RAIRO-Oper. Res. 51 (2017) 391–416
DOI: 10.1051/ro/2016031

RAIRO Operations Research
www.rairo-ro.org

UNIT COMMITMENT UNDER UNCERTAINTY IN AC TRANSMISSION SYSTEMS *VIA* RISK AVERSE SEMIDEFINITE STOCHASTIC PROGRAMS

Rüdiger Schultz[1] and Tobias Wollenberg[1]

Abstract. This paper addresses unit commitment under uncertainty of load and power infeed from renewables in alternating current (AC) power systems. Beside traditional unit-commitment constraints, the physics of power flow are included. To gain globally optimal solutions a recent semidefinite programming approach is used, which leads us to risk averse two-stage stochastic mixed integer semidefinite programs for which a decomposition algorithm is presented.

Mathematics Subject Classification. 90C15, 90C35, 90C22.

Received April 18, 2015. Accepted April 23, 2016.

1. Introduction

In recent years, research in stochastic programming has moved into various new directions. This concerns both theory and applications. Without aiming at completeness one could mention on the theory side: risk aversion with risk measures or stochastic orders, stochastic programs in mixed-integer, semidefinite, bilevel, or PDE constrained optimization as well as scenario tree construction and reduction. Fields of applications include finance, logistics, and energy optimization in the broadest sense.

The present paper shall contribute to this development by bringing together stochastic programming, semidefinite optimization, risk aversion, and optimal power flow in AC networks. In doing so, the paper draws on seminal work in risk neutral semidefinite stochastic programming [3, 31], on recent progress in power flow optimization [24], and on risk aversion by forming objective functions involving risk measures [39, 44].

To capture risk aversion in a minimization context we resort to an intuitive measure which is the probability of a random quantity to exceed a preassigned critical level. This measure, called excess or exceedance probability has been analysed in two-stage stochastic linear mixed-integer stochastic programming in [44]. In reliability analysis it has a role in various fields of engineering of which a more recent one is seismic risk analysis [18]. Here the total cost of damage and retrofit caused by an earthquake is the random quantity of interest, and risk is measured by the probability of this quantity to exceed a threshold value.

Our motivation to investigate risk aversion from the viewpoint of excess probabilities comes from the simultaneous treatment of unit commitment and AC load flow under uncertainty of power demand and infeed from

Keywords. Stochastic programming, semidefinite programming, AC power flow.

[1] Department of Mathematics, University of Duisburg-Essen, Thea-Leymann-Straße 9, 45127 Essen, Germany.
ruediger.schultz@uni-due.de; tobias.wollenberg@uni-due.de

renewables in power management. The "geographic split of the two", meaning that the locations where electricity is produced from renewables and the locations where electricity is consumed are distant apart, has given transportation *via* the electric grid increased importance. Therefore it is reasonable to expect that previous unit commitment models neglecting the grid at all or using DC approximations of the AC load flow are too coarse.

Another recent development, this time in power flow methodology, has spurred our interest in incorporating risk aversion into power flow optimization models. In [24] the authors formulate AC load flow by means of convex semidefinite constraints and some rank condition (semidefinite programming for optimal power flow problems was first presented in [4]). With a fixed commitment of generating units and for a fixed point in time, they solve the dual to the mentioned convex program. When heading for a primal solution, a good many times, their proposed solution approach has the ability to retrieve the relaxed rank-one condition, such that it enables the opportunity to solve (nonconvex) power flow problems to global optimality.

Motivated by the uncertain parameters typically developing in time, we study unit commitment over some time horizon. We extend existing unit commitment models by putting simultaneous consideration of AC load flow and stochastic uncertainty on top of the model. Together with the semidefinite programming approach in [24] this will lead to two-stage mixed-integer stochastic semidefinite programs, for which a decomposition algorithm will be presented.

Our work is organized as follows. First, in Section 2, our basic deterministic unit commitment problem is formulated. An efficient solution approach, based on a combination of the semidefinite reformulation [24] and a Benders decomposition approach [1], is presented in Section 3. Inclusion of uncertainties is considered in Section 4. This especially implies introduction of risk aversion *via* excess probabilities in two-stage stochastic semidefinite programs. We discuss particularities of these stochastic programs and present a decomposition algorithm for their solution. Finally, computational results and concluding remarks are given in Section 5.

2. Formulating the basic problem

We consider an AC power system that interconnects various power production units (such as coal fired blocks, gas turbines, pumped-storage units, and wind parks) to consumers. For some preassigned planning horizon, the challenge is to provide "optimal service" to the consumers in economically efficient, technologically feasible, and operationally reliable manner.

From the mathematical-optimization perspective, these three targets concern main branches of current research. Economic aspects, usually addressed under the key words of power dispatch and unit commitment, lead into large scale mixed-integer (linear) optimization. While here linearity often provides an acceptable compromise for model precision, this no longer holds for the technological aspects capturing generation and transmission of electricity subject to the physical laws and engineering constraints. Jointly, these features are addressed as optimal power flow. As an additional difficulty one faces the nonlinearity inevitably arising in its nonconvex fashion. Finally, the reliability issue, in the widest sense, leads into optimization under uncertainty with robust and stochastic optimization as major lines of development.

Given the breadth of topics with seminal contributions dating back for 50 years and more, *e.g.* the first model for optimal power flow due to [11], there is a vast literature on the above themes. Therefore, we here confine ourselves to refer to the recent very useful primer [14] and the excellent bibliographical review in [15] and [16]. Although all three papers mainly circle around different aspects of optimal power flow, coverage of the economic aspects and the uncertainty issue is substantial as well.

2.1. Basic traditional UC model

To begin with, we introduce principal characteristics of the unit commitment part of our full model. Drawing on [9] and [17] the presentation is fairly detailed, mainly to be self-contained, but also to introduce the quite complex notation needed subsequently.

Throughout, boldfaced symbols in mathematical formulas stand for variables, symbols in normal font for problem data.

Consider a power gird with the set of buses $\mathcal{N} := \{1, \dots, n\}$, the set of generators $\mathcal{G} \subseteq V$, and the set of flow lines $\mathcal{L} \subseteq \mathcal{N} \times \mathcal{N}$. Assume that for $(l, m) \in \mathcal{L}$, we also have $(m, l) \in \mathcal{L}$. The set of all generator buses $\mathcal{G}$ decomposes into coal fired blocks, attached gas turbines, and installed pumped-storage units, denoted by $i = 1, \dots, I$, $r = 1, \dots, R$, and $h = 1, \dots, H$, respectively. Wind power is modeled by positive infeed at wind farm buses, such that these units are not considered as controllable production devices. We will optimize over a time horizon which is discretized into finitely many hourly planning intervals $t = 1, \dots, T$. The boolean decision variables:

$$\mathbf{u}_i^t \in \{0, 1\}, \quad i = 1, \dots, I, \quad t = 1, \dots, T,$$

then indicate whether the coal fired block i is on- or off-line during time interval t. Analogously, there are the variables $\mathbf{u}_r^t \in \{0, 1\}, r = 1, \dots, R; t = 1, \dots, T$ for the gas turbines as well as the nonnegative continuous variables:

$$\mathbf{p}_i^t, \mathbf{q}_i^t,\ \mathbf{p}_r^t, \mathbf{q}_r^t,\ \mathbf{p}_h^t, \mathbf{q}_h^t, \mathbf{w}_h^t, \overline{\mathbf{w}}_h^t, \quad i = 1, \dots, I,\ r = 1, \dots, R,\ h = 1, \dots, H,\ t = 1, \dots, T,$$

representing the output levels, in both active and reactive power, for the coal fired thermal units, the gas turbines, the pumped-storage units in generation and in pumping modes. For each of the coal fired units and gas turbines we assume a quadratic cost function with given nonnegative coefficients accounting for the fuel cost in terms of active power generation, *i.e.* the fuel costs and thus the objective to be minimized is given by:

$$\sum\nolimits_{t=1}^{T} \left(\sum\nolimits_{k \in I \cup R} f_k(\mathbf{p}_k^t, \mathbf{u}_k^t, \mathbf{r}_k^t) \right), \tag{2.1}$$

with

$$f_k(\mathbf{p}_k^t, \mathbf{u}_k^t, \mathbf{r}_k^t) = c_{k2} \left(\mathbf{p}_k^t\right)^2 + c_{k1} \mathbf{p}_k^t + c_{k0} \mathbf{r}_k^t,$$

where $\mathbf{r}_k^t = \max\{\mathbf{u}_k^{t-1} - \mathbf{u}_k^t, 0\}$. Further, we denote by $P_i^{\min}$, $P_i^{\max}$, $Q_i^{\min}$, $Q_i^{\max}$, $P_r^{\max}$, $Q_r^{\max}$, $P_h^{\max}$, $Q_h^{\max}$, $W_h^{\max}$, $\overline{W}_h^{\max}$, the minimal and maximal outputs of the particular power production units. All outputs have to be within these bounds, where the natural lower bound for gas turbines, and pumped-storage units (in generation and pumping mode) is zero. This yields for all $t = 1, \dots, T$:

$$P_k^{\min} \cdot \mathbf{u}_k^t \le \mathbf{p}_k^t \le P_k^{\max} \cdot \mathbf{u}_k^t, \quad Q_k^{\min} \cdot \mathbf{u}_k^t \le \mathbf{q}_k^t \le Q_k^{\max} \cdot \mathbf{u}_k^t, \qquad \forall k \in I \cup R, \tag{2.2}$$

$$-W_h^{\min} \le \mathbf{p}_h^t - \mathbf{w}_h^t \le P_h^{\max}, \quad -\overline{W}_h^{\min} \le \mathbf{q}_h^t - \overline{\mathbf{w}}_h^t \le Q_h^{\max}, \qquad \forall h \in H. \tag{2.3}$$

Beside these production bounds, the coal fired blocks must adhere to minimum down times to avoid excessive thermal strains. These are expressed by the following inequalities:

$$\mathbf{u}_i^{t-1} - \mathbf{u}_i^t \le 1 - \mathbf{u}_i^l, \qquad \begin{array}{l} i = 1, \dots, I, \quad t = 2, \dots, T-1, \\ l = t+1, \dots, \min\{t + \tau_i - 1, T\}, \end{array} \tag{2.4}$$

where the τ_i represent the required down times. Furthermore, there are variables $\mathbf{l}_h^t$, $h = 1, \dots, H; t = 1, \dots, T$, specifying the fill (in active power) of the upper dam at pumped-storage unit h at the end of time interval t. At all times, the (nonnegative) fill levels must not exceed the maximum fills $l_h^{\max}$ and, together with generation and pumping, the fill has to meet the following balances:

$$\mathbf{l}_h^0 = l_h^{\text{in}},\ \mathbf{l}_h^T = l_h^{\text{end}},\ \mathbf{l}_h^t = \mathbf{l}_h^{t-1} - (\mathbf{p}_h^t - \eta_h \mathbf{w}_h^t)\,, \ \ h = 1, \dots, H,\ t = 1, \dots, T, \tag{2.5}$$

where l_h^{in}, l_h^{end} are the initial and final fills, respectively, and $0 \le \eta_h < 1$ indicates the pumping efficiency.

2.2. AC load flow extension

Turning attention to AC load flow, for every network bus $k \in \mathcal{N}$, we consider its apparent power[2] $\mathbf{s}_k^t = \mathbf{p}_k^t + j\mathbf{q}_k^t$ at time $t = 1, \dots, T$, where $\mathbf{p}_k^t$ denotes its active and $\mathbf{q}_k^t$ its reactive power, respectively. The apparent power is subject to Kirchhoff's first law, *i.e.* at any node in an electrical network, the sum of currents flowing into that node is equal to the sum of currents flowing out of it:

$$\mathbf{p}_k^t = \sum\nolimits_{l \in \mathcal{N}(k)} \mathbf{p}_{kl}^t, \qquad \forall k \in \mathcal{N}, \quad t = 1, \dots, T, \tag{2.6}$$

$$\mathbf{q}_k^t = \sum\nolimits_{l \in \mathcal{N}(k)} \mathbf{q}_{kl}^t, \qquad \forall k \in \mathcal{N}, \quad t = 1, \dots, T, \tag{2.7}$$

where $\mathbf{p}_{kl}^t$ and $\mathbf{q}_{kl}^t$ are the active and reactive power, respectively, transferred from k to the rest of the network through line $(k, l) \in \mathcal{L}$, and $\mathcal{N}(k)$ denotes the set of all buses directly connected to k. The apparent power $\mathbf{s}_k^t$ can also be written as the difference between complex infeed $\mathbf{s}_{G_k}^t := \mathbf{p}_{G_k}^t + j\mathbf{q}_{G_k}^t$ and complex load $\mathbf{s}_{D_k}^t := \mathbf{p}_{D_k}^t + j\mathbf{q}_{D_k}^t$, such that together with (2.6) and (2.7) we arrive at the following power balance equations:

$$\mathbf{p}_{G_k}^t - \sum\nolimits_{l \in \mathcal{N}(k)} \mathbf{p}_{kl}^t = p_{D_k}^t, \qquad \forall k \in \mathcal{G}, \qquad t = 1, \dots, T, \tag{2.8}$$

$$\mathbf{q}_{G_k}^t - \sum\nolimits_{l \in \mathcal{N}(k)} \mathbf{q}_{kl}^t = q_{D_k}^t, \qquad \forall k \in \mathcal{G}, \qquad t = 1, \dots, T, \tag{2.9}$$

$$- \sum\nolimits_{l \in \mathcal{N}(k)} \mathbf{p}_{kl}^t = p_{D_k}^t, \qquad \forall k \in \mathcal{N} \backslash \mathcal{G}, \quad t = 1, \dots, T, \tag{2.10}$$

$$- \sum\nolimits_{l \in \mathcal{N}(k)} \mathbf{q}_{kl}^t = q_{D_k}^t, \qquad \forall k \in \mathcal{N} \backslash \mathcal{G}, \quad t = 1, \dots, T, \tag{2.11}$$

where, the active and reactive electrical load $\{(p_D^t, q_D^t) : t = 1, \dots, T\}$ in terms of demand and infeed of renewables is given in advance and has to be covered (exactly).

To represent the energy flows, one possibility, for others see [14], is to select as variables the voltage angle $\boldsymbol{\theta}_k^t$ and the voltage magnitude $\mathbf{U}_k^t$ at every bus $k \in \mathcal{N}$. Then, there needs to be at least one slack bus with specified voltage magnitude and angle. It is used to balance apparent power, in such a way that it compensates system losses by emitting and absorbing active power and reactive power to and from the system, respectively. In selecting the slack bus, it is important to ensure that a powerful bus[3] is chosen, which can absorb all uncertainties arising from the system. Here, we pick bus $1 \in \mathcal{N}$ as slack bus and additionally demand $\theta_1 = 0$. Furthermore, to represent the energy flows along the lines, we introduce the complex voltages:

$$\mathbf{V}_k^t := \mathbf{U}_k^t e^{j\boldsymbol{\theta}_k^t} \in \mathbb{C}, \qquad \forall k \in \mathcal{N},$$

with variable voltage magnitudes $\boldsymbol{U}_k^t \in \mathbb{R}_+$ and voltage angles $\boldsymbol{\theta}_k^t$, respectively. Moreover, we establish the variables $\boldsymbol{\theta}_{lm}^t \in \mathbb{R}$ as the difference in voltage angle between the lth and mth bus, *i.e.* $\boldsymbol{\theta}_{lm}^t := \boldsymbol{\theta}_l^t - \boldsymbol{\theta}_m^t$. Without going into detail, using the above notations, a fairly accurate approximation of the steady-state behavior of the energy flows along the lines $(l, m) \in \mathcal{L}$ can be modeled by the following trigonometric expressions (*cf.* [51] and [14]):

$$\mathbf{p}_{lm}^t = (\mathbf{U}_l^t)^2 g_{lm} - \mathbf{U}_l^t \mathbf{U}_m^t g_{lm} \cos \boldsymbol{\theta}_{lm}^t - \mathbf{U}_l^t \mathbf{U}_m^t b_{lm} \sin \boldsymbol{\theta}_{lm}^t, \tag{2.12}$$

$$\mathbf{q}_{lm}^t = \mathbf{U}_l^t \mathbf{U}_m^t b_{lm} \cos \boldsymbol{\theta}_{lm}^t - \mathbf{U}_l^t \mathbf{U}_m^t g_{lm} \sin \boldsymbol{\theta}_{lm}^t - (\mathbf{U}_l^t)^2 (b_{lm} + b_{lm}^0), \tag{2.13}$$

where the given conductances $g_{lm} \in \mathbb{R}_+$, susceptances $b_{lm} \in \mathbb{R}_-$, and shunts b_{lm}^0 specify the line transmission capabilities. In doing so, the existing transformers are implicitly taken into account, since due to their existence, transmission capabilities (conductances, susceptances, and shunts) will be improved, such that the corresponding parameters can be readjusted.

[2] Here, j denotes the imaginary unit. This is to avoid confusion with the unit for electrical current.

[3] Normally, a load bus or the most powerful generator bus is chosen as slack bus.

For the grid, we claim that voltage magnitudes $|\mathbf{V}_k^t|$ $(= \mathbf{U}_k^t)$ have to be within particular bounds

$$V_k^{\min} \leq |\mathbf{V}_k^t| \leq V_k^{\max}, \qquad \forall k \in \mathcal{N}, \tag{2.14}$$

where we have $V_1^{\min} = V_1^{\max}$ at the slack bus, and that lines $(l, m) \in \mathcal{L}$ may not be overstrained, *i.e.* power flow is limited by the maximum transmission capacities $S_{lm}^{\max}, P_{lm}^{\max}, \varDelta V_{lm}^{\max} \in \mathbb{R}_+$:

$$(\mathbf{p}_{lm}^t)^2 + (\mathbf{q}_{lm}^t)^2 \leq (S_{lm}^{\max})^2, \qquad \forall (l, m) \in \mathcal{L}, \tag{2.15}$$

$$(\mathbf{p}_{lm}^t)^2 \leq (P_{lm}^{\max})^2, \qquad \forall (l, m) \in \mathcal{L}, \tag{2.16}$$

$$|\mathbf{V}_l^t - \mathbf{V}_m^t| \leq \varDelta V_{lm}^{\max}, \qquad \forall (l, m) \in \mathcal{L}. \tag{2.17}$$

It may happen that some of the constraints (2.14)–(2.17) are not needed in certain modeling situations. Then the vacuous constraints can be removed by setting their lower/upper bounds to $\pm\infty$.

3. Deterministic solution approach

The nonconvex AC power flow constraints (2.8)–(2.17) have been intensively studied in the literature and a multitude of algorithms have been proposed for solving optimization problems, taking into account these nonlinear restrictions [36, 37]. Most of these solution methods are based on solving the corresponding Karush–Kuhn–Tucker (KKT) conditions and thus at best guarantee local optimality.

Rather than to work with equations (2.12) and (2.13) directly, these are relaxed and approximated, respectively. The DC (direct current) power flow model, for instance, assumes that the difference of voltage angles is zero, that all voltage magnitudes are equal to one, and that the reactive power may be neglected (*cf*. Sect. 4.3 of [14] and [5]).

The DC power flow model being lossless, including these losses at least approximately will improve the model. In [42, 43], and [23], the DC model is refined by inclusion of Ohmic losses. These are modeled by trigonometric equations becoming relaxed to inequalities for computations. The relaxation is such that it overestimates losses and leads to convexity of the constraint set. Numerical optimization procedures heading for the minimization of losses then have the tendency to drive the overestimation back to zero, thus fulfilling the inequality as an equation.

In recent years, several convex relaxations were proposed, which are tight under certain conditions and thus provide a significantly better approximation of AC power flow than the DC approach and its extensions. These include Second Order Cone (SOC) [22], SDP [24], Convex-DistFlow (CDF) [13], and Quadratic Convex (QC) [21] relaxations. A comprehensive comparison of these relaxations is presented by Coffrin, Hijazi, and Hentenryck in [12]. There, it has been confirmed that the SDP relaxation is the tightest among the mentioned relaxations.

A wide class (2.12) of AC power flow models is presented in [24], where a convexification *via* a semidefinite programming relaxation may lead to globally optimal solutions. It is noted that this approach works for all IEEE benchmark systems (*cf*. [47]), provided a small resistance (10^{-5} per unit) is added to every transformer that originally is assumed to have zero resistance. This convexification does not work for all power grids – its limitations are examined in [26] as well as in [8].

For semidefinite programming, an accessible introduction is Chapter 2 of [20] as well as [48]. The recent state-of-the-art can be obtained from [2].

To solve the introduced deterministic unit commitment problem, we suggest a combination of the semidefinite programming (SDP) based algorithm by Lavaei and Low with a traditional Benders decomposition (tackling these programs by a form of a Benders algorithm can also be found in a recent work by Amjady and Ansari [1]). The basic idea is to separate the restrictions to the generators from the nonlinear conditions to the power grid, such that the latter can be tackled by the mentioned semidefinite approach.

3.1. Benders decomposition

The transformation of the above unit commitment problem into the required SDP format is outlined in Appendix A. Let us denote by $\mathfrak{W}$ the set of those matrices $\mathbf{W} \in \mathcal{S}_+^n$ (where, $\mathcal{S}_+^n$ denotes the set of symmetric and positive semidefinite matrices in $\mathbb{R}^{n\times n}$) fulfilling (A.4)–(A.7). This set describes the physical limits of the underlying grid with respect to voltage magnitude bounds (2.14) as well as line limitations (2.15)–(2.17) provided one claims in addition that $\mathbf{W} \in \mathfrak{W}$ has rank-one. Using this notation, our basic unit commitment problem can be equivalently expressed as:

$$
\begin{array}{l}
\min \sum_{t=1}^{T} \left(\sum_{k\in\mathcal{G}\backslash H} c_{k2}(Y_k \bullet \mathbf{W}^t + p_{D_k}^t)^2 + c_{k1}(Y_k \bullet \mathbf{W}^t + p_{D_k}^t) + c_{k0}\cdot \mathbf{r}_k^t \right) \\
\text{s.t.} \left.\begin{array}{l} \mathbf{u}_k^t P_k^{\min} \le \ Y_k \bullet \mathbf{W}^t + p_{D_k}^t \le \mathbf{u}_k^t P_k^{\max}, \quad \forall k \in \mathcal{G}\backslash H, \\ \mathbf{u}_k^t Q_k^{\min} \le \overline{Y}_k \bullet \mathbf{W}^t + q_{D_k}^t \le \mathbf{u}_k^t Q_k^{\max}, \quad \forall k \in \mathcal{G}\backslash H, \end{array}\right\} \begin{array}{l}\text{output bounds}\\ \text{and load coverage}\\ \text{at nonrenewables}\end{array} \\
\qquad \left.\begin{array}{l} -W_h^{\max} \le \ Y_h \bullet \mathbf{W}^t + p_{D_h}^t \le P_h^{\max}, \forall h \in H, \\ -\overline{W}_h^{\max} \le \overline{Y}_h \bullet \mathbf{W}^t + q_{D_h}^t \le Q_h^{\max}, \forall h \in H, \end{array}\right\} \begin{array}{l}\text{output bounds and}\\ \text{satisfaction of load at}\\ \text{pumped-storage plants}\end{array} \\
\qquad \left.\begin{array}{l} Y_n \bullet \mathbf{W}^t + p_{D_n}^t = 0, \quad \forall n \in \mathcal{N}\backslash\mathcal{G}, \\ \overline{Y}_n \bullet \mathbf{W}^t + q_{D_n}^t = 0, \quad \forall n \in \mathcal{N}\backslash\mathcal{G}, \end{array}\right\} \begin{array}{l}\text{load coverage at wind farms}\\ \text{and non-generator buses}\end{array} \\
\qquad \left.\begin{array}{l} \mathbf{u}_i^{t-1} - \mathbf{u}_i^t \le 1 - \mathbf{u}_i^l, \quad \forall i \in I, \quad t = 2,\dots,T-1, \\ \qquad\qquad\qquad\qquad l = t+1,\dots,\min\{t+\tau_i - 1, T\}, \end{array}\right\} \begin{array}{l}\text{min down}\\ \text{times for coal}\\ \text{fired blocks}\end{array} \\
\qquad \left.\begin{array}{ll} \mathbf{l}_h^0 = l_h^{\text{in}}, \quad \mathbf{l}_h^T = l_h^{\text{end}}, & \forall h \in H, \\ \mathbf{l}_h^t = \mathbf{l}_h^{t-1} - (\mathbf{p}_h^t - \eta_j \mathbf{w}_h^t) \le l_h^{\max}, & \forall h \in H, \\ \mathbf{p}_h^t - \mathbf{w}_h^t = Y_h \bullet \mathbf{W}^t + p_{D_h}^t, & \forall h \in H, \\ \mathbf{p}_h^t \ge 0, \quad \mathbf{w}_h^t \ge 0,, & \forall h \in H, \end{array}\right\} \begin{array}{l}\text{holding of max dam fills}\\ \text{plus considering of inter-}\\ \text{connections in pumped-}\\ \text{storage units}\end{array} \\
\qquad \mathbf{r}_k^t = \max\{\mathbf{u}_k^{t-1} - \mathbf{u}_k^t, 0\}, \quad \mathbf{u}_k^t \in \{0,1\}, \quad \forall k \in \mathcal{G}\backslash H, \\
\qquad \mathbf{W}^t \in \mathfrak{W}, \quad \operatorname{rank}(\mathbf{W}^t) = 1, \qquad\qquad\qquad \forall t \in \{1,\dots,T\}.
\end{array} \tag{3.1}
$$

Relaxing the rank-one conditions and linearizing the conditions to the grid in $\mathfrak{W}$ as well as its objective ($cf.$ Appendix A), this model becomes a mixed-integer linear semidefinite program.

The rank relaxation in (3.1) permits overload any time at any network bus ($cf.$ [46]). This may be beneficial in stressed network situations, and therefore could result in an infeasible commitment/dispatch decision. Still, it has been shown to be tight for tree networks [7,49] and for cyclic networks if every cycle contains a line with a controllable phase shifter [46].

By fixing all switching states together with the pumped-storage operation, coupling over time disappears, and (3.1) decomposes into T independent problems which are closely related to the continuous optimal power flow (OPF) problem.

Now, as in [1], the first step of our Benders decomposition algorithm treats the following mixed-integer linear programming (MILP) master problem:

$$
\begin{array}{c}
\mu_M := \min \ \sum_{t=1}^{T} \left(\sum_{k\in\mathcal{G}\backslash H} c_{k0}\cdot \mathbf{r}_k^t + \boldsymbol{\eta}_{Obj}^t \right) \\
\text{s.t.} \ \ (2.3), (2.4), \text{ and } (2.5),
\end{array} \tag{3.2}
$$

where $\mathcal{G}\backslash H$ is the set of all thermal generators, and $\boldsymbol{\eta}_{Obj}^t$ are additional nonnegative variables introduced for the objective cuts (measuring the power production costs for a feasible binary on/off assignment of the online thermal units). Note that (3.2) takes into account all constraints regarding unit commitment switching decisions as well as active power generation at pumped-storage units. The model relaxes the nonlinear conditions to the grid. After solving this master problem, its solution forms the input to a first set of subproblems. These subproblems emerge from fixing the solution to (3.2) in (3.1). This implies decoupling of time intervals and decomposition into the following $t = 1, \ldots, T$ SDP subproblems:

$$
\begin{aligned}
\mu_{Obj}^t := \min_{\mathbf{W}^t \in \mathfrak{W}} & \sum\nolimits_{k \in \mathcal{G} \setminus H} c_{k2}(Y_k \bullet \mathbf{W}^t + p_{D_k}^t)^2 + c_{k1}(Y_k \bullet \mathbf{W}^t + p_{D_k}^t) \\
\text{s.t. } & \mathbf{u}_k^t \cdot P_k^{\min} \leq Y_k \bullet \mathbf{W}^t + p_{D_k}^t \leq \mathbf{u}_k^t \cdot P_k^{\max}, \ \forall k \in \mathcal{G} \setminus H, \\
& \mathbf{u}_k^t \cdot Q_k^{\min} \leq \overline{Y}_k \bullet \mathbf{W}^t + q_{D_k}^t \leq \mathbf{u}_k^t \cdot Q_k^{\max}, \ \forall k \in \mathcal{G} \setminus H, \\
& Y_h \bullet \mathbf{W}^t + p_{D_h}^t = \mathbf{p}_h^t - \mathbf{w}_h^t, \quad \forall h \in H, \\
& -W_h^{\max} \leq \mathbf{p}_h^t - \mathbf{w}_h^t \leq P_h^{\max}, \quad \forall h \in H, \\
& -\overline{W}_h^{\max} \leq \overline{\mathbf{Y}}_h \bullet \mathbf{W}^t + q_{D_h}^t \leq Q_h^{\max}, \quad \forall h \in H, \\
& Y_n \bullet \mathbf{W}^t + p_{D_n}^t = 0, \quad \forall n \in N \setminus \mathcal{G}, \\
& \overline{Y}_n \bullet \mathbf{W}^t + q_{D_n}^t = 0, \quad \forall n \in N \setminus \mathcal{G}, \\
& \mathbf{u}_k^t = \widetilde{\mathrm{u}}_k^t, \ \forall k \in \mathcal{G} \setminus H, \quad \mathbf{p}_h^t - \mathbf{w}_h^t = \widetilde{\mathrm{p}}_{D_h}^t, \ \mathbf{p}_h^t, \mathbf{w}_h^t \geq 0, \ \forall h \in H,
\end{aligned}
\tag{3.3}
$$

where $\widetilde{\mathrm{u}}_k^t$ and $\widetilde{\mathrm{p}}_{D_h}^t := \widetilde{\mathrm{p}}_h^t - \widetilde{\mathrm{w}}_h^t$ denote the optimal solution to (3.2) delivering switching decisions for the installed thermal units and power output/consumption at pumped-storage units, respectively (here, the nonlinear objective once again can be linearized as described in (A.9)). If for $t \in \{1, \ldots, T\}$ its corresponding first subproblem becomes feasible, the following objective cut is added to (3.2):

$$
\boldsymbol{\eta}_{Obj}^t \geq \mu_{Obj}^t + \sum\nolimits_{k \in \mathcal{G}} \lambda_{Obj,k}^t (\mathbf{u}_k^t - \widetilde{\mathrm{u}}_k^t) + \sum\nolimits_{h \in H} \lambda_{Obj,h}^t (\mathbf{p}_h^t - \mathbf{w}_h^t - \widetilde{\mathrm{p}}_{D_h}^t), \tag{3.4}
$$

where $\lambda_{Obj,k}^t$ and $\lambda_{Obj,h}^t$ are the optimal dual variables with respect to the inserted constraints $\mathbf{u}_k^t = \widetilde{\mathrm{u}}_k^t$, $\forall k \in \mathcal{G} \setminus H$ and $Y_h \bullet \mathbf{W}^t + d_{D_h}^t = \widetilde{\mathrm{p}}_{D_h}^t$, $\forall h \in H$, respectively.

If otherwise, for $t \in T$ its associated first subproblem is infeasible, its infeasibility in terms of active power bounds at generator buses, voltage restrictions at net nodes, and network line limitations is measured by an appropriate second subproblem. To this end, the nonnegative auxiliary variables $\boldsymbol{z}_k^t, \boldsymbol{v}_n^t, \boldsymbol{p}_{lm}^t, \boldsymbol{m}_{lm}^t$, as well as $\boldsymbol{s}_{lm}^t$ are introduced, to reflect the violation of active power production bounds in (2.2) as well as the failure of the network limitations in (2.14), (2.15), (2.16), and (2.17) by means of the inequalities:

$$
\boldsymbol{z}_k^t \geq \mathbf{u}_k^t P_k^{\min} - (Y_k \bullet \mathbf{W}^t + d_{D_k}^t), \qquad \forall k \in \mathcal{G} \setminus H, \tag{3.5}
$$

$$
\boldsymbol{z}_k^t \geq (Y_k \bullet \mathbf{W}^t + d_{D_k}^t) - \mathbf{u}_k^t P_k^{\max}, \qquad \forall k \in \mathcal{G} \setminus H, \tag{3.6}
$$

$$
\boldsymbol{v}_n^t \geq (V_n^{\min})^2 - M_k \bullet \mathbf{W}^t, \quad \boldsymbol{v}_n^t \geq M_k \bullet \mathbf{W}^t - (V_n^{\max})^2, \qquad \forall n \in \mathcal{N}, \tag{3.7}
$$

$$
\boldsymbol{p}_{lm}^t \geq Y_{lm} \bullet \mathbf{W}^t - P_{lm}^{\max}, \quad \boldsymbol{m}_{lm}^t \geq M_{lm} \bullet \mathbf{W}^t - (\varDelta V_{lm}^{\max})^2, \qquad \forall (l,m) \in \mathcal{L}, \tag{3.8}
$$

$$
\boldsymbol{s}_{lm}^t \geq (Y_{lm} \bullet \mathbf{W})^2 + \left(\overline{Y}_{lm} \bullet \mathbf{W}\right)^2 - (S_{lm}^{\max})^2, \qquad \forall (l,m) \in \mathcal{L}, \tag{3.9}
$$

where once again (3.9) could be linearized ($cf.$ (A.8)). This leads to the following set of second subproblems:

$$\mu^t_{Feas} := \min \sum\nolimits_{i\in I} \boldsymbol{z}^t_k + \sum\nolimits_{n\in N} \boldsymbol{v}^t_n + \sum\nolimits_{(l,m)\in L} (\boldsymbol{p}^t_{lm} + \boldsymbol{m}^t_{lm} + \boldsymbol{s}^t_{lm})$$

$$\begin{aligned} &\text{s.t. (3.5), (3.6), (3.7), (3.8), and (3.9),} \\ &\quad \mathbf{u}^t_k \cdot Q^{\min}_k \le \overline{Y}_k \bullet \mathbf{W}^t + q^t_{D_k} \le \mathbf{u}^t_k \cdot Q^{\max}_k, \forall k \in \mathcal{G}\backslash H, \\ &\quad -W^{\max}_h \le \overline{Y}_h \bullet \mathbf{W}^t + q^t_{D_h} \le Q^{\max}_h, \qquad \forall h \in H, \\ &\quad Y_n \bullet \mathbf{W}^t + p^t_{D_n} = 0, \quad \forall n \in N\backslash(\mathcal{G}\cup H), \\ &\quad \overline{Y}_n \bullet \mathbf{W}^t + q^t_{D_n} = 0, \quad \forall n \in N\backslash(\mathcal{G}\cup H), \\ &\quad \mathbf{u}^t_k = \widetilde{\mathrm{u}}^t_k, \forall k \in \mathcal{G}, \quad Y_h \bullet \mathbf{W}^t + p^t_{D_h} = \widetilde{\mathrm{p}}^t_{D_h}, \forall h \in H, \\ &\quad \boldsymbol{z}^t_k,\ \boldsymbol{v}^t_n,\ \boldsymbol{p}^t_{lm},\ \boldsymbol{m}^t_{lm},\ \boldsymbol{s}^t_{lm} \ge 0, \quad \mathbf{W}^t \succeq 0. \end{aligned} \tag{3.10}$$

After solving this linear SDP, the subsequent feasibility cut (which will guarantee that the current commitment will be cut off from the feasible region of our master problem) is added to (3.2):

$$0 \ge \mu^t_{Feas} + \sum\nolimits_{k\in\mathcal{G}} \lambda^t_{Feas,k}(\mathbf{u}^t_k - \widetilde{\mathrm{u}}^t_k) + \sum\nolimits_{h\in H} \lambda^t_{Feas,h}(\mathbf{p}^t_h - \mathbf{w}^t_h - \widetilde{\mathrm{p}}^t_h). \tag{3.11}$$

The suggested solution procedure can now be summarized by the following algorithm framework.

Algorithm framework for unit commitment in AC grids.

Initialize:	Accuracy parameter $\epsilon > 0$; set $\varphi_{UB} := \infty$, and $\varphi_{LB} := 0$; solve (3.2) with $\boldsymbol{\eta}^t_{Obj} = 0$ and obtain $\widetilde{\mathrm{u}}$ as well as $\widetilde{\mathrm{p}}_H$ from its solution.
Step 1.	Update lower bound $\varphi_{LB} := \mu_M - \sum_{t=1}^T \widetilde{\eta}^t_{Obj}$; solve (for $t = 1,\dots,T$) the first set of subproblems (3.3); if subproblem t becomes feasible, keep μ^t_{Obj}, $\lambda^t_{Obj,k}$, $\lambda^t_{Obj,h}$; else, solve second subproblem (3.10) and keep μ^t_{Feas}, $\lambda^t_{Feas,k}$, $\lambda^t_{Feas,h}$.
Step 2.	If all first set subproblems become feasible, update upper bound: $\varphi_{UB} := \mu_M + \sum_{t=1}^T \mu^t_{Obj} - \sum_{t=1}^T \widetilde{\eta}^t_{Obj}$; if $\frac{\lvert\varphi_{UB}-\varphi_{LB}\rvert}{\lvert\varphi_{LB}\rvert} < \epsilon$ GOTO Step 4;
Step 3.	Add all generated cuts to the master problem; solve this new master problem and update $\mu_M, \widetilde{\eta}_{Obj}, \widetilde{\mathrm{u}}$ plus $\widetilde{\mathrm{p}}_H$; GOTO Step 1.
Step 4.	Try to recover the rank-one conditions (see Appendix B).

4. Inclusion of uncertainties

In this section we will focus on planning a unit commitment schedule under uncertainty of both power demand and output of renewables. Hence, the uncertainties, at time interval t occur at the nodes (buses) and concern the active and reactive (apparent) power, denoted by $p^t(\omega)$ and $q^t(\omega)$ for $t = 1,\dots,T$, respectively.

We assume that $z(\omega) = (p^t(\omega), q^t(\omega))$ is a random variable whose probability distribution is known at the beginning of the optimization horizon. The latter, alone, already is non-trivial, and obtaining meaningful probability distributions from statistical data is a field of active research in stochastic programming and beyond, see [19], for instance. In our case we will adopt a finite event space where the realizations and their probabilities are obtained from recorded load profiles of the past.

For modeling and tackling programs involving uncertain data, stochastic programming provides several approaches. Moreover, it is a useful tool for making discrete decisions under uncertainty. For an introduction into basic aspects of stochastic programming, we refer to the books [6, 45], and [41].

In power planning, multi-stage stochastic programming, [45], became more and more established in recent years. For the scope adopted in the present paper with its elaborate model of power flow, however, the multi-stage approach still seems premature, at least computationally. Here, two-stage models still pose challenging research questions.

Concerning their operational flexibility the on/off decisions of the coal fired thermal units are the most inertial ones. Even when making decisions with respect to a rather coarse, hourly time discretization, for instance, it is not possible to follow a random load-and-renewables profile by on/off determinations of thermal blocks alone. This observation leads to modeling the switching decisions of the coal fired thermal units as first-stage variables. The second-stage is formed by the remaining short term on/off decision for gas turbines and by the operation levels of the on-line thermal and pumped-storage units.

Denoting then by $\mathbf{u}_I = \{\mathbf{u}_i\}_{i\in I}$ and $\mathbf{u}_R = \{\mathbf{u}_r\}_{i\in R}$ the boolean vectors for switching decisions of coal fired blocks as well as gas turbines, and in addition by $\mathcal{U}_I$ and $\mathcal{U}_R$ their feasible sets, this leads to a random two-stage optimization problem of the following principal shape:

$$\min\left\{ c_I^T\mathbf{u}_I + \mathcal{H}(\mathbf{u}_R,\mathbf{W}) \ : \begin{array}{l} T\mathbf{u}_I + \mathcal{W}(\mathbf{u}_R,\mathbf{W}) = z(\omega),\ \mathbf{u}_I \in \mathcal{U}_I, \\ (\mathbf{u}_R,\mathbf{W}) \in \mathcal{U}_R \times \mathfrak{W},\ \operatorname{rank}(\mathbf{W}) = 1 \end{array} \right\}. \tag{4.1}$$

Here $T, \mathcal{W}$, and $\mathcal{H}$ are the appropriate linear operators, describing the conditions and the objective in (3.1), respectively. At this point, it is emphasized that the above program is not well-posed. Namely, as long as the realizations of the random variable $z(\omega)$ are unknown, it exhibits for every fixed first-stage determination $\mathbf{u}_I$ another random variable. Hence, "minimization" in (4.1) can be seen as selecting the "best" member among the resulting family of random variables, and this in turn raises the question of how to rank this random variables. The stochastic literature offers several different possibilities of ranking or comparing of random variables (for deeper insights into comparison methods for random variables, we refer to the book by Müller and Stoyan [38]). Beneath this ranking opportunities we will pick up ranking by statistical parameters in terms of risk aversion *via* excess probabilities in mean-risk models.

4.1. Risk aversion *via* excess probabilities in two-stage stochastic semidefinite programs

Relaxing the nonlinear rank constraint in the stochastic program (4.1), we arrive at the following general two-stage stochastic (mixed integer) linear semidefinite program:

$$\min\left\{ C \bullet \boldsymbol{X} + H \bullet \boldsymbol{Y} \ : \ \mathcal{T}\boldsymbol{X} + \mathcal{W}\boldsymbol{Y} = z(\omega),\ \boldsymbol{X} \in \mathcal{X},\ \boldsymbol{Y} \in \mathcal{Y} \right\}, \tag{4.2}$$

where $\mathcal{X} \subseteq \mathcal{S}_+^{m_1}$, $\mathcal{Y} \subseteq \mathcal{S}_+^{m_2}$ are nonempty spectrahedra (intersections of solution sets of affine matrix inequalities with the cone of positive semidefinite matrices) with possibly additional integer requirements to some or all matrix entries. Moreover, $\mathcal{T} : \mathcal{S}^{m_1} \to \mathbb{R}^s$ as well as $\mathcal{W} : \mathcal{S}^{m_2} \to \mathbb{R}^s$ are linear operators, and the right-hand side $z(\omega)$ is a random vector on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}^s$. The distribution of $z(\omega)$ does not depend on the first-stage decisions X. Here, the second-stage variable Y can be interpreted as compensating or recourse action, which, in the seminal stochastic programming literature, led to the notion of a stochastic program with recourse. To emphasize the two-stage character of decision making we rewrite program (4.2) as:

$$\min_X \left\{ C \bullet \boldsymbol{X} + \underbrace{\min_{\boldsymbol{Y}} \left\{ H \bullet \boldsymbol{Y} \ : \ \mathcal{W}\boldsymbol{Y} = z(\omega) - \mathcal{T}\boldsymbol{X},\ \boldsymbol{Y} \in \mathcal{Y} \right\}}_{=:\Phi(z(\omega)-\mathcal{T}\boldsymbol{X})} \ : \ \boldsymbol{X} \in \mathcal{X} \right\}. \tag{4.3}$$

Here, the function

$$\Phi : \mathbb{R}^s \to \mathbb{R}, \qquad \boldsymbol{t} \mapsto \min\left\{ H \bullet \boldsymbol{Y} : \mathcal{W}\boldsymbol{Y} = \boldsymbol{t}, \boldsymbol{Y} \in \mathcal{Y} \right\} \tag{4.4}$$

is the optimal-value function of the inner semidefinite program seen as a parametric optimization problem with parameter t.

Defining random variables $f(\boldsymbol{X},\omega) := C \bullet \boldsymbol{X} + \Phi(z(\omega) - \mathcal{T}\boldsymbol{X}),\ \boldsymbol{X} \in \mathcal{X}$, the random program in (4.3) turns into a minimization problem over a family of random variables.

In the present paper, we address this minimization by the mean-risk model

$$\min\{\mathcal{Q}_{\mathbb{E}}(\boldsymbol{X}) + \rho \cdot \mathcal{Q}_{\mathbb{P}^{\eta}}(\boldsymbol{X}) : \boldsymbol{X} \in \mathcal{X}\}, \tag{4.5}$$

where $\mathcal{Q}_{\mathbb{E}}$ denotes the mean value and $\mathcal{Q}_{\mathbb{P}^{\eta}}$ stands for the risk measure

$$\mathcal{Q}_{\mathbb{P}^{\eta}}(\boldsymbol{X}) := \mathbb{P}[\{\omega : f(\boldsymbol{X},\omega) > \eta\}], \tag{4.6}$$

i.e., the probability of exceeding a prescribed target level $\eta \in \mathbb{R}$.

As in stochastic programming, we have basic assumptions ensuring that model ingredients are well-defined. More specifically, let $\mathcal{Y} = \mathcal{S}_+^{m_2}$, and assume what is called complete recourse in stochastic programming

$$\mathcal{W}(\mathcal{S}_+^{m_2}) = \mathbb{R}^s.$$

This serves the purpose to have a non-empty feasible set for the second-stage optimization problem for any right-hand side. If, furthermore

$$M_D := \{u \in \mathbb{R}^s : \mathcal{W}^T u \prec H\} \neq \emptyset,$$

then, by duality the second-stage problem is always solvable.

Assume that the underlying random variable $z(\omega)$ follows a finite discrete probability distribution with realizations z_ω and probabilities $\pi_\omega, \omega = 1,\dots,S$. Then (4.5) adopts a block structure unveiled in the following theorem.

Theorem 4.1. *Assume $\mathcal{W}(\mathcal{S}_+^{m_2}) = \mathbb{R}^s$, $M_D := \{u \in \mathbb{R}^s : \mathcal{W}^T u \prec H\} \neq \emptyset$, and that $\mathcal{X}$ is compact. Then there exists a constant $M > 0$ such that the Excess Probability mean-risk model (4.5) is equivalent to*

$$\begin{aligned}
&\min C \bullet \boldsymbol{X} + \sum\nolimits_{\omega=1}^{S} \pi_\omega H \bullet \boldsymbol{Y}_\omega + \rho \cdot \sum\nolimits_{\omega=1}^{S} \pi_\omega \boldsymbol{\theta}_\omega \\
&\ s.t.\ \mathcal{T} \bullet \boldsymbol{X} + \mathcal{W} \bullet \boldsymbol{Y}_\omega = z_\omega, \\
&\qquad C \bullet \boldsymbol{X} + H \bullet \boldsymbol{Y}_\omega - M\boldsymbol{\theta}_\omega \le \eta, \\
&\qquad \boldsymbol{X} \in \mathcal{X}, \quad \boldsymbol{Y}_\omega \succeq 0, \quad \boldsymbol{\theta}_\omega \in \{0,1\}, \qquad \omega = 1,\dots,S.
\end{aligned} \tag{4.7}$$

Proof. Before showing the equivalence of the mentioned models, let us confirm that compactness of $\mathcal{X}$ yields existence of the required constant M. Indeed, let

$$M > \sup\{C \bullet \boldsymbol{X} + \Phi(z_\omega - \mathcal{T}\boldsymbol{X}) : \boldsymbol{X} \in \mathcal{X}, \omega = 1,\dots,S\} - \eta. \tag{4.8}$$

To see that the supremum on the right is bounded consider for each $\omega = 1,\dots,S$ the estimate

$$\sup_{\boldsymbol{X}\in\mathcal{X}} C \bullet \boldsymbol{X} + \Phi(z_\omega - \mathcal{T}\boldsymbol{X}) \le \sup_{\boldsymbol{X}\in\mathcal{X}} \|C\| \cdot \|\boldsymbol{X}\| + \sup_{\boldsymbol{X}\in\mathcal{X}} \max_{\boldsymbol{u}\in M_D^{\preceq}} (z_\omega - \mathcal{T}\boldsymbol{X})^T \boldsymbol{u}, \tag{4.9}$$

here $M_D^{\preceq}$ denotes the set $\{u \in \mathbb{R}^s : \mathcal{W}^T u \preceq H\}$, which is closed due to the continuity of the eigenvalue and the fact, that a matrix is positive semidefinite if and only if all its eigenvalues are non-negative. If now $M_D^{\preceq}$ were compact, so were $\mathcal{X} \times M_D^{\preceq}$. Since both $\|C\| \cdot \|\boldsymbol{X}\|$ and $(z_\omega - \mathcal{T}\boldsymbol{X})^T\boldsymbol{u}$ are continuous functions in $(\boldsymbol{X},\boldsymbol{u})$, finiteness in (4.8) would follow *via* (4.9) from Weierstrass theorem.

In order to show that $M_D^{\preceq}$ is bounded, let us assume that there exists a sequence $(v_n)_{n\in\mathbb{N}} \in M_D^{\preceq}$ with $\|v_n\| \to \infty$. If we further define $\widetilde{u}_n := v_n/\|v_n\|$, then $\|\widetilde{u}_n\| \to 1$, such that there is a subsequence converging to some $\widetilde{u} \neq 0$. The latter limit satisfies $\mathcal{W}^T \widetilde{u} \preceq 0$ which is seen as follows: By $v_n \in M_D^{\preceq}$ we have $H - \mathcal{W}^T v_n \succeq 0, \forall n \in \mathbb{N}$. Dividing by $\|v_n\|$ yields

$$\underbrace{\lim_{n\to\infty} \frac{1}{\|v_n\|} H - \mathcal{W}^T \frac{v_n}{\|v_n\|}}_{=-\mathcal{W}^T \widetilde{u}} \succeq 0.$$

Now, $\overline{u} \in M_D^{\preceq}$ implies $\overline{u} + \alpha \widetilde{u} \in M_D^{\preceq}$ for $\alpha \geq 0$. Therefore,

$$\widetilde{u}^T (\overline{u} + \alpha \widetilde{u}) = \widetilde{u}^T \overline{u} + \alpha \|\widetilde{u}\|^2 \to \infty, \quad \text{for } \alpha \to \infty,$$

verifying

$$\sup \left\{ \widetilde{u}^T u \ : \ \mathcal{W}^T u \preceq H \right\} = \infty.$$

By duality, the primal feasible set $\left\{ Y \in S_+^{m_2} \ : \ \mathcal{W}Y = \widetilde{u} \right\}$ then has to be empty which contradicts our assumption $\mathcal{W}(\mathcal{S}_+^{m_2}) = \mathbb{R}^s$.

Now let us turn to the equivalence of the models (4.5) and (4.7). Let $\overline{X}$ be an optimal solution to (4.5) and assume there is a feasible (X^*, Y^*, θ^*) to (4.7) whose objective value in (4.5) is less than $\mathcal{Q}_{\mathbb{E}}(\overline{X}) + \rho \cdot \mathcal{Q}_{\mathbb{P}^\eta}(\overline{X})$.

By the definition of Φ, see (4.4), it holds $\Phi(z_\omega - \mathcal{T} X^*) \leq H \bullet Y_\omega^*, \forall \omega$. This yields

$$\mathcal{Q}_{\mathbb{E}}(X^*) = C \bullet X^* + \textstyle\sum_{\omega=1}^S \pi_\omega \Phi(z_\omega - \mathcal{T} X^*) \leq C \bullet X^* + \sum_{\omega=1}^S \pi_\omega H \bullet Y_\omega^*$$

and, moreover, the following implication holds:

$$\theta_\omega^* = 0 \quad \Rightarrow \quad C \bullet X^* + \Phi(z_\omega - \mathcal{T} X^*) \leq \eta.$$

Thus, we obtain the inclusion

$$\{\omega : C \bullet X^* + \Phi(z_\omega - \mathcal{T} X^*) > \eta\} \subseteq \{\omega : \theta_\omega^* = 1\},$$

which yields $\mathcal{Q}_{\mathbb{P}^\eta}(X^*) \leq \sum_{\omega=1}^S \pi_\omega \theta_\omega^*$. Altogether, we get

$$\begin{aligned} \mathcal{Q}_{\mathbb{E}}(X^*) + \rho \cdot \mathcal{Q}_{\mathbb{P}^\eta}(X^*) &\leq C \bullet X^* + \sum_{\omega=1}^S \pi_\omega H \bullet Y_\omega^* + \sum_{\omega=1}^S \pi_\omega \theta_\omega^* \\ &< \mathcal{Q}_{\mathbb{E}}(\overline{X}) + \rho \cdot \mathcal{Q}_{\mathbb{P}^\eta}(\overline{X}), \end{aligned}$$

contradicting the optimality of $\overline{X}$ in (4.5), *i.e.* the optimal value of (4.7) is, in any case, an upper bound.

Furthermore, to see equality, let *vice versa* $\overline{X}$ be optimal in (4.5). Set

$$\overline{Y}_\omega \in \arg\min\{H \bullet \boldsymbol{Y}_\omega : \mathcal{W}\boldsymbol{Y}_\omega = z_\omega - \mathcal{T}\overline{X}, \ \boldsymbol{Y}_\omega \succeq 0\},$$

and

$$\overline{\theta}_\omega = \begin{cases} 0 & \text{if} \quad C \bullet \overline{X} + H \bullet \overline{Y}_\omega - \eta \leq 0, \\ 1 & \text{otherwise,} \end{cases}$$

for $\omega = 1, \ldots, S$. Then

$$\mathcal{Q}_{\mathbb{E}}(\overline{X}) + \rho \cdot \mathcal{Q}_{\mathbb{P}^\eta}(\overline{X}) = C \bullet \overline{X} + \textstyle\sum_{\omega=1}^S \pi_\omega H \bullet \overline{Y}_\omega + \rho \cdot \sum_{\omega=1}^S \pi_\omega \overline{\theta}_\omega,$$

where in addition $(\overline{X}, \overline{Y}, \overline{\theta})$ is feasible to (4.7). This completes the proof. □

Note that the above proposition remains valid in case of additional integer requirements to second stage variables. Indeed, by passing to its SDP relaxation, estimation (4.9) and compactness of the dual feasible set $M_D^{\preceq}$ hold true as well, whereas the rest of the proof continues completely analogously.

The structure of (4.7) has similarity to two-stage chance-constrained models. While (4.7) penalizes the violation of $f(\boldsymbol{X},\omega) > \eta$ by a multiple of its probability, chance-constrained models limit the choice of $\boldsymbol{X} \in \mathcal{X}$ by postulating that the probability $\mathbb{P}[\{\omega : f(\boldsymbol{X},\omega) > \eta\}]$ is less or equal to a given threshold. A recent work considering two-stage chance-constrained models is [27].

4.2. Algorithmic treatment: Lagrangean relaxation of nonanticipativity

For simplicity, let us first neglect the risk measure functional $\mathcal{Q}_{\mathbb{P}^\eta}$ in (4.5) and consider the risk neutral model

$$\min_{X \in \mathcal{X}} \mathcal{Q}_{\mathbb{E}}(X). \tag{4.10}$$

According to Theorem 4.1, (4.10) can be equivalently expressed by:

$$\begin{array}{lllll} \min & C \bullet \boldsymbol{X} + \sum_{\omega=1}^{S} H_\omega \bullet \boldsymbol{Y}_\omega & & & \\ \text{s.t.} & \mathcal{T} \bullet \boldsymbol{X} + \mathcal{W} \bullet \boldsymbol{Y}_1 & & & = z_1, \\ & \vdots & \ddots & & \vdots \\ & \mathcal{T} \bullet \boldsymbol{X} & & +\mathcal{W} \bullet \boldsymbol{Y}_S & = z_S, \\ & \boldsymbol{X} \in \mathcal{X}, \quad \boldsymbol{Y}_1 \succeq 0, & \ldots, & \boldsymbol{Y}_S \succeq 0. & \end{array} \tag{4.11}$$

Here, we have tacitly denoted $H_\omega := \pi_\omega H$. Since, with growing number of scenarios, the dimension of (4.11) quickly becomes too large for being handled in all-at-once manner by general SDP solvers, decomposition methods come to the fore. Mehrotra and Özevin [30] propose an extension of [50] to semidefinite programs, leading to a Benders decomposition based interior point method. While the latter works well if there are no integer variables, it fails with integer requirements to second-stage variables. The problems studied in the present paper, however, contain substantial numbers of Boolean decision variables in the second stage.

To solve problem (4.11), we will pursue the concept of Carøe and Schultz (see [10]). Here, Lagrangean relaxation (*cf.* [25]) of the nonanticipative first-stage decision is recommended, which then leads to decomposition into smaller subproblems of tractable dimensions. In implementing this idea, we will closely follow [28]. To this end, we introduce an additional matrix variable $\boldsymbol{X}^*$ plus copies $\boldsymbol{X}_\omega, \omega = 1,\ldots,S$ of the first-stage variable $\boldsymbol{X}$, and add the requirements

$$\boldsymbol{X}_\omega - \boldsymbol{X}^* = 0, \qquad \omega = 1,\ldots,S. \tag{4.12}$$

In doing so, we obtain the following equivalent reformulation of (4.11):

$$\begin{array}{llll} \min & \sum_{\omega=1}^{S} C \bullet \boldsymbol{X}_\omega + H_\omega \bullet \boldsymbol{Y}_\omega & & \\ \text{s.t.} & \mathcal{T} \bullet \boldsymbol{X}_1 + \mathcal{W} \bullet \boldsymbol{Y}_1 & & = z_\omega, \\ & & \ddots & \vdots \\ & & \mathcal{T} \bullet \boldsymbol{X}_S + \mathcal{W} \bullet \boldsymbol{Y}_S & = z_S, \\ & & \boldsymbol{X}_\omega - \boldsymbol{X}^* & = 0, \quad \omega = 1,\ldots,S, \\ & \boldsymbol{X}_1 \in \mathcal{X},\ \boldsymbol{Y}_1 \succeq 0, & \ldots,\ \boldsymbol{X}_S \in \mathcal{X},\ \boldsymbol{Y}_S \succeq 0. & \end{array}$$

Relaxing nonanticipativity (4.12) leads to S independent subproblems, each corresponding to a particular scenario. In context of semidefinite programming, we arrive at the following Lagrangean function:

$$L\big(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_S,\boldsymbol{X}^*,\boldsymbol{Y}_1,\ldots,\boldsymbol{Y}_S\big) := \sum_{\omega=1}^{S} L_\omega\big(\boldsymbol{X}_\omega,\boldsymbol{X}^*,\boldsymbol{Y}_\omega\big), \tag{4.13}$$

with $L_\omega\big(\boldsymbol{X}_\omega,\boldsymbol{X}^*,\boldsymbol{Y}_\omega\big) := C \bullet \boldsymbol{X}_\omega + H_\omega \bullet \boldsymbol{Y}_\omega + \Lambda_\omega \bullet \big(\boldsymbol{X}_\omega - \boldsymbol{X}^*\big)$. Thus, with the dual function

$$D(\boldsymbol{\Lambda}) := \min \left\{ \sum_{\omega=1}^{S} L_\omega\big(\boldsymbol{X}_\omega,\boldsymbol{X}^*,\boldsymbol{Y}_\omega\big) \;\; \begin{array}{l} : \mathcal{T}\boldsymbol{X}_\omega + \mathcal{W}\boldsymbol{Y}_\omega = z_\omega,\ \omega = 1,\ldots,S \\ \boldsymbol{X}_\omega \in \mathcal{X}, \quad \boldsymbol{Y}_\omega \succeq 0,\ \omega = 1,\ldots,S \end{array} \right\},$$

we obtain the associated Lagrangean dual:

$$\max\left\{D\big(\boldsymbol{\Lambda}_1,\ldots,\boldsymbol{\Lambda}_S\big):\boldsymbol{\Lambda}_\omega\in\mathcal{S}^{m_1}\right\}.$$

Now, as the auxiliary variable $\boldsymbol{X}^*$ is unconstrained, its coefficients must cancel out when forming the sum for $\omega=1,\ldots S$, *i.e.* $\sum_{\omega=1}^{S}\boldsymbol{\Lambda}_\omega=0$. We further mention that the dual function is separable, *i.e.* by determining

$$D_\omega\big(\boldsymbol{\Lambda_\omega}\big):=\min_{\boldsymbol{X}_\omega,\boldsymbol{Y}_\omega}\left\{\begin{aligned}C\bullet\boldsymbol{X}_\omega+H_\omega\bullet\boldsymbol{Y}_\omega+\boldsymbol{\Lambda}_\omega\bullet\boldsymbol{X}_\omega:\mathcal{T}\boldsymbol{X}_\omega+\mathcal{W}\boldsymbol{Y}_\omega=z_\omega,\\ \boldsymbol{X}_\omega\in\mathcal{X},\quad\boldsymbol{Y}_\omega\succeq 0\end{aligned}\right\},\tag{4.14}$$

for $\omega=1,\ldots,S$, we obtain $D\big(\boldsymbol{\Lambda}_1,\ldots,\boldsymbol{\Lambda}_S\big)=\sum_{\omega=1}^{S}D_\omega\big(\boldsymbol{\Lambda}_\omega\big)$. Using this notation, the Lagrangean dual, arising by Lagrangean relaxation of the nonanticipativity condition (4.12), can be expressed by

$$\max\left\{\sum\nolimits_{\omega=1}^{S}D_\omega\big(\boldsymbol{\Lambda}_\omega\big):\sum\nolimits_{\omega=1}^{S}\boldsymbol{\Lambda}_\omega=0\right\},$$

and this in turn is equivalent to

$$\max_{\theta,\Lambda}\left\{\sum\nolimits_{\omega=1}^{S}\boldsymbol{\theta}_\omega:\sum\nolimits_{\omega=1}^{S}\boldsymbol{\Lambda}_\omega=0,\quad\boldsymbol{\theta}_\omega\le D_\omega\big(\boldsymbol{\Lambda}_\omega\big),\ \omega=1,\ldots,S\right\}.\tag{4.15}$$

In order to solve the above problem, we will apply proximal bundle methods (*cf.* [40]). The basic idea is to approximate the constraints in (4.15) by cutting planes and adding a regularization term to the objective. At each iteration K, this leads to:

$$\begin{aligned}\max_{\theta,\Lambda}\quad&\sum\nolimits_{\omega=1}^{S}\boldsymbol{\theta}_\omega-\frac{1}{2}\tau\sum\nolimits_{\omega=1}^{S}\big\|\boldsymbol{\Lambda}_\omega-\Lambda_\omega^+\big\|_F^2\\ \text{s.t.}\quad&\sum\nolimits_{\omega=1}^{S}\boldsymbol{\Lambda}_\omega=0,\\ &\boldsymbol{\theta}_\omega\le D_\omega\big(\Lambda_\omega^{(k)}\big)+X_\omega^{(k)}\bullet\Big(\boldsymbol{\Lambda}_\omega-\Lambda_\omega^{(k)}\Big),\quad\omega=1,\ldots,S,\quad k=1,\ldots,K.\end{aligned}\tag{4.16}$$

Taking into account that $-D_\omega$ is convex for all $\omega=1,\ldots,S$, $-X_\omega^{(k)}$ is selected as a member of the subdifferential $\partial[-D_\omega](\Lambda_\omega^{(k)})$, given by

$$\left\{\boldsymbol{X}\in\mathcal{S}^{m_1}:D_\omega(\boldsymbol{\Lambda})-D_\omega(\Lambda_\omega^{(k)})+\boldsymbol{X}\bullet(\boldsymbol{\Lambda}-\Lambda_\omega^{(k)})\le 0,\ \forall\boldsymbol{\Lambda}\in\mathcal{S}^{m_1}\right\},$$

where $(-1)\partial[-D_\omega](\Lambda_\omega^{(k)})$ coincides with the $\boldsymbol{X}_\omega$-part of the optimal solution set to program (4.14). The Point $(\Lambda_\omega^+,\ldots,\Lambda_\omega^+)$ is the current proximal center, fulfilling $\sum_{\omega=1}^{S}\Lambda_\omega^+=0$, and τ is some regularization parameter which is normally adjusted at each iteration.

Converting the arising matrices into vector form by the so called vec operator that stacks the matrix columns on top of each other, (4.16) becomes a quadratic program (QP), *i.e.* it can be tackled by well-established algorithms (here, one could briefly mention active set strategies, trust region methods, conjugate gradient methods, and interior point methods).

Finally, we arrive at the following decomposition method:

Decomposition based proximal bundle method

Initialize: Accuracy parameter $\epsilon > 0$; $m = 0.1$; $K := 1$;
set for $\omega = 1, \dots, S$, $\Lambda_\omega^+ := 0$ as well as $\Lambda_\omega^{(K)} := 0$;
solve $D_\omega(\Lambda_\omega^{(K)}), \omega = 1, \dots, S$, save optimal solution $X_\omega^{(K)}$;
and put $curObj := \sum_{\omega=1}^S D_\omega(\Lambda_\omega^{(K)})$.

Step 1. Solve (4.16), obtaining optimal θ_ω^* and Λ_ω^*, for $\omega = 1, \dots, S$.

Step 2. Let $v = \left(\sum_\omega^S \theta_\omega^*\right) - curObj$.
If $v/(1 + |curObj|) < \epsilon$ terminate; else continue.

Step 3. $K := K + 1$; solve $D_\omega(\Lambda_\omega^*), \omega = 1, \dots, S$, save its optimal value $D_\omega(\Lambda_\omega^{(K)})$ and its corresponding solution $X_\omega^{(K)}$;
$newObj := \sum_{\omega=1}^S D_\omega(\Lambda_\omega^{(K)})$; $u := 2\tau(1 - (newObj - curObj)/v)$,
$\tau := \min(\max(u, \tau/10, 10^{-4}), 10\tau)$;
if $(newObj - curObj > m \cdot v)$, then update $\Lambda_\omega^+ := \Lambda_\omega^*$ and
$curObj := newObj$; GOTO Step 1.

4.3. Embedding into branch-and-bound - Enhanced by heuristics

At first, it is shown that the risk averse program (4.7) can be equivalently transformed into (4.11) with the exception that it contains additional integer requirements to some second-stage variables. To this end, we introduce for $\omega = 1, \dots, S$ the second-stage variables

$$\boldsymbol{Y}_{EP\omega} := \operatorname{diag}(\boldsymbol{Y}_\omega, \boldsymbol{\theta}_\omega, \boldsymbol{s}_\omega),$$

and extend the linear matrix operator $\mathcal{T}$ to the linear operator $\mathcal{T}_{EP} : \mathcal{S}^{m_1} \to \mathbb{R}^{s+1}$, defined by $\mathcal{T}_{EP}\boldsymbol{X} = [(\mathcal{T}\boldsymbol{X})^T, C \bullet \boldsymbol{X}]^T$. Moreover, let us define the linear operator $\mathcal{W}_{EP} : \mathcal{S}^{m_2} \to \mathbb{R}^{s+1}$, given by the following modified recourse matrices:

$$W_{EP_1} := \operatorname{diag}(W_1, 0, 0), \ \dots, \ W_{EP_s} := \operatorname{diag}(W_s, 0, 0), \ W_{EP_{s+1}} := \operatorname{diag}(H, -M, 1).$$

Finally, by setting $H_{EP\omega} := \operatorname{diag}(\pi_\omega H, \rho, 0)$ and $z_{EP\omega} := (z_\omega^T, \eta)^T$ for $\omega = 1, \dots, S$, we obtain that (4.7) is indeed equivalent to:

$$\min \left\{ C \bullet \boldsymbol{X} + \sum_{\omega=1}^{S} H_{EP\omega} \bullet \boldsymbol{Y}_{EP\omega} : \begin{array}{ll} \mathcal{T}_{EP}\boldsymbol{X} + \mathcal{W}_{EP}\boldsymbol{Y}_{EP\omega} = z_{EP\omega}, & \forall\omega \\ \boldsymbol{X} \in \mathcal{X}, \quad \boldsymbol{Y}_{EP\omega} \succeq 0, & \forall\omega \\ \boldsymbol{Y}_{EP\omega}(m_2+1, m_2+1) \in \{0,1\}, & \forall\omega \end{array} \right\}, \tag{4.17}$$

which obviously has the same structure as the risk neutral model (4.11) in the sense that there are no constraints involving second-stage variables from different scenarios.

Tackling the nonconvex program (4.17) by the proposed proximal bundle methods may result in a solution that does not meet the nonanticipativity condition. If so, the solution to the Lagrangean dual (4.15) provides us a lower bound. To measure the quality of this lower bound, upper bounds in terms of feasible points to (4.17) are required. Since the previously relaxed constraints (4.12) are quite simple, namely, we have to make all first-stage copies identical, ideas for heuristics come up straightforwardly. In this paper we have picked from $X_\omega^{opt}, \omega = 1, \dots, S$ a candidate, by averaging over them all and rounding to integers.

If the resulting gap is unsatisfactory, we recommend the embedding into a Branch-and-Bound scheme, where the underlying two-stage stochastic program is understood as a nonconvex global minimization problem. For the considered unit commitment problems this results in the following algorithm framework.

Dual decomposition for unit commitment under uncertainty

Initialize:	Let $\mathbf{P}$ be the list of current problems. Denote for $P \in \mathbf{P}$ by $\varphi_{\mathrm{LD}}(P)$ its Lagrangean lower bound[a] that is obtained by the proximal bundle methods presented in Section 4.2, where the decomposed programs $D_\xi(\boldsymbol{\Lambda}_\xi)$ are solved by the Benders decomposition approach from Section 3.1. Put $\overline{\varphi} = +\infty$ and add the underlying problem to the list $\mathbf{P}$.
Step 1.	If $\mathbf{P} = \emptyset$ then $\overline{u}$ with $\overline{\varphi} = \mathcal{Q}_{\mathbb{E}}(\overline{u}) + \rho \cdot \mathcal{Q}_{\mathbb{P}_\eta}(\overline{u})$ is optimal; Else GOTO Step 2.
Step 2.	Select and delete from the list $\mathbf{P}$ a problem $P \in \mathbf{P}$ and solve its Lagrangean dual. If $\varphi_{\mathrm{LD}}(P)$ is $+\infty$, GOTO Step 1; Otherwise GOTO Step 3.
Step 3.	If $\varphi_{\mathrm{LD}}(P) \geq \overline{\varphi}$, then GOTO Step 1.
Step 3.1	The scenario solutions $u_\xi^{opt}, \xi = 1, \ldots, S$, gained by solving the Lagrangean dual are identical[b], *i.e.* $u_1^{opt} = \ldots = u_S^{opt}$. If further $\mathcal{Q}_{\mathbb{E}}(u_1^{opt}) + \rho \cdot \mathcal{Q}_{\mathbb{P}_\eta}(u_1^{opt}) < \overline{\varphi}$, then $\overline{\varphi} := \mathcal{Q}_{\mathbb{E}}(u_1^{opt}) + \rho \cdot \mathcal{Q}_{\mathbb{P}_\eta}(u_1^{opt})$ and $\overline{u} := u_1^{opt}$. Delete from $\mathbf{P}$ all problems P' with $\varphi_{\mathrm{LD}}(P') \geq \overline{\varphi}$; GOTO Step 1.
Step 3.2	If the scenario solutions $u_\xi^{opt}, \xi = 1, \ldots, S$ differ, then run a feasibility heuristic. If its outcome $\widehat{u}$ is feasible and $\mathcal{Q}_{\mathbb{E}}(\widehat{u}) + \rho \cdot \mathcal{Q}_{\mathbb{P}_\eta}(\widehat{u}) < \overline{\varphi}$, then $\overline{\varphi} := Q_{\mathbb{E}}(\widehat{u})$ and $\overline{u} := \widehat{u}$. Delete from $\mathbf{P}$ all problems P' with $\varphi_{\mathrm{LD}}(P') \geq \overline{\varphi}$; GOTO Step 4.
Step 4.	Select a component $(\mathbf{u}_I)_i$ of $\mathbf{u}_I$ and add two new problems to $\mathbf{P}$ which arise from P by adding the constraints $(\mathbf{u}_I)_i = 0$ and $(\mathbf{u}_I)_i = 1$, respectively; GOTO Step 1.

[a] Here, the Lagrangean lower bound means the lower bound that is obtained by Lagrangean relaxation of the nonanticipativity constraint (4.12).

[b] This implies that u_1^{opt} is feasible for the SDP relaxation of (4.1).

5. Computational results

Some first computational tests have been performed using MATLAB R2013a with MILP's and QP's solved by CPLEX Studio 12.51 (connector to MATLAB). For the arising SDPs we have employed SeDuMi 1.3. A simple hardware set has been used, namely consisting of an Intel(R) Core(TM) i7-2640M CPU @ 2.80GHz 2.80GHz processor with 4 GB of RAM running under Windows 7 Professional.

5.1. Power system data

To exhibit the efficiency of our proposed decomposition approach, it is tested with the well-known 14-bus IEEE network, the 30-bus IEEE network, the 39-bus New England test system, and the 57-bus IEEE network.

We adopt a daily planning horizon subdivided into 24 equidistant time intervals. Starting from the IEEE load data as a reference we first developed a practically relevant deterministic load profile for an individual day.

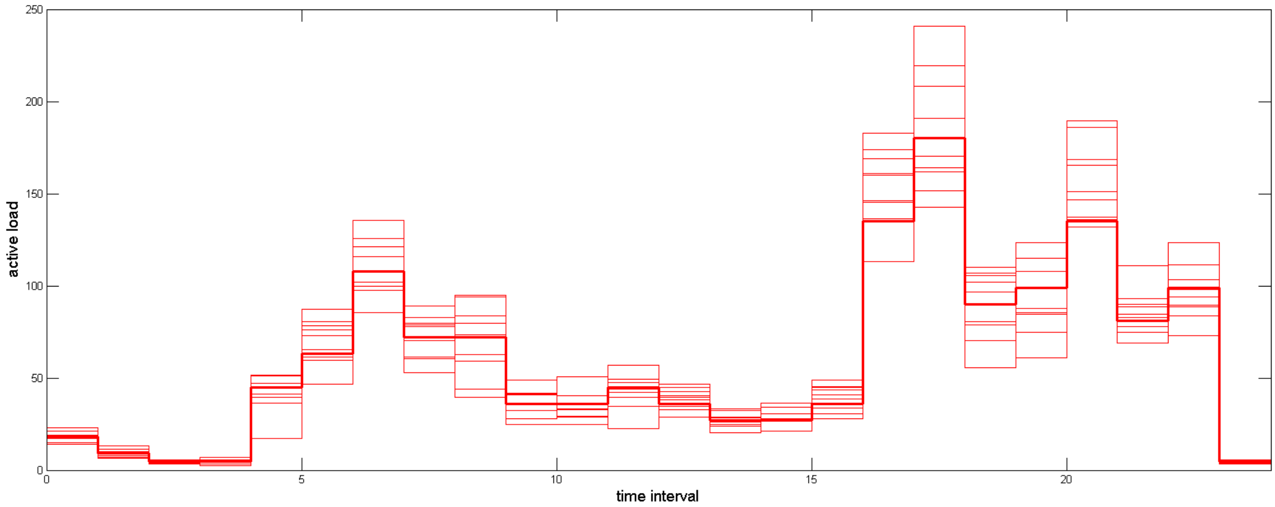

FIGURE 1. Active load scenarios.

In order to generate scenarios for the required finite discrete probability distribution, the preassigned daily load profile (p, q) has been perturbed as follows: with S denoting the number of realizations, scenarios

$$(p_\omega, q_\omega) = (p, q) + n_\omega \cdot (p, q),\ \omega = 1, \dots, S \tag{5.1}$$

are formed, where n_ω is a random number sampled from the standard normal distribution. Figure 1 displays scenarios obtained in this way.

In our tests, we have started with two basic deterministic network infrastructures: power systems with purely thermal generation and with pumped-storage plants added.

Stochastic expansions of these models were obtained by introducing random load values and assigning roles to variables making them members of the first and second stages, respectively. In this way, decisions in the first stage comprise on/off switching for the coal fired blocks. Variables in the second stage represent output levels of the coal fired units, switching decisions and output levels at gas turbines, and, if present pumping and generation modes in the pumped storage plants.

5.2. Preliminary computational results

In all numerical tests, the relaxed rank-one conditions are successfully recovered by the algorithm of Lavaei and Low (*cf*. Appendix B), implying that these unit commitment problems are solved to the specified optimality gaps.

For the considered IEEE networks, our implemented OPF solver, *i.e.* the solver that is used for solving the subproblems (3.3), provides the same solution as the SDP based OPF solvers by Madani, Asharphijuo, and Lavaei [29] and by Molzahn et al. [34,35].

5.2.1. *Computational results for the deterministic Benders approach*

Table 1 reports our computational results for the deterministic Benders approach from Section 3.1. Here, the deterministic load data equals the expected value of the random data, *i.e.* $(p, q) = \mathbb{E}_\omega[(p_\omega, q_\omega)]$. The stopping criterion has been set to 10^{-2}, *i.e.* in all of the tests listed below the overall (mixed-integer) SDP is solved to less then a 1% optimality gap.

Starting from left, the following information is “encoded” in the columns of Table 1: model number, IEEE benchmark network, numbers of generators and hydro units, as well as resulting numbers of continuous and

TABLE 1. Computational results for the deterministic Benders approach.

Model	Network	Generator	Hydro Units	Variables (Binaries)	Constr.	Iter.	CPU	Gen. Cuts	Costs
D1	IEEE 14	5	0	11 496 (120)	3319	14	132.68 s	336	278 328.52
D2	IEEE 14	5	1	11 569 (120)	3465	21	184.27 s	504	277 770.19
D3	IEEE 14	5	2	11 642 (120)	3611	35	329.35 s	840	277 814.05
D4	IEEE 30	6	0	46 656 (144)	5250	11	204.15 s	264	17 597.26
D5	IEEE 30	6	1	46 729 (144)	5396	18	324.57 s	432	17 597.61
D6	IEEE 30	6	2	46 802 (144)	5442	20	372.09 s	480	17 604.62
D7	NE 39	10	0	77 976 (240)	7694	4	94.71 s	96	918 751.08
D8	NE 39	10	1	78 049 (240)	7840	1	22.73 s	24	557 948.54
D9	NE 39	10	2	78 122 (240)	7986	1	22.18 s	24	557 948.54
D10	IEEE 57	7	0	161 568 (168)	8237	19	944.08 s	456	1 117 780.35
D11	IEEE 57	7	1	161 641 (168)	8383	11	521.80 s	264	988 369.13
D12	IEEE 57	7	2	161 714 (168)	8529	13	570.24 s	312	988 301.96

integers (binary) variables, and constraints. The remaining four columns display the numbers of iterations, CPU time, cuts generated, and optimal costs.

Our Benders decomposition approach has been significantly improved by adding the constraints

$$\sum\nolimits_{k\in\mathcal{N}} p^t_{D_k} \le \sum\nolimits_{k\in\mathcal{G}} \mathbf{u}^t_k \cdot P^{\max}_k + \sum\nolimits_{h\in\mathcal{H}} (\mathbf{p}^t_h - \mathbf{w}^t_h), \quad t = 1,\dots,T, \tag{5.2}$$

to the master problem (3.2). For each time interval, these requirements guarantee that there is a sufficient number of on-line thermal generators to produce the required active power. Doing so, a lot of switching decisions become inferior, already in the master problem (3.2), such that they do not have to be cut off by solving the (computationally expensive) semidefinite subproblems (3.10).

5.2.2. Computational results for the stochastic dual decomposition algorithm

Tables 2 and 3 exhibit numerical results for the dual decomposition algorithm from Section 4.2. These tables are structured as follows: the left most seven columns corresponds to those of Table 1 with the exception that now scenarios have to listed (in column three). Colmuns eight to ten correspond to the number of SeDuMi calls, the CPU time (with SeDuMi share in brackets), and the optimal costs. For each stochastic program listed the overall (mixed-integer) semidefinite relaxation has been solved to a 1% duality gap.

The execution of our dual decomposition algorithm requires to solve almost identical scenario specific unit commitment problems again and again. In order to improve computational performance of our method, we stored the generated Benders cuts and reused them in each iteration. This has decreased CPU time considerably. Having solved a scenario specific unit commitment problem once, which takes minutes (*cf.* Tab. 1), it will be solved in a few seconds in future iterations. Cut deletion, clearly, would have been an option. This has not been pursued, since the share of solving master problems (including all previous cuts) amounted to a mere 3% or less of the total computation time.

The above tables show that most time (on average about 95%) is spent solving the arising OPF problems by SeDuMi. Essentially, this is due to the vast number of cuts that have to be generated to solve the scenario problems.

Observe that the total times for solving the risk averse models are approximately the same as for the risk neutral ones except for the test instances containing (second-stage) gas turbines where the total time has roughly doubled. This is due to the complexity of the single-scenario unit commitment problems. We merely add to each of them just one big-M constraint and penalize its violation, such that, in this case, (4.7) is not that much harder to solve than (4.11).

Table 2. Computational results for the risk neutral model.

Model	Network	Scenarios	Generator	Hydro	Var.	Constr.	SeDuMi Calls	CPU (SeDuMi)	Costs
Networks exclusive of gas turbines (No second-stage integers)									
E1	IEEE 14	10	5(0)	0	113 760	30 040	5743	2128 s (2083 s)	292 565.64
E2	IEEE 14	10	5(0)	1	114 610	31 500	10 276	3853 s (3718 s)	293 056.19
E3	IEEE 14	10	5(0)	2	115 340	32 960	17 285	6545 s (6128 s)	293 156.36
E4	IEEE 14	50	5(0)	0	568 920	148 800	20 730	7668 s (7523 s)	314 167.29
E5	IEEE 14	50	5(0)	1	572 570	156 100	49 298	16 119 s (15526 s)	314 174.97
E6	IEEE 14	50	5(0)	2	576 220	163 400	59 774	22 530 s (21373 s)	314 279.54
E7	IEEE 30	10	6(0)	0	465 264	48 720	7668	5705 s (5661 s)	18079.13
E8	IEEE 30	10	6(0)	1	465 994	50 180	9028	6066 s (5938 s)	17821.18
E9	IEEE 30	10	6(0)	2	466 724	50 640	4323	3276 s (3229 s)	17 822.27
E10	IEEE 30	50	6(0)	0	2 325 744	241 920	15 252	11 235 s (11127 s)	18 643.85
E11	IEEE 30	50	6(0)	1	2 329 394	249 220	16 623	11 247 s (11057 s)	18 645.24
E12	IEEE 30	50	6(0)	2	2 333 044	251 520	16 840	11 464 s (11257 s)	18 645.03
Networks including gas turbines (second-stage integers).									
Model	Network	Scenarios	Generator	Hydro	Var.	Constr.	SeDuMi Calls	CPU (SeDuMi)	Costs
E13	IEEE 14	10	5(3)	0	114 408	30 040	6350	2135 s (2083 s)	258 080.00
E14	IEEE 14	10	5(3)	1	115 258	31 500	6402	2144 s (2084 s)	247 850.00
E15	IEEE 14	10	5(3)	2	115 988	32 960	9772	3572 s (3424 s)	240 280.00
E16	IEEE 14	50	5(3)	0	572 448	148 800	16 864	5992 s (5905 s)	281 230.00
E17	IEEE 14	50	5(3)	1	576 098	156 100	26 261	9815 s (9556 s)	278 960.00
E18	IEEE 14	50	5(3)	2	579 748	163 400	56 706	18 353 s (17 589 s)	271 023.66
E19	IEEE 30	10	6(4)	0	466 128	48 720	9969	6955 s (6765 s)	18 077.00
E20	IEEE 30	10	6(4)	1	466 858	50 180	3771	2770 s (2736 s)	17 816.00
E21	IEEE 30	10	6(4)	2	467 588	50 640	3698	2818 s (2783 s)	17 816.00
E22	IEEE 30	50	6(4)	0	2 330 448	241 920	6926	4762 s (4736 s)	17 896.00
E23	IEEE 30	50	6(4)	1	2 334 098	249 220	14 126	10 086 s (9934 s)	18 606.00
E24	IEEE 30	50	6(4)	2	2 337 748	251 520	14 368	10 681 s (10 509 s)	18 607.00

Table 3. Computational results for the risk averse model.

Model	Network	Scenarios	Generator	Hydro	Var.	Constr.	SeDuMi Calls	CPU (SeDuMi)	Costs
EP1	IEEE 14	10	5(0)	0	113 770	30 050	5292	1462 s (1394 s)	296 117.52
EP2	IEEE 14	10	5(3)	0	114 418	30 050	13618	3237 s (3136 s)	287 906.38
EP3	IEEE 14	50	5(0)	0	568 970	148 850	20 730	5399 s (5212 s)	319 247.29
EP4	IEEE 14	50	5(3)	0	572 498	148 850	56 654	12 265 s (120 00 s)	281 702.41
EP5	IEEE 30	10	6(0)	0	465 274	48 730	3448	1935 s (1908 s)	18 494.48
EP6	IEEE 30	10	6(4)	0	466 138	48 730	23 132	10 773 s (10 609 s)	18 150.40
EP7	IEEE 30	50	6(0)	0	2 325 794	241 970	15 252	6369 s (6294 s)	19 167.85
EP8	IEEE 30	50	6(4)	0	2 330 498	241 970	21 251	12 733 s (12 563 s)	19 052.32

The value of the stochastic solution. When exploring the usefulness of the stochastic programming approach in (4.1) achieving second stage feasibility turned out difficult. In fact, in our instances E1–E6, none of the optimal solutions to the averaged models has been feasible for all individual scenarios. This is reflected by Table 4, where column 2 displays infeasibility (regarding both numbers of scenarios and accumulated probability).

Furthermore, column 2 of Table 4 reveals that deterministic solutions could be very misleading as they may be infeasible in a considerable number of cases. By contrast, the solutions of the stochastic programs E1–E6

TABLE 4. Evaluation of the solutions of the expected value models.

Model	Infeasibility	Wait-and-See	EVPI	Gap
E1	3 (19.1%)	249 098.16	43 467.48	17.45%
E2	4 (24.8%)	240 916.12	52 140.07	21.64%
E3	3 (19.1%)	234 096.86	59 059.5	25.22%
E4	9 (15.0%)	273 349.90	40 817.39	14.93%
E5	19 (36.7%)	271 728.20	42 446.77	15.62%
E6	9 (15.0%)	271 111.36	43 168.18	15.92%

are quite robust with respect to changes in the data. Moreover, calculating the gap between the wait-and-see (WS) solution[4] (the WS solution for E1–E6 is displayed in column 3 of Tab. 4) and the optimal value of E1–E6, it turned out, that, after the fact, the stochastic solutions are not optimal, but fairly good. These gaps and in addition the expected value of perfect information[5] (EVPI) are reported in column 5 and 4 of Table 4, respectively.

The benefit of the risk averse approach. In order to push the effect of the risk averse approach, compared to the risk neutral one, we have considered instances whose scenarios were less power consuming. To this end, we just divided the scenario load profiles (from the results above) by four. Than much more first-stage solutions become feasible (more generators may be in off-state).

The bar charts in Figure 2 illustrate the impact of different, risk neutral and risk averse, stochastic criteria on the shapes of the optimal solutions. It displays our results for the 14-bus IEEE network with two (first-stage) coal fired blocks and three (second-stage) gas turbines, as well as 50 scenarios. Here, each individual column symbolizes one of the 50 scenarios where the particular height refers to the (single-scenario) objective value and the width to the corresponding probability.

The expectation based model minimizes the sum over all single-scenario problems (where each single-scenario problem is weighted by its probability). This implies that scenarios with high costs may be compensated by scenarios with lower costs. In doing so, variability is neglected at all. Hence, this may result in a solution whose associated random variable highly fluctuates and takes unfavorable values "too often". These drawbacks are illustrated in Figure 2. For the considered test instance, the solution of the expectation based model varies strongly and incurs costs higher than 91 000 in five of the fifty single-scenarios (corresponding probability is 14.1%). The objective value is 49 253.70. Despite the solution of the excess probability mean-risk model causing objective costs of 66 073.50 in the expected value model, its single-scenario objectives do not vary that much and none of them exceeds the threshold $\eta = 91\,000$.

6. CONCLUDING REMARKS

In the present paper we have brought together unit commitment in AC transmission systems with risk averse stochastic optimization employing semidefinite programming. The latter recently was boosted by rank relaxations of semidefinite programs that lead to (globally) solvable optimal power flow problems. More specifically, relaxations of rank-one conditions could be recuperated for certain classes of electricity networks including among others popular IEEE OPF test instances.

Our focus has been to explore the potential of the recent findings in power flow when addressed under data uncertainty. The computations in the present paper confirm in principal that such a model extension remains computationally feasible provided proper decomposition techniques are integrated into the algorithmic treatment.

[4] The expected value $\mathbb{E}\left[\min_{\mathbf{X}\in\mathcal{X}} f(\mathbf{X},\omega)\right]$ is called wait-and-see solution.

[5] The value $\min_{\mathbf{X}\in\mathcal{X}} \mathbb{E}\left[f(\mathbf{X},\omega)\right] - \mathbb{E}\left[\min_{\mathbf{X}\in\mathcal{X}} f(\mathbf{X},\omega)\right]$ is called the expected value of perfect information.

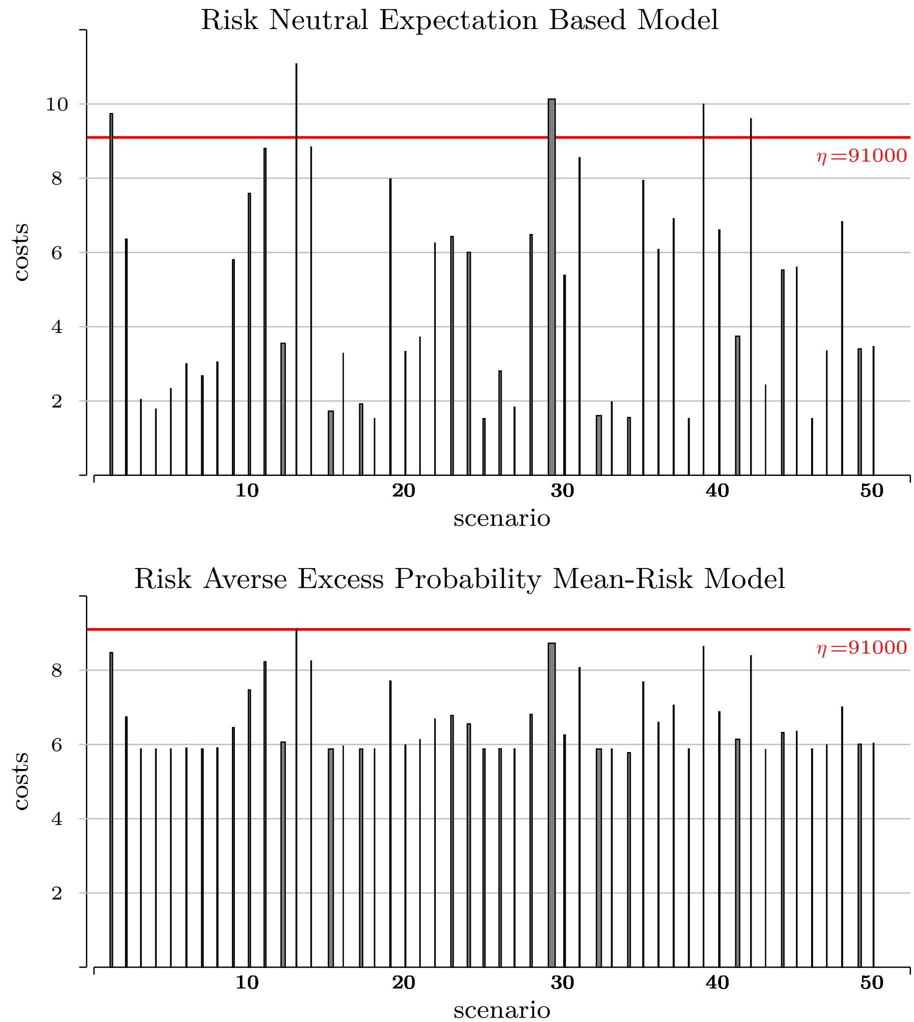

FIGURE 2. IEEE 14-bus system (2 coal fired blocks, 3 gas turbines) and 50 scenarios: Objective values for each of the single-scenario problems.

In this context, we refer to very recent publications on sufficient conditions for the semidefinite approximation of OPF finally enabling solution of the original problem, see the doctoral thesis [33] and survey paper [32].

Appendix A. Equivalent rank constrained SDP representation

The inclusion of the AC power flow constraints introduced in Section 2.1 leads to a mixed-integer nonlinear program. Whenever these conditions enter into an optimization problem, its feasible set becomes nonconvex and the problem itself NP-hard [24]. Nevertheless, following Lavaei and Low, the introduced network conditions (2.8)–(2.17) and thus our whole unit commitment problem may be solved by considering a (rank constraint) mixed-integer linear semidefinite program. To this, it takes the introduction of some parameters: For lines $(l, m) \in \mathcal{L}$, the complex parameter $y_{lm} := g_{lm} + jb_{lm}$ is referred to as the admittance between the nodes l and m. This definition is extended to all $l \neq m$ by putting y_{lm} equal to zero, whenever bus l and m are not directly linked. The parameter y_{kk} denotes the admittance-to-ground at $k \in \mathcal{N}$, it is defined as the sum over all connected line admittances added by the line shunt admittances. In addition, it is essential to introduce the admittance matrix $Y = G + jB \in \mathbb{C}^{n\times n}$ defined by $y_{ll} + \sum_{m\in N(l)} y_{lm}$ for diagonal elements and $-y_{lm}$ otherwise. Moreover, the following parameters for every $k \in \mathcal{N}$ and $(l, m) \in \mathcal{L}$ are required:

$$\widetilde{Y}_k := e_k e_k^T Y, \qquad \widetilde{Y}_{lm} := (b^0_{lm} + y_{lm}) e_l e_l^T - (y_{lm}) e_l e_m^T,$$
$$M_k := \begin{bmatrix} e_k e_k^T & 0 \\ 0 & e_k e_k^T \end{bmatrix}, \qquad M_{lm} := \begin{bmatrix} (e_l - e_m)(e_l - e_m)^T & 0 \\ 0 & (e_l - e_m)(e_l - e_m)^T \end{bmatrix},$$

where Y is the net corresponding admittance matrix, b^0_{lm} are the given shunts, y_{lm} are the admittances, and $e_1, \ldots, e_n$ are the standard basis vectors in $\mathbb{R}^n$. With these agreements, we besides need the subsequent auxiliary matrices:

$$Y_k := \frac{1}{2}\begin{bmatrix} \mathrm{Re}\left(\widetilde{Y}_k + \widetilde{Y}_k^T\right) & \mathrm{Im}\left(\widetilde{Y}_k^T - \widetilde{Y}_k\right) \\ \mathrm{Im}\left(\widetilde{Y}_k - \widetilde{Y}_k^T\right) & \mathrm{Re}\left(\widetilde{Y}_k + \widetilde{Y}_k^T\right) \end{bmatrix},$$
$$\overline{Y}_k := -\frac{1}{2}\begin{bmatrix} \mathrm{Im}\left(\widetilde{Y}_k + \widetilde{Y}_k^T\right) & \mathrm{Re}\left(\widetilde{Y}_k - \widetilde{Y}_k^T\right) \\ \mathrm{Re}\left(\widetilde{Y}_k^T - \widetilde{Y}_k\right) & \mathrm{Im}\left(\widetilde{Y}_k + \widetilde{Y}_k^T\right) \end{bmatrix},$$
$$Y_{lm} := \frac{1}{2}\begin{bmatrix} \mathrm{Re}\left(\widetilde{Y}_{lm} + \widetilde{Y}_{lm}^T\right) & \mathrm{Im}\left(\widetilde{Y}_{lm}^T - \widetilde{Y}_{lm}\right) \\ \mathrm{Im}\left(\widetilde{Y}_{lm} - \widetilde{Y}_{lm}^T\right) & \mathrm{Re}\left(\widetilde{Y}_{lm} + \widetilde{Y}_{lm}^T\right) \end{bmatrix},$$
$$\overline{Y}_{lm} := -\frac{1}{2}\begin{bmatrix} \mathrm{Im}\left(\widetilde{Y}_{lm} + \widetilde{Y}_{lm}^T\right) & \mathrm{Re}\left(\widetilde{Y}_{lm} - \widetilde{Y}_{lm}^T\right) \\ \mathrm{Re}\left(\widetilde{Y}_{lm}^T - \widetilde{Y}_{lm}\right) & \mathrm{Im}\left(\widetilde{Y}_{lm} + \widetilde{Y}_{lm}^T\right) \end{bmatrix}.$$

Further, define the real voltage vector $\mathbf{X} := [\mathrm{Re}(\mathbf{V})^T, \mathrm{Im}(\mathbf{V})^T]$ and for $k \in \mathcal{N}$, the injected net active and reactive powers $\mathbf{p}_{k,\mathrm{inj}}$ and $\mathbf{q}_{k,\mathrm{inj}}$, respectively. Here, the latter are defined by $\mathbf{p}_{k,\mathrm{inj}} := \mathbf{p}_{G_k} - p_{D_k}$, $\mathbf{q}_{k,\mathrm{inj}} := \mathbf{q}_{G_k} - q_{D_k}$ for all $k \in \mathcal{G}$ as well as $\mathbf{p}_{k,\mathrm{inj}} := -p_{D_k}$, $\mathbf{q}_{k,\mathrm{inj}} := -q_{D_k}$ for all $k \in \mathcal{N}\backslash\mathcal{G}$. Then, with these notations, the following equations are valid (*cf.* [24]):

$$\begin{aligned} \mathbf{p}_{k,\mathrm{inj}} &:= Y_k \bullet \left(\mathbf{X}\mathbf{X}^T\right), \quad \mathbf{q}_{k,\mathrm{inj}} := \overline{Y}_k \bullet \left(\mathbf{X}\mathbf{X}^T\right) \\ |\mathbf{V}_k|^2 &:= M_k \bullet \left(\mathbf{X}\mathbf{X}^T\right), \quad |\mathbf{V}_l - \mathbf{V}_m|^2 := M_{lm} \bullet \left(\mathbf{X}\mathbf{X}^T\right), \\ \mathbf{p}_{lm} &:= \mathbf{Y}_{lm} \bullet \left(\mathbf{X}\mathbf{X}^T\right), \quad |\mathbf{S}_{lm}|^2 := \left(Y_{lm} \bullet \left(\mathbf{X}\mathbf{X}^T\right)\right)^2 + \left(\overline{Y}_{lm} \bullet \left(\mathbf{X}\mathbf{X}^T\right)\right)^2. \end{aligned} \tag{A.1}$$

If now all generator on/off decisions are fixed and in addition the active and reactive power outputs/consumptions, respectively, at all pumped-storage plants are given in advance, then we obtain the following arithmetic reformulation (in terms of $\mathbf{X}$) of the production bounds (2.2) and power flow constraints (2.8)–(2.17):

$$P_k^{\min} \le Y_k \bullet \left(\mathbf{X}\mathbf{X}^T\right) + p_{D_k} \le P_k^{\max}, \qquad \forall k \in \mathcal{N}, \tag{A.2}$$
$$Q_k^{\min} \le \overline{Y}_k \bullet \left(\mathbf{X}\mathbf{X}^T\right) + q_{D_k} \le Q_k^{\max}, \qquad \forall k \in \mathcal{N}, \tag{A.3}$$
$$(V_k^{\min})^2 \le M_k \bullet \left(\mathbf{X}\mathbf{X}^T\right) \le (V_k^{\max})^2, \qquad \forall k \in \mathcal{N} \tag{A.4}$$
$$\left(\mathbf{Y}_{lm} \bullet \left(\mathbf{X}\mathbf{X}^T\right)\right)^2 + \left(\overline{\mathbf{Y}}_{lm} \bullet \left(\mathbf{X}\mathbf{X}^T\right)\right)^2 \le (S_{lm}^{\max})^2, \qquad \forall (l,m) \in \mathcal{L}, \tag{A.5}$$
$$Y_{lm} \bullet \left(\mathbf{X}\mathbf{X}^T\right) \le P_{lm}^{\max}, \qquad \forall (l,m) \in \mathcal{L}, \tag{A.6}$$
$$M_{lm} \bullet \left(\mathbf{X}\mathbf{X}^T\right) \le (\varDelta V_{lm}^{\max})^2, \qquad \forall (l,m) \in \mathcal{L}, \tag{A.7}$$

where we have to extend the definition of $P_k^{\min}$, $P_k^{\max}$, $Q_k^{\min}$, $Q_k^{\max}$ from $k \in \mathcal{G}$ to every network bus $k \in \mathcal{N}$, by putting $P_k^{\min} = P_k^{\max} = Q_k^{\min} = Q_k^{\max} = 0$ if $k \in \mathcal{N}\backslash\mathcal{G}$. Further, we have to adjust the active and reactive load/infeed at pumped-storage buses. Here, observe that all conditions in (A.2)–(A.7) become linear in $\mathbf{XX}^T$, except the quadratic inequalities in (A.5). However, with the aid of Schur's complement and by introducing for all lines $(l,m) \in \mathcal{L}$ an artificial matrix variable $\mathbf{Z}_{lm} \in \mathcal{S}^3$, these non-quadratic inequalities may be represented by the following linear matrix equations:

$$\begin{aligned}
&\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \bullet \mathbf{Z}_{lm} = (S_{lm}^{\max})^2, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \bullet \mathbf{Z}_{lm} = 1, \\
&\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \bullet \mathbf{Z}_{lm} = 1, \qquad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \bullet \mathbf{Z}_{lm} = 0, \\
&\begin{bmatrix} 0 & 1/2 & 0 \\ 1/2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \bullet \mathbf{Z}_{lm} + Y_{lm} \bullet \mathbf{XX}^T = 0, \\
&\begin{bmatrix} 0 & 0 & 1/2 \\ 0 & 0 & 0 \\ 1/2 & 0 & 0 \end{bmatrix} \bullet \mathbf{Z}_{lm} + \overline{Y}_{lm} \bullet \mathbf{XX}^T = 0, \quad \mathbf{Z}_{lm} \succeq 0.
\end{aligned} \tag{A.8}$$

Similarly, adopting the reformulation (A.1) of net active power injected at nonrenewable generators to the quadratic cost function (2.1), while simultaneously introducing for each of the nonrenewable generators $k \in \mathcal{G}\backslash H$, the matrix variable $\mathbf{A}_k \in \mathcal{S}^2$, it can be transformed into linear shape as well:

$$\begin{aligned}
&\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \bullet \mathbf{A}_k + c_{k1} Y_k \bullet \mathbf{XX}^T - \mathbf{a}_k = -a_k, \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \bullet \mathbf{A}_k = 1, \\
&\begin{bmatrix} 0 & 1/2 \\ 1/2 & 0 \end{bmatrix} \bullet \mathbf{A}_k + \sqrt{c_{k2}} Y_k \bullet \mathbf{XX}^T = -b_k, \quad \mathbf{A}_k \succeq 0,
\end{aligned} \tag{A.9}$$

where $\mathbf{a}_k$ denotes the operating costs occurring at generator k, *i.e.* the objective is $\sum_{k \in \mathcal{G}\backslash H} (\mathbf{a}_k + c_{k0} \cdot \mathbf{r}_k)$. Moreover, in the same way, the injected active power at pumped-storage units $h \in H$ may be depict by $\mathbf{p}_h - \mathbf{w}_h := Y_h \bullet \mathbf{XX}^T + p_{D_h}$, such that the box constraints

$$-W_h^{\max} \leq Y_h \bullet \mathbf{XX}^T + p_{D_h} \leq P_h^{\max}, \text{ and } -\overline{W}_h^{\max} \leq \overline{Y}_h \bullet \mathbf{XX}^T + q_{D_h} \leq \overline{P}_h^{\max},$$

at these particular network nodes have to be satisfied (*cf.* (2.3)). Hence, since that a given matrix $\mathbf{W}$ can be written as $\mathbf{XX}^T$ for some (nonzero) vector $\mathbf{X}$ if and only if $\mathbf{W}$ is both symmetric positive semidefinite and rank 1, we finally receive an equivalent rank constrained linear SDP reformulation. Thus, by applying this substitution to our initial unit commitment problem from Section 2.1, it is indeed equivalent to a mixed-integer rank constrained SDP.

Appendix B. A strategy for solving OPF problems

If we apply the reformulation described in Appendix A to the unit commitment problem introduced in Section 2.1 and further fix all generator switching decisions as well as the active and reactive power outputs/consumptions, respectively, at all pumped-storage plants, then, for $t \in \{1, \ldots, T\}$, the dual to the resulting decomposed OPF is given by (*cf.* [24]):

$$\left.\begin{aligned}\max_{\boldsymbol{\lambda}^t\geq 0}\Big\{h(\boldsymbol{\lambda}^t,\boldsymbol{r}^t) : A(\boldsymbol{\lambda}^t,\boldsymbol{r}^t) \succeq 0,&\\ \begin{bmatrix} 1 & \boldsymbol{r}^t_{k,1}\\ \boldsymbol{r}^t_{k,1} & \boldsymbol{r}^t_{k,2}\end{bmatrix} \succeq 0, \qquad \forall k \in \mathcal{G}\backslash H&\\ \begin{bmatrix} \boldsymbol{r}^t_{lm,1} & \boldsymbol{r}^t_{lm,2} & \boldsymbol{r}^t_{lm,3}\\ \boldsymbol{r}^t_{lm,2} & \boldsymbol{r}^t_{lm,4} & \boldsymbol{r}^t_{lm,5}\\ \boldsymbol{r}^t_{lm,3} & \boldsymbol{r}^t_{lm,5} & \boldsymbol{r}^t_{lm,6}\end{bmatrix} \succeq 0, \forall (l,m) \in L&\end{aligned}\right\}, \tag{B.1}$$

with

$$\begin{aligned} A(\boldsymbol{\lambda}^t,\boldsymbol{r}^t) := \;& \sum\nolimits_{k\in N} \left(\overline{\boldsymbol{\lambda}}^t_k - \underline{\boldsymbol{\lambda}}^t_k\right) Y_k + \left(\overline{\boldsymbol{\gamma}}^t_k - \underline{\boldsymbol{\gamma}}^t_k\right) \overline{Y}_k + \left(\overline{\boldsymbol{\mu}}^t_k - \underline{\boldsymbol{\mu}}^t_k\right) M_k\\ &+ \sum\nolimits_{k\in\mathcal{G}\backslash H} (c_{k1} + 2\sqrt{c_{k2}}\boldsymbol{r}^t_{k,1})Y_k\\ &+ \sum\nolimits_{(l,m)\in L} (\boldsymbol{\lambda}^t_{lm} + 2\boldsymbol{r}^t_{lm,2})Y_{lm} + 2\boldsymbol{r}^t_{lm,3}\overline{Y}_{lm} + \boldsymbol{\mu}^t_{lm}M_{lm}, \end{aligned} \tag{B.2}$$

$$\begin{aligned} h(\boldsymbol{\lambda}^t,\boldsymbol{r}^t) := \;& \sum\nolimits_{k\in N} \underline{\boldsymbol{\lambda}}^t_k P^{\min}_k - \overline{\boldsymbol{\lambda}}^t_k P^{\max}_k + \left(\overline{\boldsymbol{\lambda}}^t_k - \underline{\boldsymbol{\lambda}}^t_k\right) p^t_{D_k}\\ &+ \sum\nolimits_{k\in N} \underline{\boldsymbol{\gamma}}^t_k Q^{\min}_k - \overline{\boldsymbol{\gamma}}^t_k Q^{\max}_k + \left(\overline{\boldsymbol{\gamma}}^t_k - \underline{\boldsymbol{\gamma}}^t_k\right) q^t_{D_k}\\ &+ \sum\nolimits_{k\in N} \underline{\boldsymbol{\mu}}^t_k \left(V^{\min}_k\right)^2 - \overline{\boldsymbol{\mu}}^t_k \left(V^{\max}_k\right)^2\\ &- \sum\nolimits_{(l,m)\in L} \boldsymbol{\lambda}^t_{lm} P^{\max}_{lm} + \boldsymbol{\mu}^t_{lm} \left(\Delta V^{\max}_{lm}\right)^2 + \boldsymbol{r}^t_{lm,1} \left(S^{\max}_{lm}\right)^2 + \boldsymbol{r}^t_{lm,4} + \boldsymbol{r}^t_{lm,6}, \end{aligned}$$

where

$$\boldsymbol{\lambda}^t = (\overline{\boldsymbol{\lambda}}^t_k, \underline{\boldsymbol{\lambda}}^t_k, \overline{\boldsymbol{\gamma}}^t_k, \underline{\boldsymbol{\gamma}}^t_k, \overline{\boldsymbol{\mu}}^t_k, \underline{\boldsymbol{\mu}}^t_k, \boldsymbol{\lambda}^t_{lm}, \boldsymbol{\mu}^t_{lm}),$$

and

$$\boldsymbol{r}^t = (\boldsymbol{r}^t_{k,1}, \boldsymbol{r}^t_{k,2}, \boldsymbol{r}^t_{lm,1}, \boldsymbol{r}^t_{lm,2}, \boldsymbol{r}^t_{lm,3}, \boldsymbol{r}^t_{lm,4}, \boldsymbol{r}^t_{lm,5}, \boldsymbol{r}^t_{lm,6}).$$

Now, provided (B.1) is solvable, let us denote by $(\lambda^t_{\text{opt}}, r^t_{\text{opt}})$ any optimal solution to this program. Assume further that X^t_{opt} is primal optimal (*i.e.*, it solves the dual to (B.1)) and that Slater's condition is satisfied. Then, due to strong duality, $A(\lambda^t_{\text{opt}}, r^t_{\text{opt}}) \bullet X^t_{opt} = 0$. This equation is valid if and only if the product of the symmetric and positive semidefinite matrices $A(\lambda^t_{\text{opt}}, r^t_{\text{opt}})$ and X^t_{opt} vanishes. Hence, writing the symmetric matrix X^t_{opt} by using its eigenvalue decomposition $P^t\Lambda^t P^{t\,T} = \sum_{i=1}^{2N} \lambda^t_i p^t_i p^{t\,T}_i$, the following equations have to hold true:

$$A(\lambda^t_{\text{opt}}, r^t_{\text{opt}})p^t_i = 0, \quad \text{for those } i \in \{1,\dots,2N\} \text{ for which } \lambda_i \neq 0. \tag{B.3}$$

This means that all of the concerned orthogonal eigenvectors (eigenvectors to nonzero eigenvalues of X^t_{opt}) must belong to the kernel of $A(\lambda^t_{\text{opt}}, r^t_{\text{opt}})$. If now the latter were of dimension one, the primal would have a rank-one solution. Hence, there were a zero duality gap between the OPF and its SDP relaxation. The same result is obtained when the kernel of $A(\lambda^t_{\text{opt}}, r^t_{\text{opt}})$ has dimension less than or equal to 2 (*cf.* [24]). Indeed, in view of Appendix A, the matrix $A(\lambda^t_{\text{opt}}, r^t_{\text{opt}})$ as a weighted sum of the matrices $Y_k, \overline{Y}_k, M_k, Y_{lm}, \overline{Y}_{lm}, M_{lm}$ has the following block structure:

$$A(\lambda^t_{\text{opt}}, r^t_{\text{opt}}) = \begin{bmatrix} \overline{A}(\lambda^t_{\text{opt}}, r^t_{\text{opt}}) & B(\lambda^t_{\text{opt}}, r^t_{\text{opt}})\\ -B(\lambda^t_{\text{opt}}, r^t_{\text{opt}}) & \overline{A}(\lambda^t_{\text{opt}}, r^t_{\text{opt}})\end{bmatrix}.$$

This implies: if the kernel of $A(\lambda^t_{\text{opt}}, r^t_{\text{opt}})$ includes

$$p^t = \left[p^{t\,T}_1 \; p^{t\,T}_2\right]^T,$$

then it also includes

$$\left[-p_2^{t\,T}\ p_1^{t\,T}\right]^T.$$

As these two vectors are orthogonal, they must be the two eigenvectors to the zero eigenvalue of $A(\lambda^t_{\text{opt}}, r^t_{\text{opt}})$. Therefore, for primal optimal X^t_{opt}, the following holds

$$X^t_{opt} = \lambda^t_1 \begin{bmatrix} p^t_1 \\ p^t_2 \end{bmatrix} \left[p_1^{t\,T}\ p_2^{t\,T}\right] + \lambda^t_2 \begin{bmatrix} -p^t_2 \\ p^t_1 \end{bmatrix} \left[-p_2^{t\,T}\ p_1^{t\,T}\right].$$

Further, due to the fact that the trace of a skew-symmetric and symmetric matrix is equal to zero, it could be observed that

$$\begin{aligned} \begin{bmatrix} A & B \\ -B & A \end{bmatrix} \bullet \begin{bmatrix} p^t_1 \\ p^t_2 \end{bmatrix} \left[p_1^{t\,T}\ p_2^{t\,T}\right] &= A \bullet p^t_1 p_1^{t\,T} + A \bullet p^t_2 p_2^{t\,T} \\ &= \begin{bmatrix} A & B \\ -B & A \end{bmatrix} \bullet \begin{bmatrix} -p^t_2 \\ p^t_1 \end{bmatrix} \left[-p_2^{t\,T}\ p_1^{t\,T}\right]. \end{aligned} \tag{B.4}$$

Hence, the rank-one matrix

$$\overline{X}^t_V = (\lambda^t_1 + \lambda^t_2) \begin{bmatrix} p^t_1 \\ p^t_2 \end{bmatrix} \left[p_1^{t\,T}\ p_2^{t\,T}\right]$$

is globally optimal for the original OPF (it satisfies all of its constraints and produces the same objective value as X^t_{opt}). Summing up, this leads to the following corollary ($cf.$ [24]).

Corollary B.1. *Assume that* $(\lambda^t_{opt}, r^t_{opt})$ *is an optimal solution to* (B.1) *and that*

$$\dim\left(ker\left(A(\lambda^t_{opt}, r^t_{opt})\right)\right) \le 2. \tag{B.5}$$

Then, for any nonzero vector p^t *in the null space of* $A(\lambda^t_{opt}, r^t_{opt})$*, there exists two real-valued scalars* λ^t_1 *and* λ^t_2 *such that*

$$\overline{X}^t_V = (\lambda^t_1 + \lambda^t_2) p^t p^{t\,T}$$

is a global optimum of the corresponding OPF problem.

In summery, the following strategy for finding a global optimum of the underlying OPF problem can be applied (see [24]):

Algorithm framework for solving OPF problems	
Step 1.	Compute a solution $(\lambda^t_{\text{opt}}, r^t_{\text{opt}})$ of (B.1).
Step 2.	If the optimal value $h(\lambda^t_{\text{opt}}, r^t_{\text{opt}})$ is $+\infty$, then the OPF problem is infeasible.
Step 3.	Find the multiplicity ψ of the zero eigenvalue of $A(\lambda^t_{\text{opt}}, r^t_{\text{opt}})$.
Step 4.	If $\psi > 2$, (B.1) depicts a lower bound for the associated OPF.
Step 5.	If $\psi \le 2$, then a (globally) optimal solution to the associated OPF can be constructed *via* Corollary B.1.

Beside the nice feature to convexify NP-hard OPF problems, applying the above SDP approach (which squares the number of voltage variables) leads to an enormous inflation of problem size. Indeed, when considering large network instances with a huge number of buses this approach yields an SDP with an enormous number of variables. Nevertheless, according to [24], for all IEEE benchmark systems, the SDP approach works very well, meaning that these problems could be solved within a few seconds. For larger network instances tree decomposition techniques have been proposed (see [34] and [29]) in order to break down the large-scale semidefinite constraint into small-sized constraints.

Acknowledgements. The authors gratefully acknowledge the helpful suggestions of the referees and the support of the German Research Foundation (DFG) within the collaborative research center TRR 154 "Mathematical Modeling, Simulation and Optimization Using the Example of Gas Networks"

References

[1] N. Amjady and M.R. Ansari, Hydrothermal unit commitment with ac constraints by a new solution method based on benders decomposition. *Energy Convers. Manage.* **65** (2013) 57–65.

[2] M.F. Anjos and J.B. Lasserre, *Handbook on Semidefinite, Conic Polynomial Optimization*. Vol. 166 of *Internat. Ser. Oper. Res. Management Sci.* Springer (2012).

[3] K.A. Ariyawansa and Y. Zhu, Stochastic semidefinite programming: A new paradigm for stochastic optimization. *4OR: A Quart. J. Oper. Res.* **4** (2004) 239–253.

[4] X. Bai, H. Wei, K. Fujisawa and Y. Wang, Semidefinite programming for optimal power flow problems. *Int. J. Electric Power Energy Syst.* **30** (2008) 383–392.

[5] D. Bienstock and S. Mattia, Using mixed-integer programming to solve power grid blackout problems. *Discrete Optim.* **4** (2007) 115–141.

[6] J.R. Birge and F. Louveaux, Introduction to Stochastic Programming. *Internat. Series in Operations Research and Financial Engineering*, 2nd edition. Springer (2011).

[7] S. Bose, D.F Gayme, S. Low and K.M.i Chandy, Optimal power flow over tree networks. In *Communication, Control and Computing (Allerton), 2011 49th Annual Allerton Conference on*. IEEE (2011) 1342–1348.

[8] W.A. Bukhsh, A. Grothey, K. McKinnon and P. Trodden, Local solutions of optimal power flow. *IEEE Trans. Power Syst.* **28** (2013) 4780–4788.

[9] C.C. Carøe and R. Schultz, *A two-stage stochastic program for unit commitment under uncertainty in a hydro-thermal power system*. Konrad-Zuse-Zentrum, Berlin (1998), ZIB-Preprint SC 98-11.

[10] C.C. Carøe and R. Schultz, Dual decomposition in stochastic integer programming. *Oper. Res. Lett.* **24** (1999) 37–45.

[11] J. Carpentier, Contribution to the economic dispatch problem. *Bull. Soc. Franc. Elect.* **3** (1962) 431–447.

[12] C. Coffrin, H.L. Hijazi and P. Van Hentenryck, The qc relaxation: Theoretical and computational results on optimal power flow. Preprint arXiv:1502.07847 (2015).

[13] M. Farivar, C.R. Clarke, S.H. Low and K.M. Chandy, Inverter var control for distribution systems with renewables. In *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on*. IEEE (2011) 457–462.

[14] S. Frank and S. Rebennack, A primer on optimal power flow: Theory, formulation, and practical examples. Working Paper, Colorado School of Mines (2012).

[15] S. Frank, I. Steponavice and S. Rebennack, A primer on optimal power flow: A bibliographic survey (i) – formulations and deterministic methods. *Energy Systems* **3** (2012) 221–258.

[16] S. Frank, I. Steponavice and S. Rebennack, A primer on optimal power flow: A bibliographic survey (ii) – non-deterministic and hybrid methods. *Energy Systems* **3** (2012) 259–289.

[17] R. Gollmer, M.P. Nowak, W. Römisch and R. Schultz, Unit commitment in power generation - a basic model and some extensions. *Ann. Oper. Res.* **96** (2000) 167–189.

[18] T. Haukaas, Probablistic models, methods, and decisions in earthquake engineering. In *Safety, Reliability, Risk and Life-Cycle Performance of Structures & Infrastructures*, edited by G. Deodatis, B.R. Ellingwood and D.M. Frangopol. CRC Press (2014) 47–66.

[19] H. Heitsch and W. Römisch, Scenario tree modeling for multistage stochastic programs. *Math. Program.* **118** (2009) 371–406.

[20] C. Helmberg, *Semidefinite programming for combinatorial optimization*. Konrad-Zuse-Zentrum, Berlin (2000), ZIB-Report ZR 00-34.

[21] H. Hijazi, C. Coffrin and P. Van Hentenryck, *Convex quadratic relaxations of mixed-integer nonlinear programs in power systems*, Tech. rep. NICTA, Canberra, ACT Australia (2013).

[22] R. Jabr *et al.*, *Radial distribution load flow using conic programming. IEEE Trans. Power Syst.* **21** (2006) 1458–1459.

[23] S. Kuhn, *Betriebsoptimierung von elektrischen Energieerzeugungsanlagen und Übertragungssystemen bei unvollständiger Information.* Ph.D. thesis, University of Duisburg-Essen (2011).

[24] J. Lavaei and S.H. Low, Zero duality gap in optimal power flow problem. *IEEE Trans. Power Syst.* **27** (2011) 92–107.

[25] C. Lemaréchal, Lagrangian relaxation. In *Computational Combinatorial Optimization*, edited by M. Jünger and D. Naddef. Vol. 2241 of *Lect. Notes Comput. Sci.* Springer (2001) 112–156.

[26] B.C. Lesieutre, D.K. Molzahn, A.R. Borden and C.L. DeMarco, Examining the limits of the application of semidefinite programming to power flow problems. In *49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)* (2011) 1492–1499.

[27] X. Liu, S. Küçükyavuz and J. Luedtke, Decomposition algorithms for two-stage chance-constrained programs. *Math. Program.* **157** (2016) 219–243.

[28] M. Lubin, K. Martin, C. Petra and B. Sandikçi, On parallelizing dual decomposition in stochastic integer programming. *Oper. Res. Lett.* **41** (2013) 252–258.

[29] R. Madani, M. Asharphijuo and J. Lavaei, SDP Solver of Optimal Power Flow User's Manual (2014). Available at http://www.columbia.edu/~rm3122/OPF_Solver_Guide.pdf.

[30] S. Mehrotra and M.G. Özevin, Decomposition-based interior point methods for two-stage stochastic semidefinite programming. *SIAM J. Optim.* **18** (2007) 206–222.

[31] S. Mehrotra and M.G. Özevin, On the implementation of interior point decomposition algorithms for two-stage stochastic conic programs. *SIAM J. Optim.* **19** (2009) 1846–1880.

[32] D. Mehta, D.K Molzahn and K. Turitsyn, Recent advances in computational methods for the power flow equations. Preprint arXiv:1510.00073 (2015).

[33] D.K Molzahn, *Application of Semidefinite Optimization Techniques to Problems in Electric Power Systems*. Ph. D. thesis, University of Wisconsin–Madison (2013).

[34] D.K Molzahn, J.T. Holzer, B.C. Lesieutre and C.L. DeMarco, Implementation of a large-scale optimal power flow solver based on semidefinite programming. *IEEE Trans. Power Syst.* **28** (2013) 3987–3998.

[35] D.K. Molzahn, B.C. Lesieutre and C.L. DeMarco, A sufficient condition for global optimality of solutions to the optimal power flow problem. *IEEE Trans. Power Syst.* **29** (2014) 978–979.

[36] J.A. Momoh, R. Adapa and M.E. El-Hawary, A review of selected optimal power flow literature to 1993 Part i: Nonlinear and quadratic programming approaches. *IEEE Trans. Power Syst.* **14** (1999) 96–104.

[37] J.A. Momoh, R. Adapa and M.E. El-Hawary, A review of selected optimal power flow literature to 1993 Part ii: Nonlinear and quadratic programming approaches. *IEEE Trans. Power Syst.* **14** (1999) 105–111.

[38] A. Müller and D. Stoyan, *Comparison Methods for Stochastic Models and Risks*. John Wiley & Sons, INC (2002).

[39] G.C. Pflug and W. Römisch, *Modeling, Measuring and Managing Risk*. World Scientific Publishing (2007).

[40] A. Ruszczyński, A regularized decomposition method for minimizing a sum of polyhedral functions. *Math. Program.* **35** (1986) 309–333.

[41] A. Ruszczyński and A. Shapiro, Stochastic Programming. Vol. 10 of *Handbooks Oper. Res. Manage. Sci.* Elsevier (2003).

[42] P. Sánchez-Martin and A. Ramos, *Modeling transmission ohmic losses in a stochastic bulk production cost model*. Instituto de Investigación Tecnológica, Universidad Pontificia Comillas, Madrid (1997).

[43] P. Sánchez-Martin, A. Ramos and J.F. Alonso, Probablistic midterm transmission planning in a liberalized market. *IEEE Trans. Power Syst.* **20** (2005) 2135–2142.

[44] R. Schultz and S. Tiedemann, Risk aversion via excess probabilities in stochastic programs with mixed-integer recourse. *SIAM J. Optim.* **14** (2003) 115–138.

[45] A. Shapiro, D. Dentcheva and A. Ruszczyński, *Lectures on Stochastic Programming, Modeling and Theory*. Society for Industrial and Applied Mathematics (2009).

[46] S. Sojoudi and J. Lavaei, Physics of power networks makes hard optimization problems easy to solve. In *Power and Energy Society General Meeting, 2012 IEEE*. IEEE (2012) 1–8.

[47] UWEE University of Washington Electrical Engineering, *Power Systems Test Case Archive* (2014). Available at http://www.ee.washington.edu/research/pstca/.

[48] H. Wolkowitz, R. Saigal and L. Vandenberghe, eds. Handbook of Semidefinite Progamming: Theory, Algorithms and Applications. Vol. 27 of *Int. Ser. Oper. Res. Manage. Sci.* Kluwer Academic Publishers (2000).

[49] B. Zhang and D. Tse, Geometry of feasible injection region of power networks. In *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*. IEEE (201) 1508–15151.

[50] G. Zhao, A log-barrier method with benders decomposition for solving two-stage stochastic linear programs. *Math. Program.* **90** (2001) 507–536.

[51] J. Zhu, Optimization of Power System Operation. *IEEE Press Series on Power Engineering*. John Wiley & Sons, INC. (2009).