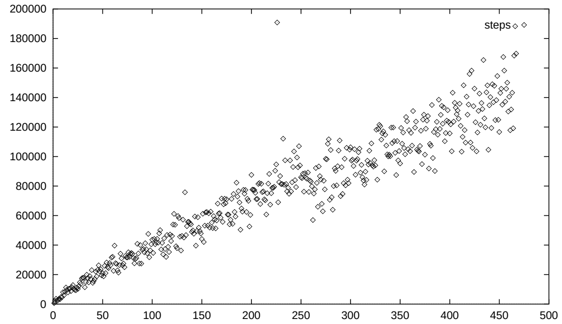

Figure: 6: Number of steps versus number of scenarios

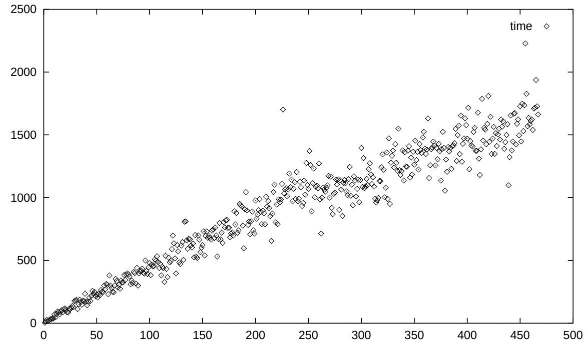

Figure: 7: Computing times [s] versus number of scenarios

4. Numerical Results

The stochastic Lagrangian relaxation algorithm was implemented in C++ except for the proximal bundle method, for which the FORTRAN-package NOA 3.0 ([20]) was used as a callable library. For testing the implementation, a test bunch of load scenario trees was generated as follows. Starting from a reference load scenario obtained from real-life data, $S - 1$ random branching points were selected successively to produce a scenario tree with S identical scenarios. Then a discretized Brownian motion was added to each node of the tree. Finally, randomly selected probabilities were assigned to each scenario. The random construction of the trees leads to different trees and different numbers of nodes (even for fixed S) for each test run. Test runs were performed for the hydro-thermal power generation system of VEAG comprising 25 thermal units and 7 pumped storage plants on an HP 9000 (780/J280) Compute-Server with 180 MHz frequency and 768 MByte main memory under HP-UX 10.20. For test runs with

Nr.	modified cost functions				original cost functions			
	opt.tol: 10^{-3}		opt.tol: 10^{-4}		opt.tol: 10^{-3}		opt.tol: 10^{-4}	
	gap/%	time/s	gap/%	time/s	gap/%	time/s	gap/%	time/s
1	0.18	34.08	0.10	89.84	0.67	31.32	0.56	86.24
2	0.25	47.82	0.12	109.92	0.60	42.67	0.61	100.07
3	0.43	44.81	0.26	111.75	0.44	35.10	0.25	102.61
4	0.34	53.86	0.14	119.84	0.94	47.40	0.96	115.55
5	0.20	78.42	0.11	157.31	0.98	73.76	0.93	151.64
6	0.39	39.52	0.11	88.35	0.66	37.42	0.54	79.88

Table 4
Influence of modified costs and of NOA opt.tol on the gap and running times

10 scenarios Table 4 shows the influence of modifying the start-up costs and of changing the optimality tolerance of the proximal bundle method on the gap and computing times. It shows that a (slight) modification of start-up costs of former identical units leads to smaller gaps. Here, the gap refers to the relative difference

$$\frac{1}{D(\boldsymbol{\lambda}^*)}\left(\sum_{t=1}^{T}\sum_{i=1}^{I}\left[C_i(p_i^t,u_i^t)+S_i^t(u_i)\right]-D(\boldsymbol{\lambda}^*)\right)$$

of the cost of the scheduling decision $(u,\, p,\, s,\, w)$ and the optimal value of the dual problem. Moreover, improving optimality tolerances leads to smaller gaps paid by increased computing times. Fig. 8 provides the final output of the algorithm and contains, in particular, the nearly optimal first-stage solution for the total thermal and hydro generation (i.e. for the periods $t = 1, \ldots, 24$). Table 5 shows how the computing time grows with increasing numbers of scenarios and nodes. Since the complexity of the model is higher compared to the stochastic programs in Sections 3.3 and 3.5, the variance of the computing time is greater than the variances expressed in Figs. 4 and 7. The reason is that the iteration numbers in the bundle method, in the method for searching a reserve feasible solution and in the economic dispatch solver depend on the input data in a very involved way. Another observation is that the gap seems to be (almost) independent of the number of scenarios.

Scenarios	Nodes	time[s]/gap[%]	Nodes	time[s]/gap[%]
10	781	31.2 / 0.274	1043	52.93 / 0.138
10	1232	50.36 / 0.201	975	54.21 / 0.723
20	1982	89.13 / 0.149	1627	93.62 / 0.101
20	1651	67.94 / 0.367	1805	84.73 / 0.066
30	2643	139.71 / 0.528	2643	138.61 / 0.528
30	2548	147.51 / 0.849	2515	162.14 / 0.175
50	4530	475.29 / 0.175	4060	274.43 / 0.096
50	4041	312.86 / 0.099	4457	288.03 / 0.430
100	9230	1183.25 / 0.108	9224	1072.18 / 0.131
100	7727	929.68 / 0.087	8867	1234.12 / 0.304

Table 5
Computing times and gaps (NOA 3.0: opt.tol= 10^{-3}, NGRAD= 50)

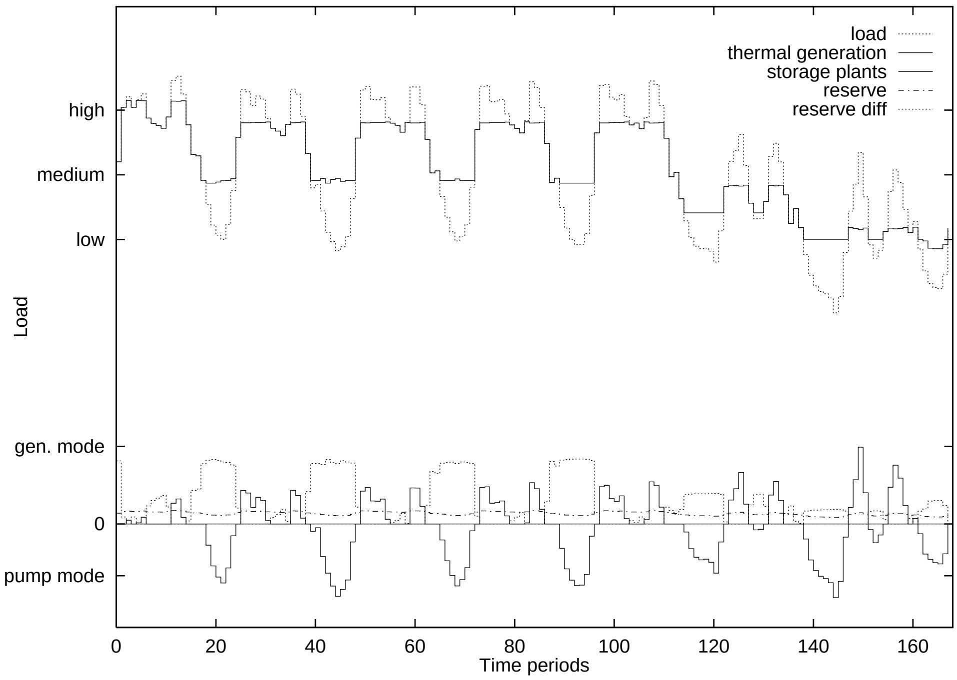

Figure 8. Approximate solution

5. Conclusions

We have elaborated a mixed-integer multi-stage stochastic programming model for power scheduling in a hydro-thermal generation system under uncertainty on the electrical load. Due to the huge size of the model, an application of state-of-the-art mixed-integer LP solvers is prevented. Therefore, we have developed a novel approach based on stochastic Lagrangian relaxation of coupling constraints. It consists of proximal bundle iterations for solving a stochastic dual followed by a Lagrangian heuristics to determine a nearly optimal primal first-stage solution. The stochastic dual decomposes into stochastic thermal and hydro subproblems, which are solved by specific fast algorithms. Our computational experience indicates that the stochastic Lagrangian relaxation algorithm is able to produce good approximate first-stage solutions for medium-size realistic power systems and 20 (100) load scenarios within less than 2 (20) minutes on a modern HP-workstation. It also indicates that the algorithm bears potential for solving more complex real-life power scheduling models under uncertainty in reasonable time.

Acknowledgements

The authors are grateful to K.C. Kiwiel (Polish Academy of Sciences, Warsaw) for the permission to use the NOA 3.0 package, to Darinka Dentcheva (Humboldt-University Berlin) for useful discussions on the Lagrangian heuristics described in Section 3.4 and to G. Schwarzbach and J. Thomas (VEAG Vereinigte Energiewerke AG) for providing us with the data used in our test runs. We extend our gratitude to two anonymous referees for their suggestions which led to an improved presentation of the material. This research was supported by the Schwerpunktprogramm *Echtzeit-Optimierung großer Systeme* of the Deutsche Forschungsgemeinschaft and by the Graduiertenkolleg *Geometrie und Nichtlineare Analysis* at the Humboldt-University Berlin.

References

[1] D.P. Bertsekas, G.S. Lauer, N.R. Sandell Jr. and T.A. Posbergh, Optimal short-term scheduling of large-scale power systems, IEEE Transactions on Automatic Control AC-28 (1983),1-11.

[2] J.R. Birge, Stochastic programming computations and applications, INFORMS Journal on Computing 9 (1997), 111-133.

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

[4] P.P.J. van den Bosch, Optimal static dispatch with linear, quadratic and nonlinear functions of the fuel costs, IEEE Transactions on Power Apparatus and Systems PAS-104 (1985), 3402-3408.

[5] P.P.J. van den Bosch and F.A. Lootsma, Scheduling of power generation via large-scale nonlinear optimization, Journal of Optimization and Applications 55 (1987), 1684-1690.

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

[7] C.C. Carøe and R. Schultz, A two-stage stochastic program for unit commitment under uncertainty in a hydro-thermal power system, DFG-Schwerpunktprogramm *Echtzeit-Optimierung großer Systeme*, Preprint 98-13, 1998.

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

[9] D. Dentcheva, R. Gollmer, A. Möller, W. Römisch and R. Schultz, Solving the unit commitment problem in power generation by primal and dual methods, in: *Progress in Industrial Mathematics at ECMI 96*, eds. M. Brøns, M.P. Bendsøe and M.P. Sørensen, Teubner, Stuttgart 1997, 332-339.

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

[11] S. Feltenmark and K.C. Kiwiel, Dual applications of proximal bundle methods, including Lagrangian relaxation of nonconvex problems, SIAM Journal on Optimization (to appear).

[12] S.-E. Fleten, S.W. Wallace and W.T. Ziemba, Portfolio management in a deregulated hydropower based electricity market, in: *Hydropower' 97*,eds. E. Broch, D.K. Lysne, N. Flatabø and E. Helland-Hansen, Balkema, Rotterdam 1997.

[13] A. Gjelsvik, T.A. Røtting and J. Røynstrand, Long-term scheduling of hydro-thermal power systems, in: *Hydropower '92*, eds. E. Broch and D.K. Lysne, A.A. Balkema, Rotterdam 1992, 539-546.

[14] R. Gollmer, A. Möller, W. Römisch, R. Schultz, G. Schwarzbach and J. Thomas, Optimale Blockauswahl bei der Kraftwerkseinsatzplanung der VEAG, in: *Optimierung in der Energieversorgung II*, VDI-Berichte 1352, Düsseldorf 1997, 71-85.

[15] N. Gröwe, W. Römisch, R. Schultz, A simple recourse model for power dispatch under uncertain demand, Annals of Operations Research 59 (1995), 135-164.

[16] J.L. Higle and S. Sen, *Stochastic Decomposition - A Statistical Method for Large Scale Stochastic Linear Programming*, Kluwer, Dordrecht, 1996.

[17] G. Infanger, *Planning Under Uncertainty - Solving Large-Scale Stochastic Linear Programs*, Boyd and Fraser, 1994.

[18] J. Jacobs, G. Freeman, J. Grygier, D. Morton, G. Schultz, K. Staschus and J. Stedinger, SOCRATES: A system for scheduling hydroelectric generation under uncertainty, Annals of Operations Research 59 (1995), 99-133.

[19] K.C. Kiwiel, Proximity control in bundle methods for convex nondifferentiable optimization, Mathematical Programming 46 (1990), 105-122.

[20] K.C. Kiwiel, User's Guide for NOA 2.0/3.0: A Fortran package for convex nondifferentiable optimization, Polish Academy of Sciences, Systems Research Institute, Warsaw (Poland), 1993/94.

[21] K.C. Kiwiel, Approximations in proximal bundle methods and decomposition of convex programs, Journal of Optimization Theory and Applications 84 (1995), 529-548.

[22] P. Kleinmann and R. Schultz, A simple procedure for optimal load dispatch using parametric programming, ZOR - Methods and Models of Operations Research 34 (1990), 219-229.

[23] C. Lemarechal, Lagrangian decomposition and nonsmooth optimization, in: *Nonsmooth Optimization Methods and Applications*, ed. F. Gianessi, Gordon and Breach, 1992, 201-216.

[24] C. Lemarechal and A. Renaud, Dual equivalent convex and nonconvex problems, Research Report, INRIA, Rocquencourt, 1996.

[25] A. Løkketangen and D.L. Woodruff, Progressive hedging and tabu search applied to mixed integer (0,1) multi-stage stochastic programming, Journal of Heuristics 2 (1996), 111-128.

[26] V.I. Norkin, G. Pflug and A. Ruszczyński, A branch and bound method for stochastic global optimization, Mathematical Programming 83 (1998), 425-450.

[27] M.P. Nowak, A fast descent method for the hydro storage subproblem in power generation, International Institute for Applied Systems Analysis (Laxenburg, Austria), Working Paper WP-96-109, 1996.

[28] M.P. Nowak and W. Römisch, Optimal power dispatch via multistage stochastic programming, in: *Progress in Industrial Mathematics at ECMI 96*, eds. M. Brøns, M.P. Bendsøe and M.P. Sørensen, Teubner, Stuttgart, 1997, 324-331.

[29] M.V.F. Pereira and L.M.V.G. Pinto, Multi-stage stochastic optimization applied to energy planning, Mathematical Programming 52 (1991), 359-375.

[30] A. Renaud, Optimizing short-term operation of power generation: a new stochastic model, Lecture presented at the VIII International Conference on Stochastic Programming, Vancouver (Canada), August 1998.

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

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

[33] W. Römisch and R. Schultz, Decomposition of a multi-stage stochastic program for power dispatch, ZAMM - Zeitschrift für Angewandte Mathematik und Mechanik 76 (1996) Suppl. 3, 29-32.

[34] A. Ruszczyński, Decomposition methods in stochastic programming, Mathematical Programming, Series B, 79 (1997), 333-353.

[35] G.B. Sheble and G.N. Fahd, Unit commitment literature synopsis, IEEE Transactions on Power Systems 9 (1994), 128-135.

[36] S. Takriti, J.R. Birge and E. Long, A stochastic model for the unit commitment problem, IEEE Transactions on Power Systems 11 (1996), 1497-1508.

[37] S. Takriti, B. Krasenbrink and L.S.-Y. Wu, Incorporating fuel constraints and electricity spot prices into the stochastic unit commitment problem, IBM Research Report RC 21066, Yorktown Heights, New York 1997 and Operations Research (to appear).

[38] R. J.-B. Wets, Stochastic Programming, in: *Handbooks in Operations Research and Management Science*, Vol. 1, Optimization, eds. G.L. Nemhauser, A.H.G. Rinnooy Kan, M.J. Todd, North-Holland, Amsterdam, 1989, 573-629.

[39] A.J. Wood and B.F. Wollenberg, *Power Generation, Operation, and Control* (Second Edition), Wiley, New York, 1996.

[40] F. Zhuang and F.D. Galiana, Towards a more rigorous and practical unit commitment by Lagrangian relaxation, IEEE Transactions on Power Systems 3 (1988), 763-773.