SIAM Review
Vol. 56, No. 3, pp. 461–495



# Chance-Constrained Optimal Power Flow: Risk-Aware Network Control under Uncertainty*

Daniel Bienstock†
Michael Chertkov‡
Sean Harnett§

**Abstract.** When uncontrollable resources fluctuate, optimal power flow (OPF), routinely used by the electric power industry to redispatch hourly controllable generation (coal, gas, and hydro plants) over control areas of transmission networks, can result in grid instability and, potentially, cascading outages. This risk arises because OPF dispatch is computed without awareness of major uncertainty, in particular fluctuations in renewable output. As a result, grid operation under OPF with renewable variability can lead to frequent conditions where power line flow ratings are significantly exceeded. Such a condition, which is borne by our simulations of real grids, is considered undesirable in power engineering practice. Possibly, it can lead to a risky outcome that compromises grid stability—line tripping. Smart grid goals include a commitment to large penetration of highly fluctuating renewables, thus calling to reconsider current practices, in particular the use of standard OPF. Our chance-constrained (CC) OPF corrects the problem and mitigates dangerous renewable fluctuations with minimal changes in the current operational procedure. Assuming availability of a reliable wind forecast parameterizing the distribution function of the uncertain generation, our CC-OPF satisfies all the constraints with high probability while simultaneously minimizing the cost of economic redispatch. CC-OPF allows efficient implementation, e.g., solving a typical instance over the 2746-bus Polish network in 20 seconds on a standard laptop.

**Key words.** optimization, power flows, uncertainty, wind farms, networks

**AMS subject classifications.** 15A15, 15A09, 15A23

**DOI.** 10.1137/130910312

The power grid, one of the greatest engineering achievements of the 20th century, delivers social development and resulting political stability of billions of people around the globe through control sophistication and careful long-term planning, with only very rare disruptions.

*Received by the editors February 20, 2013; accepted for publication (in revised form) April 9, 2014; published electronically August 7, 2014. The work at LANL was carried out under the auspices of the National Nuclear Security Administration of the U.S. Department of Energy at Los Alamos National Laboratory under contract DE-AC52-06NA25396. The second and third authors were partially supported by DTRA Basic Research grant BRCALL08-Per3-D-2-0022. The first and third authors were partially supported by DOE award DE-SC0002676.

http://www.siam.org/journals/sirev/56-3/91031.html

†Department of Industrial Engineering and Operations Research and Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027 (dano@columbia.edu).

‡Theoretical Division and Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM 87545 (chertkov@lanl.gov).

§Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027, and Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM 87545 (srh2144@columbia.edu).

However, the grid is under growing stress and the premise of secure electrical power may become less certain. Despite massive investments, large-scale power outages occur unpredictably and with increasing frequency. Automatic grid control and regulations achieve remarkable robustness of operation under normal fluctuations, in particular under approximately predicted interday demand trends, or even single points of failure, such as the failure of a generator or tripping of a single line. Larger, unexpected events can prove difficult to overcome. Despite remarkable sophistication, in particular highly efficient distributed frequency and voltage controls, unusual conditions often make current grid operation dependent on human input. Additionally, only some real-time data is actually used by the grid to respond to evolving conditions.

A benefit of a cost-effective migration toward increased reliance on sophisticated algorithms concerns the effective integration of renewables. This is a critical issue because large-scale introduction of renewables brings with it the risk of large, random variability—a condition that the current grid was not developed to accommodate.

This point becomes clear when we consider so-called optimal power flow (OPF) or economic dispatch, used to set generator output in typically 15-minute windows (more frequently in some cases). OPF sets generator outputs so as to meet demand at minimum cost, under operating limitations of generators and transmission lines. Estimates of the typical loads for the upcoming time window are employed in this computation. In real time, the generator outputs computed by OPF are modulated by the AGC system (see section 1.1) to account for demand variation. This scheme can fail, dramatically, when renewables are part of the generation mix and fluctuations in their output become large. By "failure" we mean not accounting for instances where a combination of generator and renewable outputs conspire to produce power flows that significantly exceed line ratings. When a line's rating is exceeded, the line becomes more likely to trip (i.e., be taken out of service). If several key lines trip, the grid likely becomes unstable and experiences a cascading failure, with large losses in served load. To prevent line tripping, an additional scheme, based on direct line flow measurements and requiring a human operator in the loop, is employed as a part of the current operational routine—after receiving a warning, the operator may initiate an emergency action, possibly disconnecting the overheated line.

Under the forthcoming high wind power penetration, significantly more frequent line overloads are likely, making the current operational paradigm clearly unsustainable. This is not an idle assumption, since firm commitments to major renewable penetration are in place throughout the world. For example, 20% renewable penetration by 2030 is a decree in the U.S. [1], and similar plans are to be implemented in Europe; see, e.g., discussions in [21, 26, 27]. At the same time, operational margins (between typical power flows and line ratings) are decreasing and expected to decrease in part due to deregulation and difficulties in expanding transmission capacity [27]. A possible failure scenario is illustrated in Figure 0.1 using as an example the U.S. Pacific Northwest regional grid data (2866 lines, 2209 buses, 176 generators, and 18 wind sources), where lines highlighted in red are jeopardized (flow becomes too high) with unacceptably high probability by fluctuating wind resources positioned along the Columbia river basin (green dots marking existing wind farms).

We propose a solution that requires somewhat accurate wind forecasts but no change in machinery or significant operational procedures. Instead, we propose to mitigate risk using an efficient algorithmic modification of the OPF approach. In computational experiments our algorithm solves realistic examples with thousands of buses and lines (such as the U.S. Pacific Northwest case) in a matter of seconds, and is

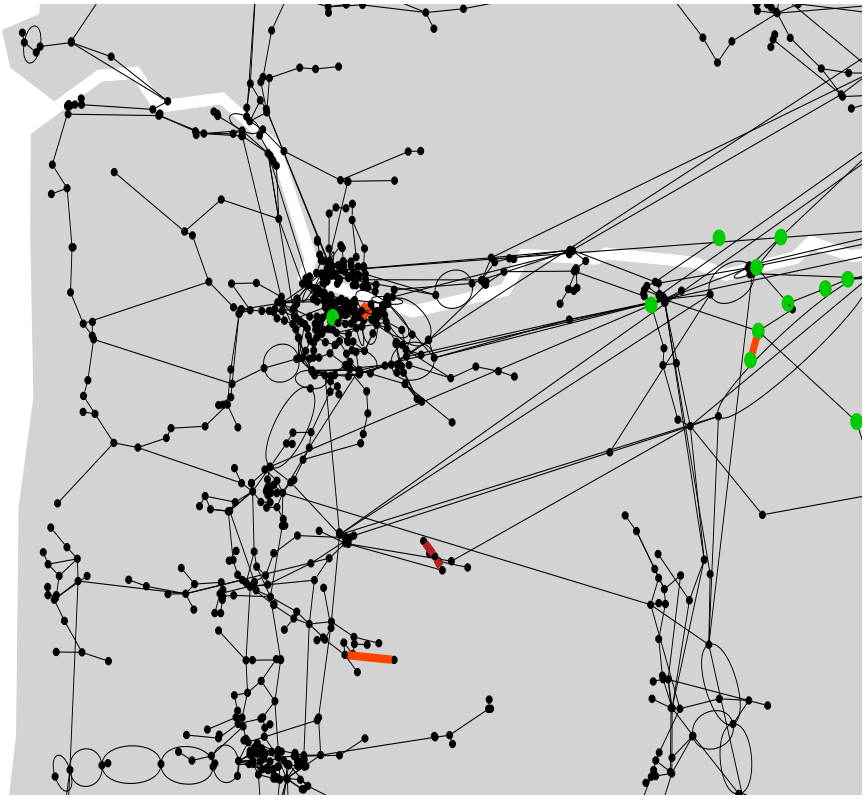

**Fig. 0.1** *Bonneville Power Administration* [52] *shown in outline under* 9% *wind penetration, where green dots mark actual wind farms. We set standard deviation to be* 0.3 *of the mean for each wind source. Our chance-constrained OPF (with* 1% *of overload set as allowable) resolved the case successfully (no overloads) and was computed in seconds, while the standard OPF showed* 8 *overloaded lines, all marked in color. Lines shown orange are at* 4% *chance of overload. There are two dark red lines which are at* 50% *of the overload while other (dark orange) lines show values of overload around* 10%.

thus only slightly slower than standard economic dispatch methods even in large-scale cases.

The nondeterministic behavior of wind makes it natural to assess the risk of an event such as a line overload in terms of probabilities; thus our proposed approach relies on stochastic optimization. In a system under stochastic risk, an extremely large variety of events that could pose a danger might emerge. Recent works [18, 19, 4] suggest that focusing on instantons, or most-likely (dangerous) events, provides a practicable route to risk control and assessment. However, there may be far too many comparably probable instantons, and, furthermore, we need a computationally efficient methodology that not only identifies dangerous, relatively probable events, but also mitigates them.

This paper suggests a new approach that searches for the most probable realizations of line overloads under renewable generation and corrects such situations through control actions, simultaneously and efficiently in one step. Our approach

relies on "chance-constrained" optimization [41]. Chance-constrained optimization problems are optimization problems under uncertainty, with constraints stating that the probability of a certain random event is kept smaller than a target value.

Our reformulation of standard OPF minimizes the average cost of generation over the random power injections, while specifying a mechanism by which (standard, i.e., controllable) generators compensate in real time for renewable power fluctuations so as to guarantee low probability that any line will exceed its rating. This last constraint is naturally formulated as a chance constraint—we term our approach chance-constrained OPF, or CC-OPF.

This paper is organized as follows. In section 1 we motivate and present the various mathematical models used to describe how the grid operates, as well as our proposed methodology. We explain how to solve the models in section 2. We then present in section 3 a number of examples to demonstrate the speed and usefulness of our approach. Section 4 summarizes the results and discusses the path forward. In the appendix we consider a data-robust version of our CC-OPF method, as well as additional experiments.

## 1. Formulating Chance-Constrained Optimum Power Flow Models.

**1.1. Transmission Grids: Controls and Limits.** In this paper we consider *transmission grids* which operate at high voltages so as to convey power economically, with minimal losses, over large distances. In contrast, *distribution systems* are typically residential, lower voltage grids used to provide power to individual consumers. From the point of view of wind power generation, smooth operation of transmission systems is key since reliable wind sources are frequently located far away from consumption.

Transmission systems balance consumption/load and generation using a complex strategy that spans three different time scales (see, e.g., [7]). At any point in time, generators produce power at a previously computed base level; power is generated (and transmitted) in alternating current (AC) form. An essential stability ingredient is that all generators operate at a common frequency (e.g., 60Hz in North America). In real time, changes in loads are registered at generators through (opposing) changes in frequency. Consider the case of a sudden load increase. In that case generator frequency will start to drop. The so-called primary frequency control, normally implemented on some gas, hydro, and coal plants, will react so as to stop frequency drift by having each responding generator convey more power to the system, proportionally to the frequency change. This reaction is swift and local, leading to stabilization of frequency across the system, however, not necessarily at the nominal 60Hz value. The task of the secondary, or automatic generation control (AGC), is to undertake the adjustment of generation levels to return frequency to the nominal value.[1] The OPF algorithm typically runs as frequently as every 15 minutes, providing information for AGC, which ultimately undertakes the adjustment of generation levels to achieve optimal (or close to optimal) control. The OPF time window thus represents the shortest time scale where actual off-line and network-wide optimal computations are employed.

**1.2. OPF—Standard Generation Dispatch.** As discussed above, OPF [31, 36, 7] is used to reset generator output levels over a control area of the transmission grid (for example, over the Bonneville Power Administration (BPA) grid shown in Figure 0.1).

[1] AGC is based on SCADA communications, which is receiving an updated signal every 4s or so; however, the actual control action is computed based on integration of the SCADA communicated signal over a longer window of several minutes.

In order to describe OPF we will employ power engineering terms such as "bus" to refer to a graph-theoretic vertex and "line" to refer to an edge. The set of all buses will be denoted by $\mathcal{V}$, the set of lines by $\mathcal{E}$, and the set of buses that house generators by $\mathcal{G}$. We let $n = |\mathcal{V}|$. A line joining buses $i$ and $j$ is denoted by $\{i, j\}$, indicating an unordered pair. We assume that the underlying graph is connected, without loss of generality.

The generic OPF problem can be stated as follows:

- The goal is to determine the vector $p \in \mathbb{R}^{\mathcal{G}}$,[2] where for $i \in \mathcal{G}$, $p_i$ is the output of generator $i$, so as to minimize an objective function $c(p)$. This function is, usually, a convex, separable quadratic function of $p$:
$$c(p) = \sum_{i \in \mathcal{G}} c_i(p_i),$$
where each $c_i$ is convex quadratic.
- The problem is endowed by three types of constraints: power flow, line limit, and generation bound constraints.[3]

Generation bounds are simple box constraints on the individual $p_i$. Thermal line limits place an upper bound on the power flows in each line. Power flow constraints, in their most general form, are Kirchhoff's circuit laws stated in terms of voltages (potentials) and power flows. In this context, for each bus $i \in \mathcal{V}$ its voltage $U_i$ is defined as $v_i e^{j\theta_i}$, where $v_i$ and $\theta_i$ are the voltage magnitude and phase angle at bus $i$. Voltages can be used to derive expressions for other physical quantities, in particular power, obtaining in the simplest case a system of quadratic equations on the voltage real and imaginary coordinates; see, e.g., [36, 7].

The AC power flow equations can constitute an obstacle to solvability of OPF (from a technical standpoint, they give rise to nonconvexities). In transmission system analysis a linear set of equations that gives an approximate representation of angles and real power flows is commonly used, the so-called DC-approximation. In this approximation (a) all voltage magnitudes are assumed fixed and rescaled to unity; (b) phase differences between neighboring nodes are assumed small $\forall \{i, j\} \in \mathcal{E} : |\theta_i - \theta_j| \ll 1$; (c) thermal losses are ignored (reactance dominates resistance for all lines). Then the (real) power flow $f_{ij}$ from $i$ to $j$ over line $\{i, j\}$, with line susceptance $\beta_{ij}$ $(= \beta_{ji})$, is related linearly to the respective phase difference,

$$f_{ij} = \beta_{ij}(\theta_i - \theta_j). \tag{1.1}$$

Note that $f_{ji} = -f_{ij}$ so that the notation is consistent. For convenience of notation we will extend the vector $p$ to include an entry for every bus $i \in \mathcal{V}$ with the proviso that $p_i = 0$ whenever $i \notin \mathcal{G}$. Likewise, denote by $d \in \mathbb{R}_+^{|\mathcal{V}|}$ the vector of (possibly zero) demands and by $\theta \in \mathbb{R}^{|\mathcal{V}|}$ the vector of phase angles. Then a vector $f$ of power flows is feasible iff

$$\sum_{j:\{i,j\} \in \mathcal{E}} f_{ij} = p_i - d_i \quad \text{for each bus } i, \tag{1.2}$$

[2]Here and in what follows, notation of the form "$x \in \mathbb{R}^S$" indicates that the entries in the vector $x$ are indexed by the set $S$, so $x$ is $|S|$-dimensional.

[3]In practice, the so-called security-constrained OPF (SC-OPF) is implemented by utility. SC-OPF modifies OPF to include additional constraints to guarantee feasibility of the generation dispatch against any possible $(N-1)$ contingency accounting for loss of a single line or single generator. However, the generalization does not lead to any principal complications as all the new SC-related constraints can be treated on an equal footing with the standard power flow, line flow, and generation constraints of the base OPF.

and in view of equation (1.1), this can be restated as

$$\theta_i \sum_{j:\{i,j\}\in\mathcal{E}} \beta_{ij} \quad - \quad \sum_{j:\{i,j\}\in\mathcal{E}} \beta_{ij}\theta_j = p_i - d_i \quad \text{for each bus } i. \tag{1.3}$$

In matrix form this equation can be rewritten as follows:

$$B\theta = p - d, \tag{1.4}$$

where the $n \times n$ matrix $B$ is a weighted-Laplacian defined as follows:

$$\forall i,j: \quad B_{ij} = \begin{cases} -\beta_{ij}, & \{i,j\} \in \mathcal{E}, \\ \sum_{k;\{k,j\}\in\mathcal{E}} \beta_{kj}, & i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{1.5}$$

For future reference, we state some well-known properties of Laplacians and the power flow system (1.4).

LEMMA 1.1. *The sum of rows of $B$ is zero, and under the connectedness assumption for the underlying graph the rank of $B$ equals $n-1$. Thus system* (1.4) *is feasible in $\theta$ iff*

$$\sum_i p_i = \sum_i d_i. \tag{1.6}$$

*In other words, under the DC model the power flow equations* (1.4) *are feasible precisely when total generation equals total demand. Moreover, if equations* (1.4) *are feasible, then for any index $1 \le j \le n$ there is a solution with $\theta_j = 0$.* □

In summary, the standard DC-formulation OPF[4] problem can be stated as the following constrained optimization problem:

$$\textbf{OPF:} \quad \min_{p,\theta} c(p) \quad \text{s.t.} \tag{1.7}$$

$$B\theta = p - d, \tag{1.8}$$

$$\forall i \in \mathcal{G}: \quad p_i^{min} \le p_i \le p_i^{max}, \tag{1.9}$$

$$\forall \{i,j\} \in \mathcal{E}: \quad |\beta_{ij}(\theta_i - \theta_j)| \le f_{ij}^{max}. \tag{1.10}$$

Note that the $p_i^{min}, p_i^{max}$ quantities can be used to enforce the convention $p_i = 0$ for each $i \notin \mathcal{G}$; if $i \in \mathcal{G}$, then $p_i^{min}, p_i^{max}$ are lower and upper generation bounds which are generator-specific. Constraint (1.10) is the line limit constraint for $\{i,j\}$; $f_{ij}^{max}$ represents the line limit (typically a thermal limit), which is assumed to be strictly enforced. This conservative condition will be relaxed in what follows.

Problem (1.7) is a convex quadratic program, easily solved using modern optimization tools. The vector $d$ of demands is *fixed* in this problem and is obtained through estimation. In practice, however, demand will fluctuate around $d$; generators then respond by adjusting their output (from the OPF-computed quantities) proportionally to the overall fluctuation as discussed above. This scheme works well in current practice, as demands do not substantially fluctuate on the time scale to which OPF applies.

[4]For brevity, we will simply refer to this scheme as the "standard OPF."

**1.3. Chance-Constrained OPF: Motivation.** To motivate our approach we outline how generator output is modulated, in real time, in response to demand fluctuations. Suppose we have computed, using OPF, the output $p_i$ for each generator $i$ assuming constant demands $d$. Let $\hat{d}(t)$ be the vector of real-time demands at time $t$. Then so-called frequency control, or more properly, primary and secondary controls in combination will achieve (altogether, on the scale of minutes) the following outputs to quantities $\hat{p}_i(t)$:

$$\hat{p}_i(t) = p_i - \rho_i \sum_j (d_j - \hat{d}_j(t)) \quad \text{for each } i \in \mathcal{G}. \tag{1.11}$$

In this equation, the quantities $\rho_i \geq 0$ satisfy

$$\sum_i \rho_i = 1.$$

Thus, from (1.11) we obtain

$$\sum_i \hat{p}_i(t) = \sum_i p_i - \sum_j (d_j - \hat{d}_j(t)) = \sum_j \hat{d}_j(t)$$

from (1.6); in other words, the requirements are met. The parameters $\rho_i \geq 0$ are generator-dependent, related to the so-called generator AGC participation factors which are typically chosen through an auxiliary optimization task which focuses on (economic) cost, but does not take into account stochastic renewable fluctuations. See [53, 3] for relevant discussions.

Thus, in effect, generator outputs are set in hierarchical fashion, using OPF to compute a base level, with real-time adjustments as per (1.11) which is furthermore risk-unaware. This scheme has worked in the past because of the slow time scales of change in uncontrolled resources (mainly loads). That is to say, frequency control and load changes are well separated. A large error in the forecast or an underestimation of possible $d$ for the next—e.g., fifteen-minute—period may lead to an operational problem (see, e.g., the discussions in [33, 37]) because even though the vector $\hat{p}(t)$ is sufficient to meet average demands, the $\hat{\theta}(t)$ computed from

$$B\hat{\theta}(t) = \hat{p}(t) - \hat{d}(t)$$

may give rise to real-time power flows

$$\hat{f}_{ij}(t) \doteq \beta_{ij}[\hat{\theta}_i(t) - \hat{\theta}_j(t)]$$

that violate constraints (1.10). Even the generator constraints (1.9) may fail to hold. This has not been considered a handicap simply because any resulting line trips are rare, primarily because the deviations $\hat{d}_i(t) - d_i$ will be small in the time scale of interest. In effect, the risk-unaware approach that assumes constant demands has worked well.

This perspective changes when renewable power sources such as wind are incorporated. We assume that a subset $\mathcal{W}$ of the buses holds uncertain power sources (wind farms); for each $j \in \mathcal{W}$, write the amount of power generated by source $j$ at time $t$ as $\mu_j + \omega_j(t)$, where $\mu_j$ is the forecast output of farm $j$ in the time period of interest. For ease of exposition, we will assume in what follows that $\mathcal{G}$ refers to the set of buses holding *controllable* generators, i.e., $\mathcal{G} \cap \mathcal{W} = \emptyset$. Renewable generation

can be incorporated into the OPF formulation (1.7)–(1.10) by simply setting $p_i = \mu_i$ for each $i \in \mathcal{W}$. Assuming constant demands but fluctuating renewable generation, the application of the frequency control yields the following analogue to (1.11):

$$\hat{p}_i(t) = p_i - \rho_i \sum_{j \in \mathcal{W}} \omega_j(t) \quad \text{for each } i \in \mathcal{G}, \tag{1.12}$$

e.g., if $\sum_{j \in \mathcal{W}} \omega_j(t) > 0$, that is to say, there is a net increase in wind output, then (controllable) generator output will proportionally decrease.

Equation (1.12) describes how generation will adjust to wind changes under current power engineering practice. The hazard embodied in this relationship is that the quantities $\omega_j(t)$ *can* be large, resulting in stochastic changes in power flows significant enough to overload power lines. The risk of such overloads can be expected to increase (see [27]); this is due to a projected increase of renewable penetration in the future, accompanied by the decreasing gap between normal operation and limits set by line capacities. Lowering of the thermal limits (the $f_{ij}^{max}$ quantities in (1.10)) can succeed in deterministically preventing overloads, but it also forces excessively conservative choices of the generation redispatch, potentially causing extreme volatility of the electricity markets. See, e.g., the discussion in [50] on abnormal price fluctuations in markets that are heavily reliant on renewables.

**1.4. Using Chance Constraints.** Power lines do not fail (i.e., trip) instantly when their flow thermal limits are exceeded. A line carrying flow that exceeds the line's thermal limit will gradually heat up, possibly sag, and, through a variety of processes (such as a contact), may then trip or be disconnected by the operator. The precise process is extremely difficult to calibrate.[5] Additionally, the rate at which a line overheats depends on its overload, which may dynamically change (or even temporarily disappear) as flows adjust due to external factors, in our case fluctuations in renewable outputs. What *can* be stated with certainty is that the longer a line stays *overheated*, the higher the probability that it will trip—to put it differently, if a line remains overheated over a duration possibly measured in minutes to ten of minutes, it will trip.

Even though an exact representation of line tripping seems difficult, we can, however, state a practicable alternative. Ideally, we would use a constraint of the form "*for each line, the fraction of the time that it exceeds its limit within a certain time window is small.*" Direct implementation of this constraint would require resolving dynamics of the grid over the generator dispatch time window of interest. Instead we propose the following static proxy of this ideal model, a *chance constraint*: we will require that the probability that a given line exceeds its limit be small.

To formalize this notion, we assume the following:

W.1. For each $i \in \mathcal{W}$, the (stochastic) amount of power generated by source $i$ is of the form $\mu_i + \boldsymbol{\omega_i}$, where

W.2. $\mu_i$ is constant, assumed known from the forecast, and $\boldsymbol{\omega_i}$ is a zero mean independent random variable with known standard deviation $\sigma_i$.

Here and in what follows, we use boldface to indicate uncertain quantities. Let $\boldsymbol{f_{ij}}$ be the flow on a given line $\{i, j\}$, and let $0 < \epsilon_{ij}$ be small. The chance constraint for line $\{i, j\}$ is

$$P(\boldsymbol{f_{ij}} > f_{ij}^{max}) < \epsilon_{ij} \quad \text{and} \quad P(\boldsymbol{f_{ij}} < -f_{ij}^{max}) < \epsilon_{ij} \qquad \forall \{i, j\}. \tag{1.13}$$

[5] We refer the reader to [48] for discussions of line tripping during the 2003 Northeast U.S.-Canada cascading failure.

One could alternatively use

$$P(|\boldsymbol{f_{ij}}| > f_{ij}^{max}) < \epsilon_{ij} \quad \forall\{i,j\}, \tag{1.14}$$

which is more conservative than (1.13). If (1.14) holds, then so does (1.13), and if the latter holds, then $P(|\boldsymbol{f_{ij}}| > f_{ij}^{max}) < 2\epsilon_{ij}$. However, (1.13) proves more tractable, and moreover we are interested in the regime where $\epsilon_{ij}$ is fairly small; thus we estimate that there is small practical difference between the two constraints; this will be verified by our numerical experiments. Likewise, for a generator $g$ we will require that

$$P(\boldsymbol{p_g} > p_g^{max}) < \epsilon_g \quad \text{and} \quad P(\boldsymbol{p_g} < p_g^{min}) < \epsilon_g. \tag{1.15}$$

The parameter $\epsilon_g$ will be chosen *extremely* small, so that for all practical purposes all generator outputs will be guaranteed to stay within respective bounds.

Chance constraints [45, 16, 38] are but one possible methodology for handling uncertain data in optimization. Broadly speaking, this methodology fits within the general field of stochastic optimization. Constraint (1.13) can be viewed as a "value-at-risk" statement; the closely related "conditional value-at-risk" concept provides a (convex) alternative, which roughly stated constrains the expected overload of a line to remain small, *conditional* on there being an overload (see [41] for definitions and details).

One alternative model would insist on imposing the (much stronger) constraint

$$P(\exists \text{ line } \{i,j\} \text{ s.t. } |\boldsymbol{f_{ij}}| > f_{ij}^{max}) < \epsilon, \tag{1.16}$$

where $0 < \epsilon < 1$. Nemirovski and Shapiro [41] develop a general framework for constructing, under appropriate assumptions, a convex optimization problem with an approximate version of constraint (1.16), using methodology inspired by the large deviations theory.

Reference [54] considers the standard OPF problem under stochastic demands, using chance constraints to guarantee high probability that the system operates within acceptable bounds. The problem is tackled using a simulation-based local optimization system, with experiments using a 5-bus and a 30-bus example.

Another related study [49] describes a scenario-based system for reserve scheduling with fluctuating wind generation, using chance constraints to limit line or generator overloads. This optimization is tackled via transformation to a convex problem plus a heuristic scheme, with no convergence to global optimum of the nonlinear problem guaranteed.

Chance-constrained optimization has also been discussed recently in relation to the unit commitment problem, which concerns discrete-time planning for operation of large generation units on the scale of hours to months, so as to account for the long-term wind-farm generation uncertainty [42, 51, 55].

**1.5. Uncertain Power Sources.** In our model we assume independence of wind power fluctuations at different sites; this is justified by the fact that the wind farms are sufficiently far away from each other. For the typical OPF time span of 15 minutes and typical wind speed of 10m/s, fluctuations of wind at the farms more than 10km apart are not correlated.

We make additional simplifying assumptions that are approximately consistent with our general physics understanding of fluctuations in atmospheric turbulence;

in particular we assume Gaussianity of $\boldsymbol{\omega_i}$.[6] We will also assume that only a standard weather forecast (coarse grained on minutes and kilometers) is available, and all available wind energy resources are transformed into power without spillage.[7]

Additionally, under the Gaussian assumption, chance constraint (1.13) can be captured using an optimization framework that proves particularly efficient. We will also consider a data-robust version of our chance-constrained problem where the parameters for the Gaussian distributions are not precisely known. This allows both for parameter mis-estimation and for model error, that is to say, the implicit approximation of non-Gaussian distributions with Gaussians; our approach, detailed in section 2.4, remains computationally sound in this robust setting.

Other fitting distributions considered in the wind-modeling literature, e.g., Weibull distributions and logistic distributions [14, 30], will be discussed later in the text as well.[8] In particular, we will demonstrate on out-of-sample tests that the computationally advantageous Gaussian modeling of uncertainty allows as well to model effects of other distributions.

**1.6. Affine Control.** Since the power injections at each bus are fluctuating, we need a control scheme to ensure that generation is equal to demand at all times within the time window of interest. We assume that all governors involved in the controls respond to fluctuations in the generalized load (actual demand which is assumed frozen minus stochastic wind resources) in a proportional way, however, with possibly different proportionality coefficients. Thus, we term the joint result of the primary frequency control and secondary frequency control the *affine* control. The stochastic version of (1.12) thus becomes

$$\forall \,\text{bus}\, i \in \mathcal{G}: \quad \boldsymbol{p_i} = \bar{p}_i - \alpha_i \sum_{j \in \mathcal{W}} \boldsymbol{\omega_j}. \quad \alpha_i \geq 0. \tag{1.17}$$

Here the quantities $\bar{p}_i \geq 0$ and $\alpha_i \geq 0$ are design variables satisfying (among other constraints) $\sum_{i \in \mathcal{G}} \alpha_i = 1$. Thus the generator output $p_i$ combines a fixed term $\bar{p}_i$ and a term which varies with wind, $-\alpha_i \sum_{j \in \mathcal{W}} \boldsymbol{\omega_j}$. Observe that $\sum_i p_i = \sum_i \bar{p}_i - \sum_i \boldsymbol{\omega_j}$, that is, the total power generated equals the average production of the generators minus any additional wind power above the average case. We allow $\alpha$ to change, leaving it to the wind fluctuations–aware optimization to decide the optimal value. (In some cases it may even be advantageous to allow negative $\alpha_i$, but we decided not to consider such a drastic change of current policy in this study.)

This affine control scheme creates the possibility of requiring a generator to produce power beyond its limits. With unbounded wind, this possibility is inevitable,

[6] Correlations of wind turbine generated power within the correlation time of 15 minures, roughly equivalent to the time span between the two consecutive OPFs, are approximately Gaussian. The assumption is not perfect, in particular because it ignores significant up and down ramps possibly extending tails of the distribution in the regime of really large deviations. Also see [14, 30] and references therein.

[7] See [27] for some empirical validation.

[8] Note that the fitting approach of [14, 30] does not differentiate between typical and atypical events and assumes that the main body and the tail should be modeled using a simple distribution with only one or two fitting parameters. Generally this assumption is not justified as the physical origin of the typical and anomalous contributions of the wind, contributing to the main body and the tail of the distribution, respectively, are rather different. Gaussian fit (of the tail)—or more accurately, faster than exponential decay of probability in the tail for relatively short-time (under one hour) forecast—would be reasonably consistent with phenomenological modeling of turbulence generating these fluctuations.

though we can restrict it to occur with arbitrarily small probability, which we will do with additional chance constraints for all controllable generators, $\forall g \in \mathcal{G}$,

$$P(p_g^{min} \leq \bar{p}_g - \alpha_g \sum_{j \in \mathcal{W}} \boldsymbol{w}_j \leq p_g^{max}) > 1 - \epsilon_g. \tag{1.18}$$

**1.7. CC-OPF: Brief Discussion of Solution Methodology.** Our methodology applies and develops general ideas on chance-constrained optimization [41] to the setting of OPF under uncertainty. In section 2.1 we will provide a generic formulation of our chance-constrained OPF problem that is valid under the assumption of linear power flow laws and statistical independence of wind fluctuations at different sites, while using control law (1.17) to specify standard generation response to wind fluctuations. Under the additional assumption of Gaussianity, in section 2.2 this formulation is reduced to a deterministic convex optimization problem, more precisely, a second-order cone program (SOCP) [15, 24]; an efficient computational implementation is discussed in section 2.3. We will term this SOCP, which assumes known values of the wind distribution parameters, the *nominal* problem. In section 2.4 we extended the formulation to account for data-related uncertainty in the parameters of the Gaussian distributions.

Many of our assumptions are not restrictive and allow natural generalizations. Using techniques from [41], one can relax the assumption of wind source Gaussianity. For example, using only the mean and variance of output at each wind farm, one can use Chebyshev's inequality to obtain a similar though more conservative formulation. And following [41] we can also obtain convex approximations to (1.13) which are tighter than Chebyshev's inequality for a large number of empirical distributions discussed in the literature. The data-robust version of our algorithm provides a methodologically sound (and computationally efficient) means to protect against data and model errors. Additionally out-of-sample experiments (below) involving the controls computed with the nominal approach (first, to investigate the effect of parameter estimation errors in the Gaussian case and, second, to gauge the impact of non-Gaussian wind distributions) indicate robustness.

## 2. Solving the Models.

**2.1. Chance-Constrained Optimal Power Flow: Formal Expression.** Following the W.1 and W.2 notation, equations (1.17) explain the *affine* control, given that the $\alpha_i$ are decision variables in our CC-OPF, additional to the standard $\bar{p}_i$ decision variables already used in the standard OPF (1.7). For $i \notin \mathcal{W}$ write $\mu_i = \sigma_i = 0$ (so that $\boldsymbol{\omega_i} = 0$), thereby obtaining vectors $\mu$, $\sigma$, $\boldsymbol{\omega} \in \mathbb{R}^n$. Likewise, extend $\bar{p}$ and $\alpha$ to vectors in $\mathbb{R}^n$ by writing $\bar{p}_i = \alpha_i = 0$ whenever $i \notin \mathcal{G}$.

DEFINITION 2.1. *We say that the pair $\bar{p}, \alpha$ is viable if the generator outputs under control law* (1.17)*, together with the uncertain outputs, always exactly match total demand.*

The following simple result characterizes this condition as well as other basic properties of the affine control. Here and below, $e \in \mathbb{R}^n$ is the vector of all 1's.

LEMMA 2.1. *Under the control law* (1.17) *the net output of bus $i$ equals*

$$\bar{p}_i + \mu_i - d_i + \boldsymbol{\omega}_i - \alpha_i(e^T \boldsymbol{\omega}), \tag{2.1}$$

*and thus the (stochastic) power flow equations can be written as*

$$B\boldsymbol{\theta} = \bar{p} + \mu - d + \boldsymbol{\omega} - (e^T \boldsymbol{\omega})\alpha. \tag{2.2}$$

*Consequently, the pair $\bar{p}, \alpha$ is viable iff*

$$\sum_{i \in \mathcal{V}} (\bar{p}_i + \mu_i - d_i) = 0, \tag{2.3}$$
$$\alpha \geq 0 \quad \textit{and} \quad \sum_i \alpha_i = 1.$$

*Proof.* Equation (2.1) follows by definition of the $\bar{p}$, $\mu$, $d$ vectors and the control law. Thus (2.2) holds. By Lemma 1.1 from (2.2) one gets that $\bar{p}, \alpha$ is viable iff

$$\begin{aligned} 0 &= \sum_{i=1}^{n} (\bar{p}_i - (e^T \boldsymbol{\omega})\alpha_i + \mu_i + \boldsymbol{\omega_i} - d_i) \\ &= \sum_i (\bar{p}_i + \mu_i - d_i), \end{aligned} \tag{2.4}$$

since by construction $\sum_i \alpha_i = 1$. ∎

*Remark.* Equation (2.3) can be interpreted as stating the condition that expected total generation must equal total demand; however, the lemma contains a rigorous proof of this fact.

As remarked before, any $(n-1) \times (n-1)$ matrix obtained by striking out the same column and row of $B$ is invertible. For convenience of notation we will assume that bus $n$ is neither a generator nor a wind farm bus, that is to say, $n \notin \mathcal{G} \cup \mathcal{W}$, and we denote by $\hat{B}$ the submatrix obtained from $B$ by removing row and column $n$ and write

$$\breve{\mathcal{B}} = \begin{pmatrix} \hat{B}^{-1} & 0 \\ 0 & 0 \end{pmatrix}. \tag{2.5}$$

Further, by Lemma 1.1 we can assume without loss of generality that $\boldsymbol{\theta_n} = 0$. The following simple result will be used in what follows.

LEMMA 2.2. *Suppose the pair $\bar{p}, \alpha$ is viable. Then under the control law* (1.17) *a vector of (stochastic) phase angles is*

$$\boldsymbol{\theta} = \bar{\theta} + \breve{\mathcal{B}}(\boldsymbol{\omega} - (e^T \boldsymbol{\omega})\alpha), \quad \textit{where} \tag{2.6}$$
$$\bar{\theta} = \breve{\mathcal{B}}(\bar{p} + \mu - d). \tag{2.7}$$

*As a consequence,*

$$\mathbb{E}_{\boldsymbol{\omega}} \boldsymbol{\theta} = \bar{\theta}, \tag{2.8}$$

*and given any line $\{i, j\}$,*

$$\mathbb{E}_{\boldsymbol{\omega}} \boldsymbol{f_{ij}} = \beta_{ij}(\bar{\theta}_i - \bar{\theta}_j). \tag{2.9}$$

*Furthermore, each quantity $\boldsymbol{\theta_i}$ or $\boldsymbol{f_{ij}}$ is an affine function of the random variables $\boldsymbol{\omega_i}$.*

*Proof.* For convenience we rewrite system (2.2): $B\boldsymbol{\theta} = \bar{p} + \mu - d + \boldsymbol{\omega} - (e^T \boldsymbol{\omega})\alpha$. Since $\bar{p}, \alpha$ is viable, this system is always feasible, and since the sum of rows of $B$ is zero, its last row is redundant. Therefore (2.6) follows since $\boldsymbol{\theta_n} = 0$, and (2.8) holds since $\boldsymbol{\omega}$ has zero mean. Since $\boldsymbol{f_{ij}} = \beta_{ij}(\boldsymbol{\theta_i} - \boldsymbol{\theta_j})$ for all $\{i, j\}$, (2.9) holds. From this fact and (2.6) it follows that $\boldsymbol{\theta}$ and $\boldsymbol{f}$ are affine functions of $\boldsymbol{\omega}$. ∎

Using this result we can now give an initial formulation to our chance-constrained problem, with discussion following.

$$\text{CC-OPF:} \quad \min \ \mathbb{E}_{\boldsymbol{\omega}} \left[ c(\bar{p} - (e^T \boldsymbol{\omega})\alpha) \right] \tag{2.10}$$

$$\text{s.t.} \quad \sum_{i\in\mathcal{G}} \alpha_i = 1, \quad \alpha \geq 0, \quad \bar{p} \geq 0, \tag{2.11}$$

$$\sum_{i\in\mathcal{V}} (\bar{p}_i + \mu_i - d_i) = 0, \tag{2.12}$$

$$B\bar{\theta} = \bar{p} + \mu - d; \tag{2.13}$$

for all lines $\{i, j\}$,

$$P\left(\beta_{ij}(\bar{\theta}_i - \bar{\theta}_j + [\breve{\mathcal{B}}(\boldsymbol{\omega} - (e^T\boldsymbol{\omega})\alpha)]_i - [\breve{\mathcal{B}}(\boldsymbol{\omega} - (e^T\boldsymbol{\omega})\alpha)]_j) > f_{ij}^{max}\right) < \epsilon_{ij}, \tag{2.14}$$

$$P\left(\beta_{ij}(\bar{\theta}_i - \bar{\theta}_j + [\breve{\mathcal{B}}(\boldsymbol{\omega} - (e^T\boldsymbol{\omega})\alpha)]_i - [\breve{\mathcal{B}}(\boldsymbol{\omega} - (e^T\boldsymbol{\omega})\alpha)]_j) < -f_{ij}^{max}\right) < \epsilon_{ij}; \tag{2.15}$$

for all generators $g$,

$$P\left(\bar{p}_g - (e^T\boldsymbol{\omega})\alpha_i > p_g^{max}\right) < \epsilon_g \quad \text{and} \quad P\left(\bar{p}_g - (e^T\boldsymbol{\omega})\alpha_i < p_g^{min}\right) < \epsilon_g. \tag{2.16}$$

The variables in this formulation are $\bar{p}, \alpha$, and $\bar{\theta}$. Constraint (2.11) simply states basic conditions needed by the affine control. Constraint (2.12) is (2.3). Constraints (2.13), (2.14), and (2.15) express our chance constraint in view of Lemma 2.2.

The objective function is the expected cost incurred by the stochastic generation vector

$$\boldsymbol{p} = \bar{p} - (e^T\boldsymbol{\omega})\alpha$$

over the varying wind power output $\boldsymbol{w}$. In standard power engineering practice generation cost is convex, quadratic, and separable; i.e., for any vector $p$, $c(p) = \sum_i c_i(p_i)$, where each $c_i$ is convex quadratic. Note that for any $i \in \mathcal{G}$ we have

$$\boldsymbol{p_i^2} = \bar{p}_i^2 + (e^T\boldsymbol{\omega})^2\alpha_i^2 - 2e^T\boldsymbol{\omega}\bar{p}_i\alpha_i,$$

from which we obtain, since the $\omega_i$ have zero mean,

$$\mathbb{E}_{\boldsymbol{w}}(\boldsymbol{p_i^2}) = \bar{p}_i^2 + \mathrm{var}(\boldsymbol{\Omega})\alpha_i^2,$$

where "var" denotes variance and $\boldsymbol{\Omega} \doteq \sum_j \boldsymbol{\omega_j}$. It follows that the objective function can be written as

$$\mathbb{E}_{\boldsymbol{\omega}} c(\boldsymbol{p}) = \sum_{i\in\mathcal{G}} \left\{c_{i2}\left(\bar{p}_i^2 + \mathrm{var}(\boldsymbol{\Omega})\alpha_i^2\right) + c_{i1}\bar{p}_i + c_{i0}\right\}, \tag{2.17}$$

where $c_{i2} \geq 0$ for all $i \in \mathcal{G}$. Consequently the objective function is convex quadratic as a function of $\bar{p}$ and $\alpha$.

The above formulation is the formal statement for our optimization problem. Even though its objective is convex in cases of interest, the formulation is not in a form that can be readily exploited by standard optimization algorithms. Below we will provide an efficient approach for solving relevant classes of problems with the above form; prior to that we need a technical result. We will employ the following notation:

- For $j \in \mathcal{W}$, the variance of $\boldsymbol{\omega_j}$ is denoted by $\sigma_j^2$.
- For $1 \le i, j \le n$ let $\pi_{ij}$ denote the $i, j$ entry of the matrix $\breve{\mathcal{B}}$ given above, that is to say,

$$\pi_{ij} = \begin{cases} [\hat{B}^{-1}]_{ij}, & i < n, \\ 0 & \text{otherwise.} \end{cases} \tag{2.18}$$

- Given $\alpha$, for $1 \le i \le n$ write

$$\delta_i \doteq [\breve{\mathcal{B}}\alpha]_i = \begin{cases} \sum_{1=j}^{n-1} \pi_{ij}\alpha_j, & i < n, \\ 0 & \text{otherwise.} \end{cases} \tag{2.19}$$

LEMMA 2.3. *Assume that the $\boldsymbol{\omega_i}$ are independent random variables. Given $\alpha$, for any line $\{i, j\}$,*

$$\operatorname{var}(\boldsymbol{f_{ij}}) = \beta_{ij}^2 \sum_{k \in \mathcal{W}} \sigma_k^2 (\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2. \tag{2.20}$$

*Proof.* Using $\boldsymbol{f_{ij}} = \beta_{ij}[\boldsymbol{\theta_i} - \boldsymbol{\theta_j}]$ and (2.6) we have that

$$\boldsymbol{f_{ij}} - \mathbb{E}_{\boldsymbol{\omega}} \boldsymbol{f_{ij}} = \beta_{ij}([\breve{\mathcal{B}}(\boldsymbol{\omega} - (e^T\boldsymbol{\omega})\alpha)]_i - [\breve{\mathcal{B}}(\boldsymbol{\omega} - (e^T\boldsymbol{\omega})\alpha)]_j) \tag{2.21}$$

$$= \beta_{ij}\left([\breve{\mathcal{B}}\boldsymbol{\omega}]_i - [\breve{\mathcal{B}}\boldsymbol{\omega}]_j - (e^T\boldsymbol{\omega})\delta_i + (e^T\boldsymbol{\omega})\delta_j\right) \tag{2.22}$$

$$= \beta_{ij} \sum_{k \in \mathcal{W}} (\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)\boldsymbol{\omega_k}, \tag{2.23}$$

since by convention $\boldsymbol{\omega_i} = 0$ for any $i \notin \mathcal{W}$. The result now follows. ∎

*Remark.* Lemma 2.3 holds for any distribution of the $\boldsymbol{\omega_i}$ so long as independence is assumed. Similar results are easily obtained for higher-order moments of the $\boldsymbol{f_{ij}}$.

**2.2. Formulating the Chance-Constrained Problem as a Conic Program.** In deriving the above formulation (2.10)–(2.16) for CC-OPF we assumed that the $\boldsymbol{\omega_i}$ random variables have zero mean. To obtain an efficient solution procedure we will additionally assume that they are (a) statistically independent, and (b) normally distributed. These assumptions were justified in section 1.5. Under the assumptions, however, since the $\boldsymbol{f_{ij}}$ are affine functions of the $\boldsymbol{\omega_i}$ (because the $\boldsymbol{\theta}$ are, by (2.6)), it turns out that there is a simple restatement of the chance constraints (2.14), (2.15), and (2.16) in a computationally practicable form. See [41] for a general treatment of linear inequalities with stochastic coefficients. For any real $0 < r < 1$ we write $\eta(r) = \phi^{-1}(1 - r)$, where $\phi$ is the cdf of a standard normally distributed random variable.

LEMMA 2.4. *Let $\bar{p}$, $\alpha$ be viable. Assume that the $\boldsymbol{\omega_i}$ are normally distributed and independent. Then the following hold:*

- *For any line $\{i, j\}$, $P(\boldsymbol{f_{ij}} > f_{ij}^{max}) < \epsilon_{ij}$ and $P(-\boldsymbol{f_{ij}} > f_{ij}^{max}) < \epsilon_{ij}$ iff*

$$\beta_{ij}|\bar{\theta}_i - \bar{\theta}_j| \le f_{ij}^{max} - \eta(\epsilon_{ij}) \left[\beta_{ij}^2 \sum_{k \in \mathcal{W}} \sigma_k^2 (\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2\right]^{1/2}, \tag{2.24}$$

*where as before $\bar{\theta} = \breve{\mathcal{B}}(\bar{p} + \mu - d)$ and $\delta = \breve{\mathcal{B}}\alpha$.*

• *For any generator* $g$, $P\left(\bar{p}_g - (e^T\boldsymbol{\omega})\alpha_i > p_g^{max}\right) < \epsilon_g$ *and* $P\left(\bar{p}_g - (e^T\boldsymbol{\omega})\alpha_i < p_g^{min}\right) < \epsilon_g$ *iff*

$$p_g^{min} + \eta(\epsilon_g)\left(\sum_{k\in\mathcal{W}}\sigma_k^2\right)^{1/2} \le \bar{p}_g \le p_g^{max} - \eta(\epsilon_g)\left(\sum_{k\in\mathcal{W}}\sigma_k^2\right)^{1/2}. \tag{2.25}$$

*Proof.* By Lemma 2.2, $\boldsymbol{f_{ij}}$ is an affine function of the $\boldsymbol{\omega_i}$; under the assumption it follows that $\boldsymbol{f_{ij}}$ is itself normally distributed. Thus, $P(\boldsymbol{f_{ij}} > f_{ij}^{max}) < \epsilon_{ij}$ iff

$$\mathbb{E}_{\boldsymbol{\omega}}\boldsymbol{f_{ij}} + \eta(\epsilon_{ij})\,\mathrm{var}(\boldsymbol{f_{ij}}) \le f_{ij}^{max}, \tag{2.26}$$

and similarly, $P(\boldsymbol{f_{ij}} < -f_{ij}^{max}) < \epsilon_{ij}$ iff

$$\mathbb{E}_{\boldsymbol{\omega}}\boldsymbol{f_{ij}} - \eta(\epsilon_{ij})\,\mathrm{var}(\boldsymbol{f_{ij}}) \ge -f_{ij}^{max}. \tag{2.27}$$

Lemma 2.2 gives $\mathbb{E}_{\boldsymbol{\omega}}\boldsymbol{f_{ij}} = \beta_{ij}(\bar{\theta}_i - \bar{\theta}_j)$, while by Lemma 2.3 $\mathrm{var}(\boldsymbol{f_{ij}}) = \beta_{ij}^2\sum_{k\in\mathcal{W}}\sigma_k^2(\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2$. Substituting these values into (2.26) and (2.27) yields (2.24). The proof of (2.25) is similar. ∎

*Remarks.* Equation (2.24) highlights the difference between, e.g., our chance constraint for lines, which requires that $P(\boldsymbol{f_{ij}} > f_{ij}^{max}) < \epsilon_{ij}$ and that $P(\boldsymbol{f_{ij}} < -f_{ij}^{max}) < \epsilon_{ij}$, and the stricter requirement that $P(|\boldsymbol{f_{ij}}| > f_{ij}^{max}) < \epsilon_{ij}$, which amounts to

$$P(\boldsymbol{f_{ij}} > f_{ij}^{max}) + P(\boldsymbol{f_{ij}} < -f_{ij}^{max}) < \epsilon_{ij}. \tag{2.28}$$

Unlike our requirement, which is captured by (2.24), the stricter condition (2.28) does not admit a compact statement.

We can now present a formulation of our chance-constrained optimization as a convex optimization problem on variables $\bar{p}, \alpha, \bar{\theta}, \delta$, and $s$. We will assume in what follows that for all lines $\{i,j\}$, $\epsilon_{ij} < 1/2$, so that $\eta(\epsilon_{ij}) > 0$:

$$\min \sum_{i\in\mathcal{G}}\left\{c_{i2}\bar{p}_i^2 + \left(\sum_k \sigma_k^2\right)c_{i2}\alpha_i^2 + c_{i1}\bar{p}_i + c_{i0}\right\}; \tag{2.29}$$

$$\text{for } 1 \le i \le n-1, \qquad \sum_{j=1}^{n-1}\hat{B}_{ij}\,\delta_j = \alpha_i; \tag{2.30}$$

$$\text{for } 1 \le i \le n-1, \quad \sum_{j=1}^{n-1}\hat{B}_{ij}\bar{\theta}_j - \bar{p}_i = \mu_i - d_i; \tag{2.31}$$

$$\sum_i \alpha_i = 1, \quad \alpha \ge 0, \quad \bar{p} \ge 0; \tag{2.32}$$

$$\bar{p}_n = \alpha_n = \delta_n = \bar{\theta}_n = 0; \tag{2.33}$$

$$\beta_{ij}|\bar{\theta}_i - \bar{\theta}_j| + \beta_{ij}\eta(\epsilon_{ij})\,s_{ij} \le f_{ij}^{max} \quad \forall\,\{i,j\}; \tag{2.34}$$

$$\left[\sum_{k\in\mathcal{W}}\sigma_k^2(\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2\right]^{1/2} - s_{ij} \le 0 \quad \forall\,\{i,j\}; \tag{2.35}$$

$$-\bar{p}_g + \eta(\epsilon_g)\left(\sum_{k\in\mathcal{W}} \sigma_k^2\right)^{1/2} \leq -p_g^{min} \quad \forall\, g \in \mathcal{G}; \tag{2.36}$$

$$\bar{p}_g + \eta(\epsilon_g)\left(\sum_{k\in\mathcal{W}} \sigma_k^2\right)^{1/2} \leq p_g^{max} \quad \forall\, g \in \mathcal{G}. \tag{2.37}$$

In this formulation, the variables $s_{ij}$ are auxiliary and introduced to facilitate the discussion below—since $\eta_{ij} \geq 0$, without loss of generality (2.35) will hold as an equality. Constraints (2.35), (2.36), and (2.37) are *second-order cone* inequalities [15]. A problem of the above form is solvable in polynomial time using well-known methods of convex optimization; several commercial software tools such as Cplex [32], Gurobi [29], and Mosek [39] are available. Constraint (2.30) is equivalent to $\delta_i = \sum_{1=j}^{n-1} \pi_{ij}\alpha_j$ (as we did in (2.19)); however, the $\pi_{ij}$ can be seen to be all nonzero, whereas $\hat{B}$ is very sparse for typical grids. Constraint (2.35) can be relatively dense—the sum has a term for each farm. However, as a percentage of the total number of buses this can be expected to be small.

**2.3. Solving the Conic Program.** Even though optimization theory guarantees that the above problem is efficiently solvable, experimental testing shows that in the case of large grids (thousands of lines) the problem proves challenging. For example, in the Polish 2003-2004 winter peak case,[9] we have 2746 buses, 3514 lines, and 8 wind farms, and Cplex [32] reports (after presolving) 36625 variables and 38507 constraints, of which 6242 are conic. On this problem, a recent version of Cplex on a (current) 8-core workstation ran for 3392 seconds (on 16 parallel threads, making use of "hyperthreading") and was unable to produce a feasible solution. On the same problem Gurobi reported "numerical trouble" after 31.1 cpu seconds, and stopped.

In fact, all of the commercial solvers [32, 29, 39] we experimented with reported numerical difficulties with problems of this size. Anecdotal evidence indicates that the primary cause for these difficulties is not simply the size, but also to a large degree *numerics* in particular poor conditioning due to the entries in the matrix $B$. These are susceptances, which are inverses of reactances, and often take values in a wide range.

To address this issue we implemented an effective algorithm for solving problem (2.29)–(2.37). For brevity we will focus on constraints (2.35) ((2.36) and (2.37) are similarly handled). For a line $\{i, j\}$ define

$$C_{ij}(\delta) \doteq \left(\sum_{k\in\mathcal{W}} \sigma_k^2(\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2\right)^{1/2}.$$

Constraint (2.35) can thus be written as $C_{ij}(\delta) \leq s_{ij}$. For completeness, we state the following result.

LEMMA 2.5. *Constraint* (2.35) *is equivalent to the infinite set of linear inequalities,*

$$C_{ij}(\hat{\delta}) + \frac{\partial C_{ij}(\hat{\delta})}{\partial \delta_i}(\delta_i - \hat{\delta}_i) + \frac{\partial C_{ij}(\hat{\delta})}{\partial \delta_j}(\delta_j - \hat{\delta}_j) \leq s_{ij} \quad \forall\, \hat{\delta} \in \mathbb{R}^n. \tag{2.38}$$

Constraints (2.38) express the *outer envelope* of the set described by (2.35) [15]. Any vector $\delta \in \mathbb{R}^n$ which satisfies (2.35) (for a given choice of $s_{ij}$) is guaranteed to

[9] Available with MATPOWER [56].

satisfy (2.38). Thus a finite subset of the inequalities (2.38), used instead of (2.35), will give rise to a relaxation of the optimization problem and thus a lower bound on the optimal objective value. Given Lemma 2.5 there are two ways to proceed, both motivated by the observation that at $\delta^* \in \mathbb{R}^n$ the most constraining inequality from among the set (2.38) (that is to say, the one whose left-hand side evaluated at $\delta^*$ is largest) is that obtained by choosing $\hat{\delta} = \delta^*$.

First, one can use inequalities (2.38) as *cutting planes* in the context of the ellipsoid method [28], obtaining a polynomial-time algorithm. A different way to proceed yields a numerically practicable algorithm. For a classical reference, see [34]. For brevity, we will omit treatment of the generator conic constraints (2.36), (2.37) (which are similarly handled). Denote by $F(\bar{p}, \alpha)$ the objective function in (2.29).

ALGORITHM 2.6. Cutting-Plane Algorithm.

*Initialization:* The linear "master" system $A(\bar{p}, \alpha, \delta, \theta, s)^T \geq b$ is defined to include constraints (2.30)–(2.34).

*Iterate:*

(1) Solve $\min\{F(\bar{p}, \alpha) \ : \ A(\bar{p}, \alpha, \delta, \theta, s)^T \geq b\}$. Let $(\bar{p}^*, \alpha^*, \delta^*, \theta^*, s^*)$ be an optimal solution.

(2) If all conic constraints are satisfied up to numerical tolerance by $(\bar{p}^*, \alpha^*, \delta^*, \theta^*, s^*)$, **Stop**.

(3) If all chance constraints are satisfied up to numerical tolerance by $(\bar{p}^*, \alpha^*)$, **Stop**.

(4) Otherwise, add to the master system the outer inequality (2.38) arising from that constraint (2.35) which is most violated.

As the algorithm iterates, the master system represents a valid relaxation of the conic program (2.30)–(2.35); thus the objective value of the solution computed in step (1) is a valid lower bound on the value of the problem. Each problem solved in step (1) is a linearly constrained, convex quadratic program. Computational experiments involving large-scale realistic cases show that the algorithm is robust and rapidly converges to an optimum.

Note that step (3) is not redundant. Recall that our formulation includes the variables $s_{ij}$; each such variable is lower bounded (via (2.35)) by the variance of flow on line $\{i, j\}$. However, in the cutting-plane algorithm we do not make explicit use of (2.35)—indeed, we progressively approximate it by means of cutting planes. At a typical intermediate iteration the current solution will still fail to satisfy (2.35); i.e., we may have, for some line (or lines) $\{i, j\}$,

$$\left[\sum_{k \in \mathcal{W}} \sigma_k^2 (\pi_{ik} - \pi_{jk} - \delta_i^* + \delta_j^*)^2\right]^{1/2} > s_{ij}^* \tag{2.39}$$

(where "*" indicates the current values of the variables), which would normally indicate that the algorithm has not terminated because condition (2.35) has not been adequately represented by the cutting planes. However, in later stages of the algorithm it may be the case that (2.39) holds for some lines, and yet the pair $(\bar{p}^*, \alpha^*)$ already satisfies the chance constraints for all lines. This situation will arise if in

**Table 2.1** *Performance of cutting-plane method on a number of large cases.*

| Case | Buses | Generators | Lines | Time (s) | Iterations | Barrier iterations |
|---|---|---|---|---|---|---|
| BPA | 2209 | 176 | 2866 | 5.51 | 2 | 256 |
| Polish1 | 2383 | 327 | 2896 | 13.64 | 13 | 535 |
| Polish2 | 2746 | 388 | 3514 | 30.16 | 25 | 1431 |
| Polish3 | 3120 | 349 | 3693 | 25.45 | 23 | 508 |

the current solution some of the $s^*_{ij}$ quantities are artificially low, i.e., they could be increased so as to match the right-hand side of (2.39) without jeopardizing feasibility of the rest of the solution, i.e., constraint (2.34). The underlying dynamic is that the conic constraints (2.35) are still somewhat loosely (outer-)approximated by a set of linear inequalities, resulting in a range of possible values for a particular variable $s_{ij}$. One such value will be picked by the underlying quadratic programming solver, and, as just argued, this value may be artificially low. A direct way to check that $(\bar{p}^*, \alpha^*)$ satisfies the chance constraint for a given line $\{i, j\}$ is straightforward—the flows $\boldsymbol{f_{ij}}$ are normally distributed (as noted in Lemma 2.4), and their means and variances can be directly computed from $(\bar{p}^*, \alpha^*)$, and so we can directly compute the probability that, under $(\bar{p}^*, \alpha^*)$, any given $\boldsymbol{f_{ij}}$ exceeds its corresponding line limit (additionally, this procedure computes said probability, a useful output parameter). If $(\bar{p}^*, \alpha^*)$ indeed satisfies all chance constraints, it is optimal, since it attains a cost that matches a lower bound to the overall problem (i.e., it attains the value of the current cutting-plane relaxation).

In our implementation, termination is declared in step (2) or step (3) when the corresponding constraint violation is less than $10^{-6}$. Table 2.1 displays typical performance of the cutting-plane algorithm on (comparatively more difficult) large problem instances. In the table, "Polish1"–"Polish3" are the three Polish cases included in MATPOWER [56] (in Polish1 we increased loads by 30%). All Polish cases have uniform random costs on [0.5, 2.0] for each generator and ten arbitrarily chosen wind sources. The average wind power penetration for the four cases is 8.8%, 3.0%, 1.9%, and 1.5%. "Iterations" is the number of linearly constrained subproblems solved before the algorithm converges. "Barrier iterations" is the total number of iterations of the barrier algorithm in CPLEX over all subproblems, and "Time" is the total (wallclock) time required by the algorithm. Line tolerances are set to two standard deviations and generator tolerances three standard deviations (tail probabilities 0.0228 and 0.00135, respectively). These instances all prove unsolvable if directly tackled by CPLEX or Gurobi. A more challenging test involving the Polish grid is reported in section 3.6.

Table 2.2 provides additional, typical numerical performance for the cutting-plane algorithm on an instance of the Polish grid model. Each row of Table 2.2 shows the maximum relative error and objective value at the end of several iterations. The total run-time was 25 seconds. Note the "flatness" of the objective. This makes the problem nontrivial—the challenge is to find a *feasible* solution (with respect to the chance constraints); at the onset of the algorithm the computed solution is quite infeasible, and it is this attribute that is quickly improved by the cutting-plane algorithm.

We note the (typical) small number of iterations needed to attain numerical convergence. Thus at termination only a very small number of conic constraints (2.35) have been incorporated into the master system. This validates the expectation that only a small fraction of the conic constraints in CC-OPF are active at optimality. The

**Table 2.2** *Typical convergence behavior of cutting-plane algorithm on a large instance.*

| Iteration | Max rel. error | Objective |
|---|---|---|
| 1 | 1.2e-1 | 7.0933e6 |
| 4 | 1.3e-3 | 7.0934e6 |
| 7 | 1.9e-3 | 7.0934e6 |
| 10 | 1.0e-4 | 7.0964e6 |
| 12 | 8.9e-7 | 7.0965e6 |

cutting-plane algorithm can be viewed as a procedure that opportunistically discovers these constraints.

**2.4. Data-Robust Chance Constraints.** Above we developed a formulation for our CC-OPF problem as the conic program (2.29)–(2.37). This approach assumed exact estimates for the mean $\mu_i$ and the variance $\sigma_i^2$ of each wind source $\boldsymbol{\omega_i}$. In practice, however, the estimates at hand might be imprecise,[10] consequently jeopardizing the usefulness of our conic program, henceforth termed the *nominal* chance-constrained problem. In particular, the performance of the control computed by the conic program might conceivably be sensitive to even small changes in the data. We will deal with this issue in two complimentary ways.

**2.4.1. Out-of-Sample Analysis.** Our first approach is the out-of-sample tests, implemented experimentally in section 3.5. We assume that the $\mu_i$ and $\sigma_i^2$ have been mis-estimated and explore the robustness of the affine control with respect to the estimation errors. The experiments of section 3.5 show that the degradation of the chance constraints is small when small data errors are experienced. This empirical observation has a rigorous explanation discussed below.

Our chance constraints are represented by convex inequalities, for example, in the case of $P(\boldsymbol{f_{ij}} > f_{ij}^{max}) < \epsilon_{ij}$ and $P(\boldsymbol{f_{ij}} < -f_{ij}^{max}) < \epsilon_{ij}$ we used (2.34) and (2.35). Suppose that we have solved the chance-constrained problem assuming (incorrect) expectations $\mu_i$ and variances $\sigma_i^2$ $(i \in \mathcal{W})$. Let $\tilde{\mu}_i$ and $\tilde{\sigma}_i^2$ $(i \in \mathcal{W})$ be the exact (or realized) values. With some abuse of notation, we will write $\xi$ (resp., $\tilde{\xi}$) for the incorrect (exact) distribution. For a given line $\{i, j\}$, write

$$m_{ij} = \mathbb{E}_{\boldsymbol{\xi}} \boldsymbol{f_{ij}} = \beta_{ij}([\breve{\mathcal{B}}(\bar{p} + \mu - d)]_j - [\breve{\mathcal{B}}(\bar{p} + \mu - d)]_j), \tag{2.40}$$

$$\sigma_{ij}^2 = \mathrm{var}_{\xi} \boldsymbol{f_{ij}} = \beta_{ij}^2 \sum_{k \in \mathcal{W}} \sigma_k^2 (\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2, \tag{2.41}$$

and similarly,

$$\tilde{m}_{ij} = \mathbb{E}_{\tilde{\boldsymbol{\xi}}} \boldsymbol{f_{ij}} = \beta_{ij}([\breve{\mathcal{B}}(\bar{p} + \tilde{\mu} - d)]_j - [\breve{\mathcal{B}}(\bar{p} + \tilde{\mu} - d)]_j), \tag{2.42}$$

$$\tilde{\sigma}_{ij}^2 = \mathrm{var}_{\tilde{\xi}} \boldsymbol{f_{ij}} = \beta_{ij}^2 \sum_{k \in \mathcal{W}} \tilde{\sigma}_k^2 (\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2. \tag{2.43}$$

Using (2.42) and (2.43) we see that the "true" probability $P(\boldsymbol{f_{ij}} > f_{ij}^{max})$ is that value $\tilde{\epsilon}$ such that

$$\tilde{m}_{ij} + \eta(\tilde{\epsilon})\tilde{\sigma}_{ij} = f_{ij}^{max}. \tag{2.44}$$

[10] In particular since they would effectively be computed in real time.

We wish to evaluate how much *larger* this (realized) value $\tilde{\epsilon}$ is compared with the target value $\epsilon_{ij}$, which was the goal in the chance-constrained computation. We will do this assuming that the estimation errors are small; more precisely, there exist nonnegative constants $M$ and $V$ such that

$$\forall i \in \mathcal{W}, \quad |\tilde{\mu}_i - \mu_i| < M\mu_i \quad \text{and} \quad |\tilde{\sigma}_i^2 - \sigma_i^2| < V^2\sigma_i^2. \tag{2.45}$$

Considering (2.44), we see that for data errors of a given magnitude, $\tilde{\epsilon}$ is maximized when $\tilde{m}_{ij}$ and $\tilde{\sigma}_{ij}$ are maximized. Further, considering (2.42) and (2.43) we see that to first order $\tilde{m}_{ij} \le m_{ij} + O(M)$, and that $\tilde{\sigma}_{ij}^2 \le (1+V^2)\sigma_{ij}^2$. From these two observations and (2.44) we obtain

$$\begin{aligned}\tilde{\epsilon} &= \frac{1}{\sqrt{2\pi}} \int_{\eta(\tilde{\epsilon})}^{+\infty} e^{-x^2/2} dx = \epsilon_{ij} + \frac{1}{\sqrt{2\pi}} \int_{\eta(\tilde{\epsilon})}^{\eta(\epsilon_{ij})} e^{-x^2/2} dx \\ &= \epsilon_{ij} + \frac{1}{\sqrt{2\pi}} e^{-\frac{\eta(\epsilon_{ij})^2}{2}} [\eta(\epsilon_{ij}) - \eta(\tilde{\epsilon})] + \text{smaller order errors.} \end{aligned} \tag{2.46}$$

The quantity in brackets in the right-hand side of (2.46) equals

$$\eta(\epsilon_{ij}) - \frac{f_{ij}^{max} - m_{ij}}{\tilde{\sigma}_{ij}} + \frac{\tilde{m}_{ij} - m_{ij}}{\tilde{\sigma}_{ij}} \tag{2.47}$$

$$< \eta(\epsilon_{ij}) \left(1 - \frac{1}{\sqrt{1+V^2}}\right) + \frac{\tilde{m}_{ij} - m_{ij}}{\tilde{\sigma}_{ij}} \tag{2.48}$$

$$< \eta(\epsilon_{ij}) O(V) + \frac{O(M)}{\sigma_{ij}}. \tag{2.49}$$

Using (2.46) and (2.49) we obtain that $\tilde{\epsilon} = \epsilon_{ij} + [\eta(\epsilon_{ij})O(V) + O(M)]e^{-\frac{\eta(\epsilon_{ij})^2}{2}}$, where the "$O$" notation contains solution-dependent constants.

**2.4.2. Efficiently Solvable Data-Robust Formulations.** As the preceding analysis makes clear, the constants $M$ and $V$ may need to be quite small, for example, if the $\sigma_{ij}$ are small. We thus seek a better guarantee of robustness. This justifies our second approach discussed below—to develop a version of CC-OPF which is methodologically guaranteed to be insensitive to data errors. This approach puts CC-OPF within the scope of robust optimization; to be more precise we will be solving an ambiguous chance-constrained problem in the language of [23].

We will write each $\mu_i$ in the form $\bar{\mu}_i + r_i$, where the $\bar{\mu}_i$ are point estimates of the $\mu_i$ and the $r_i$ are "errors" which are constrained to lie in some fixed set $\mathcal{M} \subseteq \mathbb{R}^{|\mathcal{W}|}$ with $0 \in \mathcal{M}$. Likewise, we assume that there is a set $\mathcal{S} \subseteq \mathbb{R}^{|\mathcal{W}|}$ including 0, such that each $\sigma_i^2$ is of the form $\bar{\sigma}_i^2 + v_i$ where the vector of errors $v_i$ belongs to $\mathcal{S}$. As an example for how to construct $\mathcal{M}$ or $\mathcal{S}$, one can use the following set parameterized by values $0 < \gamma_i$ and $0 < \Gamma$:

$$U(\gamma, \Gamma) = \left\{ r \in \mathbb{R}^{\mathcal{W}} \,:\, |r_i| \le \gamma_i \; \forall i \in \mathcal{W}, \;\; \sum_{i \in \mathcal{W}} \frac{|r_i|}{\gamma_i} \le \Gamma \right\}. \tag{2.50}$$

This set was introduced in [9]. Another candidate is the ellipsoidal set

$$E(A, b) = \left\{ r \in \mathbb{R}^{|\mathcal{W}|} \,:\, r^T A r \le b \right\}, \tag{2.51}$$

where $A \in \mathbb{R}^{|\mathcal{W}\times\mathcal{W}|}$ is positive definite and $b \geq 0$ is a real; see [5, 25]. We can now formally proceed as follows.

DEFINITION 2.2. *Let the estimates $\bar{\mu}$, $\bar{\sigma}^2$ and the sets $\mathcal{M}$ and $\mathcal{S}$ be given.*

1. *For each pair $r \in \mathcal{M}$ and $v \in \mathcal{S}$ we will consider a random variable $\xi = \xi(r, v)$ with coordinatewise mean $\mu(\xi) = \bar{\mu} + r$ and variance $\sigma^2(\xi) = \bar{\sigma}^2 + v$. We term $\xi$ a* realization*, and we denote by $\mathcal{D} = \mathcal{D}(\bar{\mu}, \mathcal{M}, \bar{\sigma}^2, \mathcal{S})$ the set of all realizations.*

2. *A pair $\bar{p}, \alpha$ is called robust with respect to the pair $\mathcal{M}$, $\mathcal{S}$ if for each line $\{i, j\}$*

$$\max_{\xi\in\mathcal{D}} P_\xi(\boldsymbol{f_{ij}} > f_{ij}^{max}) < \epsilon_{ij} \quad \textit{and} \tag{2.52}$$

$$\max_{\xi\in\mathcal{D}} P_\xi(\boldsymbol{f_{ij}} < -f_{ij}^{max}) < \epsilon_{ij}, \tag{2.53}$$

*where we denote by $P_\xi$ the probability function under realization $\xi$.*

Our task will be to replace, in our optimization problem formulation, the chance constraint (1.13) with one asking for robustness as in (2.52)–(2.53).[11] If the uncertainty sets $\mathcal{M}$ and $\mathcal{S}$ consist of a single point estimate each, then we recover the nominal chance-constrained problem we discussed above. As $\mathcal{M}$ or $\mathcal{S}$ become larger, the robust approach becomes more insensitive to estimation errors, albeit at the cost of becoming more conservative. Thus, a reasonable balance should be attained by choosing $\mathcal{M}$ and $\mathcal{S}$ small but of positive measure, thereby preventing trivial sensitivity of the control to changes in the data.

To indicate our approach, we focus on one of the statements for our chance constraint for lines.

Consider a particular realization $\xi \in \mathcal{D}$ and denote by $\bar{\theta}(\xi)$ the vector of average phase angles under $\xi$ (below we will provide an expression for this vector). Robustness criterion (2.52) applied to a given line $\{i, j\}$ requires that the chance constraint $P_\xi(\boldsymbol{f_{ij}} > f_{ij}^{max}) < \epsilon_{ij}$ hold. By Lemma 2.4 this statement can be equivalently phrased as

$$\beta_{ij}\left\{\bar{\theta}_i(\xi) - \bar{\theta}_j(\xi) + \eta(\epsilon_{ij})\left[\sum_{k\in\mathcal{W}} \sigma_k^2(\xi)(\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2\right]^{1/2}\right\} \leq f_{ij}^{max}. \tag{2.54}$$

It follows that robustness criterion (2.52) can be succinctly stated as

$$\beta_{ij}\max_{\xi\in\mathcal{D}}\left\{\bar{\theta}_i(\xi) - \bar{\theta}_j(\xi) + \eta(\epsilon_{ij})\left[\sum_{k\in\mathcal{W}} \sigma_k^2(\xi)(\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2\right]^{1/2}\right\} \leq f_{ij}^{max}. \tag{2.55}$$

We can see that (2.55) consists of a (possibly infinite) set of convex constraints; thus the data-robust chance-constrained problem is a convex problem. However, we need to exploit this fact in a computationally favorable manner. Toward this end, our next task is to obtain a more convenient restatement for (2.54). Given a specific $\xi \in \mathcal{D}$, recall that by definition we have

$$\mu(\xi) = \bar{\mu} + r \quad \text{and} \quad \sigma^2(\xi) = \bar{\sigma}^2 + v \tag{2.56}$$

[11] And likewise with the generator chance constraints (2.36), (2.37).

for some $r \in \mathcal{M}$ and $v \in \mathcal{S}$. Hence

$$\bar{\theta}(\xi) = \breve{\mathcal{B}}(\bar{p} + \mu(\xi) - d) = \breve{\mathcal{B}}(\bar{p} + \bar{\mu} - d) \ + \breve{\mathcal{B}}r = \bar{\theta}(\bar{\mu}) \ + \breve{\mathcal{B}}r. \tag{2.57}$$

Thus, we see that (2.54) can be equivalently rewritten as

$$\begin{aligned} &\beta_{ij}(\bar{\theta}(\bar{\mu})_i - \bar{\theta}(\bar{\mu})_j) + \beta_{ij} e_{ij}^T \breve{\mathcal{B}} r \\ &+ \eta(\epsilon_{ij})\beta_{ij} \left[ \sum_{k \in \mathcal{W}} \bar{\sigma}_k^2 (\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2 + \sum_{k \in \mathcal{W}} v_k (\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2 \right]^{1/2} \\ &\leq f_{ij}^{max}. \end{aligned} \tag{2.58}$$

Here, $e_{ij}^T \in \mathbb{R}^n$ is the vector with a $+1$ entry in position $i$, a $-1$ entry in position $j$, and 0 elsewhere. In order for (2.52) to hold, (2.58) must hold for all $\xi \in \mathcal{D}$; in other words, it must hold for all pairs $(r, v)$ with $r \in \mathcal{M}$ and $v \in \mathcal{S}$. It follows that criterion (2.52) can be stated as

$$\begin{aligned} &\beta_{ij}(\bar{\theta}(\bar{\mu})_i - \bar{\theta}(\bar{\mu})_j) + \beta_{ij} \max_{r \in \mathcal{M}} \left\{ e_{ij}^T \breve{\mathcal{B}} r \right\} \\ &+ \eta(\epsilon_{ij})\beta_{ij} \left[ \sum_{k \in \mathcal{W}} \bar{\sigma}_k^2 (\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2 + \max_{v \in \mathcal{S}} \left\{ \sum_{k \in \mathcal{W}} v_k (\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2 \right\} \right]^{1/2} \\ &\leq f_{ij}^{max}. \end{aligned} \tag{2.59}$$

Note that if in the left-hand side of (2.59) we ignore the term involving $r$ and the second term inside the square brackets, we obtain the nominal (i.e., nonrobust) version of chance constraint (2.52); see (2.34), (2.35). A similar constraint (with the $\bar{\theta}$ terms switched in sign) is obtained from equations (2.53).

Considering (2.59), we see that the maximum over $\mathcal{M}$ is independent of all decision variables and can be computed in advance. In what follows we write, for each line $\{i, j\}$,

$$R_{ij} \doteq \max_{r \in \mathcal{M}} \left\{ e_{ij}^T \breve{\mathcal{B}} r \right\}. \tag{2.60}$$

The second maximum in (2.59) presents a challenge. In a traditional robust optimization approach the next task would be to represent that maximum with an equivalent, compact (i.e., moderate-size) system of equivalent convex constraints. Such a transformation would rely on linear programming duality in the case of the uncertainty model $U(\gamma, \Gamma)$, and on the conic duality or the S-lemma [15, 25] in the case of the ellipsoidal model $E(A, b)$. However, it can be shown that in the case of (2.59) such an approach will fail—it will produce a large formulation, which is additionally nonconvex. We refer the reader to the appendix for a complete discussion. We next describe an alternate approach that proves efficient.

**2.4.3. Efficient Solution of the Data-Robust Problem Using Cutting Planes.** To derive an algorithm for the data-robust chance-constrained problem that (a) has some theoretical justification and (b) can prove numerically tractable, we note that if we replace the set $\mathcal{S}$ with a finite subset $\tilde{\mathcal{S}} \subseteq \mathcal{S}$ we obtain a valid relaxation for the optimization problem. In other words the system made up of the following constraints, for each line $\{i, j\}$,

$$\beta_{ij}(\bar{\theta}(\bar{\mu})_i - \bar{\theta}(\bar{\mu})_j) + \beta_{ij} R_{ij} + \beta_{ij} \eta(\epsilon_{ij})\, s_{ij} \leq f_{ij}^{max}, \tag{2.61}$$

$$\beta_{ij}(\bar{\theta}(\bar{\mu})_j - \bar{\theta}(\bar{\mu})_i) - \beta_{ij}R_{ij} + \beta_{ij}\eta(\epsilon_{ij})\, s_{ij} \le f_{ij}^{max}, \tag{2.62}$$

$$\left[\sum_{k\in\mathcal{W}} (\bar{\sigma}_k^2 + v_k)(\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2\right]^{1/2} \le s_{ij} \quad \forall\, v \in \mathcal{S}, \tag{2.63}$$

constitutes a valid relaxation of the chance constraints (2.34), (2.35) of the nominal formulation for each line $\{i, j\}$ for any given finite $\tilde{\mathcal{S}} \subseteq \mathcal{S}$. This observation can be used to formally obtain a polynomial-time algorithm for the data-robust problem in the cases of interest. For a given $v \in \mathcal{S}$ let

$$L_{ij}^{v}(\delta) \doteq \left(\sum_{k\in\mathcal{W}} (\bar{\sigma}_k^2 + v_k)(\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2\right)^{1/2}$$

(i.e., the expression inside the "max" in (2.63)). For completeness, we state the following lemma.

LEMMA 2.7. *In the case of uncertainty sets of the form $U(\gamma, \Gamma)$ or $E(A, b)$ the data-robust chance-constrained problem can be solved in polynomial time.*

*Proof.* Suppose we are given quantities $\hat{\delta}_i$ for each $i \in \mathcal{V}$ and $\hat{s}_{ij}$ for each line $\{i, j\}$. Then, as argued before, if $\mathcal{S}$ is of the form $U(\gamma, \Gamma)$ or $E(A, b)$ one can check in polynomial time whether $\max_{v\in\mathcal{S}} \left\{L_{i,j}^{v}(\delta)\right\} \le \hat{s}_{ij}$. If the condition is violated for $\bar{v} \in \mathcal{S}$, then trivially

$$L_{ij}^{\bar{v}}(\hat{\delta}) + \frac{\partial L_{ij}^{\bar{v}}(\hat{\delta})}{\partial \delta_i}(\delta_i - \hat{\delta}_i) + \frac{\partial L_{ij}^{\bar{v}}(\hat{\delta})}{\partial \delta_j}(\delta_i - \hat{\delta}_i) \le s_{ij} \tag{2.64}$$

is violated at $\hat{\delta}, \hat{s}$. Since (2.64) is valid for the data-robust chance-constrained problem (by convexity) the result follows by relying on the ellipsoid method [28]. ∎

Lemma 2.7 describes a formally "good" algorithm. For computational purposes we would instead rely on a cutting-plane algorithm much like Algorithm 2.6 developed in section 2.3. Details are provided in the appendix.

**3. Experiments/Results.** Here we will describe qualitative aspects of our affine control on small systems; in particular we focus on the contrast between standard OPF (meaning standard DC OPF here and below) and nominal CC-OPF, on problematic features that can arise because of fluctuating wind sources and on out-of-sample testing of the CC-OPF solution, including the analysis of non-Gaussian distributions. Some of our tests involve the BPA grid and Polish grid, which are large; we present additional sets of tests to address the scalability of our solution methodology to the large cases. Additional experiments are provided in the appendix, section 5.3.

Above (see (1.7)) we introduced the so-called standard OPF method for setting traditional generator output levels. When renewables are present, the natural extension of this approach would make use of some fixed estimate of output (e.g., mean output) and to handle fluctuations in renewable output through the same method used to deal with changes in load: ramping output of traditional generators up or down in proportion to the net increase or decrease in renewable output. This feature could seamlessly be handled using today's control structure, with each generator's output adjusted at a fixed (preset) rate. For the sake of simplicity, in the experiments below we assume that all droop constants are equal.

Different assumptions on these fixed rates will likely produce different numerical results; however, this general approach entails an inherent weakness. The key

point here is that mean generator output levels *as well as* in particular the ramping rates would be chosen *without* considering the stochastic nature of the renewable output levels. Our experiments are designed to highlight the limitations of this "risk-unaware" approach. In contrast, our CC-OPF produces control parameters (the $\bar{p}$ and the $\alpha$) that are risk-aware and, implicitly, also topology-aware—in the sense of network proximity to wind farms.

As a technical point, we note that generator adjustment under standard OPF can be viewed as a special case of our affine control mechanism, but with fixed $\alpha$ values. Above (Lemmas 2.2 and 2.3) we have provided expressions for the mean and standard deviation of the flow on any given line, under the independence assumption for the $\boldsymbol{\omega_i}$. It follows that under the Gaussianity assumption, for any given vectors $\bar{p}$ and $\alpha$ we can compute the probability that any given line $\{i, j\}$ is overloaded (both under CC-OPF and standard OPF): if $z$ is a standard normal variable, and $a > 0$, $\text{Prob}(z > a) = .5 * (1 - \text{erf}(z/\sqrt{2}))$, where erf is the error function [2].

**3.1. Failure of Standard OPF.** We first consider the IEEE 118-bus model with a quadratic cost function, and four sources of wind power added at arbitrary buses to meet 5% of demand in the case of average wind. The standard OPF solution is safely within the thermal capacity limits for all lines in the system. Then we account for fluctuations in wind assuming Gaussian and site-independent fluctuations with standard deviations set to 30% of the respective means. The results, which are shown in Figure 3.1, illustrate that under standard OPF five lines (marked in red) frequently become overloaded, exceeding their limits 8% or more of the time. This situation translates into an unacceptably high risk of failure for any of the five red lines. This problem occurs for grids of all sizes; similar results hold on the 2746-bus

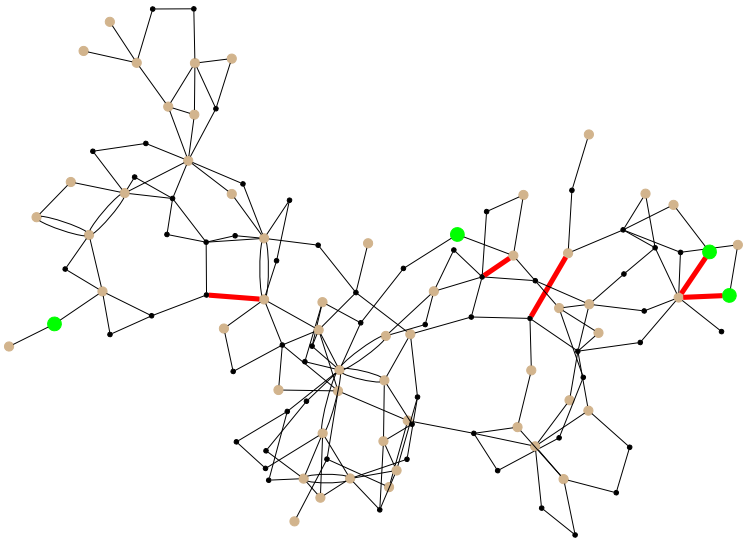

**Fig. 3.1** *118-bus case with four wind farms (green dots; brown are generators, black are loads). Shown is the standard OPF solution against the average wind case with penetration of* 5%. *Standard deviations of the wind are set to* 30% *of the respective average cases. Lines in red exceed their limit* 8% *or more of the time.*

Polish grid (from MATPOWER [56]). In this case, after scaling up all loads by 10% to simulate a more highly stressed system, we added wind power to ten buses for a total of 2% penetration. The standard solution results in six lines exceeding their limits over 45% of the time, and in one line over 10% of the time. For an additional and similar experiment using the Polish grid, see section 3.6.

**3.2. Cost of Reliability under High Wind Penetration.** If we maintain the current operational paradigm, congestion of transmission lines may force temporary shutdown of wind farms even during times of high wind. Our methodology suggests, as an alternative solution to curtailment of wind power, an appropriate reconfiguration of standard generators. If successful, this solution can use the available wind power without curtailment, and thus result in cheaper operating costs.

As a (crude) proxy for curtailment, we perform the following experiment, which considers different levels of renewable penetration. Here, the mean power outputs of the wind sources are kept in a fixed proportion to one another and proportionally scaled so as to vary the total amount of penetration, and likewise with the standard deviations. Under a high penetration level (e.g., 30%), the *standard* OPF solution may yield significant probability of line overloads. In order to obtain comparisons with CC-OPF, we perform experiments that involve two controls: (1) increasing line limits by some proportion and (2) decreasing wind penetration. Assuming zero cost for wind power, the difference in cost for a high-penetration CC-OPF solution and the low-penetration standard solution represents the savings produced by our model (generously, given the line limit increase).

We performed experiments involving the 39-bus case. For this test we used the standard cost function associated with this case (provided with MATPOWER), which is $\sum_{g\in\mathcal{G}} \left(0.01\, p_g^2 + 0.3\, p_g\right)$ (we are ignoring constant terms), and $\epsilon_{ij} = 0.01$ for all lines. Table 3.1 summarizes the outcome of this experiment. In the table, "OPF type" refers to either standard ("STD") or CC-OPF, "Line limit boost" is the percentage by which line limits are increased, and "Overload prob." is the maximum line overload probability. Thus, in the case of 30% penetration, standard OPF experiences high probability of line overloads, even with a 10% increase in line limits (two lines are overloaded with very high probability). Reducing penetration to 5% removes this problem, but at a major cost increase. In contrast, even under 40% penetration, CC-OPF achieves low overload probability and a significant cost decrease.

**Table 3.1** *Impact of penetration,* 39*-bus case.*

| OPF type | Penetration | Line limit boost | Overload prob. | Cost |
|---|---|---|---|---|
| STD | 30% | 10% | 21.8% | $2.05 \times 10^4$ |
| STD | 5% | 0% | 0.0% | $3.72 \times 10^4$ |
| CC | 30% | 0% | 0.3% | $2.16 \times 10^4$ |
| CC | 40% | 0% | 0.5% | $1.80 \times 10^4$ |

**3.3. Nonlocality.** We have established that under fluctuating power generation, some lines may exceed their flow limits an unacceptable fraction of the time. Is there a simple solution to this problem, for instance, by carefully adjusting (a posteriori of the standard OPF) the outputs of the generators near the violated lines? The answer is no. Power systems exhibit significant nonlocal behavior. Consider Figure 3.2.

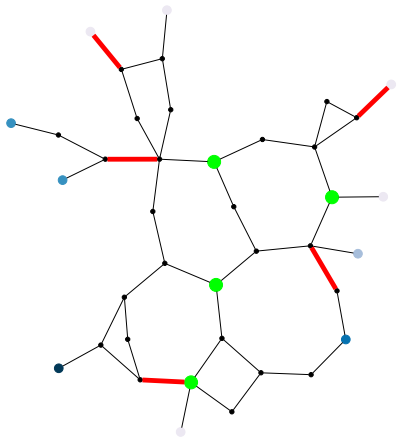

**Fig. 3.2** *39-bus case. Red lines indicate high probability of flow exceeding the limit under the standard OPF solution. Generators are shades of blue, with darker shades indicating greater absolute difference between the chance-constrained solution and the standard solution.*

In this example, the major differences in generator outputs between the standard OPF solution and our CC-OPF model's solution are not obviously associated with the different line violations. Notice that nonlocality of the power flow response is not specific to CC-OPF, and it may also be seen in experiments with standard OPF testing sensitivity of power flows, e.g., to the generation shift factors. However, our point in here is that nonlocality caused by stochastic factors, accounted for in CC-OPF, may be of a new type which does not allow intuitive explanation in terms of the standard OPF. In general, it seems that it would be difficult to bypass CC-OPF and make small local adjustments to relieve the stressed lines. On the positive side, even though CC-OPF is not local and requires a centralized computation, it is also only slightly more difficult than the standard OPF in terms of implementation.

**3.4. Increasing Penetration.** Current plans for the power system in the United States call for 30% of wind energy penetration by 2030 [1]. Investments necessary to achieve this ambitious target may focus on both software (improving operations) and hardware (building new lines, substations, etc.), with the former obviously representing a much less expensive and thus economically attractive option. Our CC-OPF solution contributes toward this option. A natural question that arises concerns the maximum level of penetration one can safely achieve by upgrading from the standard OPF to our CC-OPF.

To answer the question we consider the 39-bus New England system (from [56]) case with four wind generators added, and line flow limits scaled by .7 to simulate a heavily loaded system. The quadratic cost terms (the $c_{i2}$ coefficients in (2.17), which are the leading coefficients for the quadratic functions $c_i$) are set to rand(0,1) + .5. We fix the four wind generator average outputs in a ratio of 5/6/7/8 and standard deviations at 30% of the mean. We first solve our model using $\epsilon = .02$ for each line and assuming zero wind power, and then increase total wind output until the optimization problem becomes infeasible. See Figure 3.3. This experiment illustrates that at least for the model considered, the 30% of wind penetration with rather strict probabilistic guarantees enforced by our CC-OPF may be feasible, but in fact lies rather close to the dangerous threshold. Pushing penetration beyond the threshold is

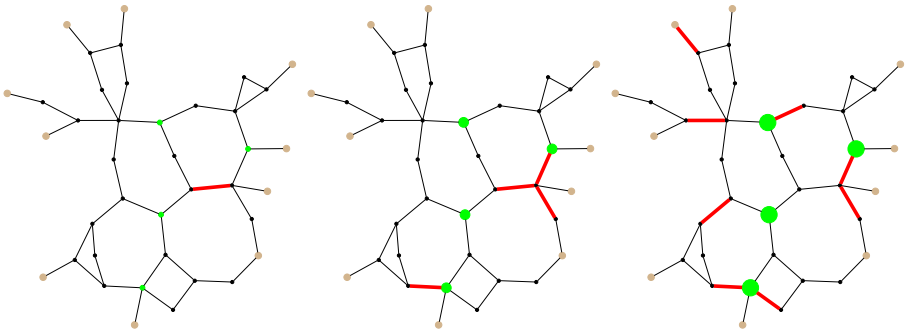

**Fig. 3.3** *39-bus case with four wind farms (green dots; brown are generators, black are loads). Lines in red are at the maximum of $\epsilon = .02$ chance of exceeding their limit. The three cases shown are, left to right, .1%, 8%, and 30% average wind penetration. With penetration beyond 30% the problem becomes infeasible.*

impossible without upgrading lines and investing in other (not related to wind farms themselves) hardware.

**3.5. Out-of-Sample Tests.** We now study the performance of the control computed using *nominal* CC-OPF when there are errors in the underlying distribution of wind power. We consider two types of errors: (1) the true distribution is non-Gaussian but our Gaussian fit is "close" in an appropriate sense, and (2) the true distribution is Gaussian but with different mean or standard deviation. The experiments in this section use as data set the BPA grid, which as noted before has 2209 buses and 2866 lines, and collected wind data, altogether constituting a realistic test case.

We first consider the non-Gaussian case, using the following probability distributions, all with fatter tails than Gaussian: (1) Laplace, (2) logistic, (3) Weibull (three different shapes), (4) t location-scale with 2.5 degrees of freedom, (5) Cauchy. For the Laplace and logistic distributions, we simply match the mean and standard deviation. For the Weibull distribution, we consider shape parameters $k = 1.2, 2, 4$ and choose the scale parameter to match the standard deviations. We then translate to match means. For the t distribution, we fix 2.5 degrees of freedom and then choose the location and scale to match mean and standard deviation. For the Cauchy distribution, we set the location parameter to the mean and choose the scale parameter so as to match the 95th percentiles.

We use our model and solve under the Gaussian assumption, seeking a solution which results in no line violations for cases within two standard deviations of the mean, i.e., a maximum of about 2.27% chance of exceeding the limit. We then perform Monte Carlo tests drawing from the above distributions to determine the actual rates of violation. See Table 3.2. The worst performer is the highly asymmetric (and perhaps unreasonable) Weibull with shape parameter 1.2, which approximately doubles the desired maximum chance of overload. Somewhat surprisingly, the fat-tailed logistic and student's t distributions result in a maximum chance of overload significantly less than desired, suggesting that our model is too conservative in these cases.

Next we consider the Gaussian case with errors. We solve with nominal values for the mean and standard deviation of wind power. We then consider the rate of violation after scaling all means and standard deviations (separately). While the solution is sensitive to errors in the mean forecast, the sensitivity is well behaved.

**Table 3.2** *Maximum probability of overload for out-of-sample tests. These are a result of Monte Carlo testing with* 10,000 *samples on the BPA case, solved under the Gaussian assumption and desired maximum chance of overload at* 2.27%.

| Distribution | Max. prob. violation |
|---|---|
| Normal | 0.0227 |
| Laplace | 0.0297 |
| Logistic | 0.0132 |
| Weibull, $k = 1.2$ | 0.0457 |
| Weibull, $k = 2$ | 0.0355 |
| Weibull, $k = 4$ | 0.0216 |
| t location-scale, $\nu = 2.5$ | 0.0165 |
| Cauchy | 0.0276 |

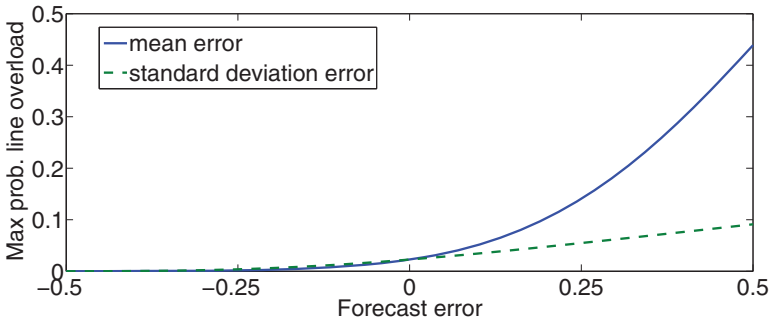


**Fig. 3.4** *BPA case solved with average penetration at* 8% *and standard deviations set to* 30% *of mean. The maximum probability of line overload desired is* 2.27%*, which is achieved with* 0 *forecast error on the graph. Actual wind power means are then scaled according to the x-axis, and maximum probability of line overload is recalculated (solid). The same is then done for standard deviations (dashed).*

With a desired safety level of $\epsilon = 2.27\%$ for each line, an error in the mean of 25% results in a maximum 15% chance of exceeding the limit. The solution is quite robust to errors in the standard deviation forecast, with a 25% error resulting in less than 6% chance of overload. See Figure 3.4.

**3.6. Scalability.** As an additional experiment illustrating scalability of the approach we studied the Polish national grid (obtained from MATPOWER as explained above) under simulated 20% renewable penetration spread over 50 wind farms, co-located with the 50 largest generators. This co-location should lessen the risk associated with renewable fluctuation (which should be partially "absorbed" by the co-located generators). For this experiment we used $\eta$ parameters of 3.0 both for lines and generators, which translates to $\epsilon_{ij}$ (and $\epsilon_g$) values of roughly $1.350 \times 10^{-3}$. Figure 3.5 studies the resulting risk exposure under standard OPF and CC-OPF. The chart shows the number of lines that attain several levels of overload probability. The performance of standard OPF is unacceptable: it would lead to frequent tripping of at least four lines. In contrast, under CC-OPF there is a drastic reduction in overload probabilities—the system is computationally stable.[12] Moreover, this is attained with

[12]The chart also shows small roundoff errors produced in the run.

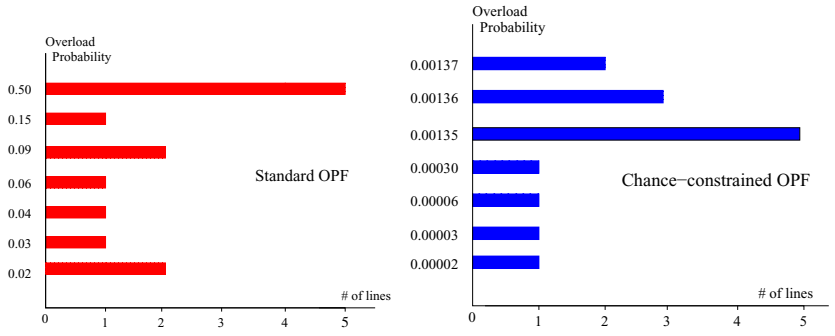


**Fig. 3.5** *Number of lines that are overloaded with given probability values in simulation of* 2746 *bus Polish power grid with* 20% *wind penetration distributed over* 50 *wind farms. Under standard OPF, five lines are overloaded half of the time, and two lines are overloaded more than* 10% *of the time, constituting a situation with unacceptable systemic risk. Under chance-constrained OPF, the largest overload probability is* fifty *times smaller than in the case of standard OPF. Moreover, the cost increase is by less than* 5%.

a minor increase in cost (less than 5%), while the computational time, on the standard laptop on which these experiments are performed, is on the order of 30 seconds. Varying the risk control parameters $\eta$ results in running times of the same order of magnitude.

**4. Discussions.** This manuscript suggests a new approach to incorporating uncertainty in the standard OPF setting routinely used in the power industry to set generation during a time window, or period. Our approach relies on chance constraints to limit the probability that any line (or generator) is overloaded, together with an optimal online control that accounts for controllable generator response to renewable fluctuation. Our key technical result is that the resulting optimization problem, CC-OPF, can be stated as a convex (conic) deterministic optimization problem. We also rely on plausible assumptions regarding the exogenous uncertainty, and approximate linearity of the underlying power flow equations. Our CC-OPF is solved quite efficiently, even on realistic large-scale instances, using a cutting-plane algorithm.

Experimental results using our algorithm are summarized as follows:

- We observe that CC-OPF delivers feasible results where standard OPF, run using the average forecast, would fail in the sense that many lines would be overloaded an unacceptably large portion of time.
- Not only is CC-OPF safer than standard OPF, but it also results in *cheaper* operation. This is demonstrated by considering the optimal cost of CC-OPF under a sufficiently high wind penetration solution (where standard OPF would fail) and a low penetration solution (corresponding to the highest possible penetration where standard OPF would not fail).
- We discover that solutions produced by CC-OPF deviate significantly from what amounts to a naive adjustment of the standard OPF obtained by correcting dispatch just at those generators which are close to overloaded lines.
- We performed out-of-sample testing by applying CC-OPF (modeling exogenous fluctuations as Gaussian) to other distributions. Overall these tests suggest that with a proper calibration of the effective Gaussian distribution our CC-OPF delivers good performance. The worst CC-OPF performance is observed for the most asymmetric distributions.

- We also presented a computationally sound data-robust version of the CC-OPF where the parameters for the Gaussian distributions are assumed unknown, but lying in a window. This allows for parameter mis-estimation and model error, and it suggests a tractable way for handling non-Gaussian fluctuations.

Our approach required a number of simplifications (arguably natural for a first study), in particular concerning static wind forecasts and power flow approximation. We also primarily focused on line congestion, while ignoring, for example, loss of synchronicity and voltage variations. Nevertheless, some extensions should be straightforward:

- Accounting for time evolving forecast/loads/etc. Wind forecast (in terms of the mean and standard deviation at the wind farm sites) changes on a time scale comparable to that of the generation redispatch interval. Our formulation (2.29)–(2.37) can be extended by splitting any single time interval into subintervals over which our control scheme would be computed.
- Accounting for synchronization bounds. Loss of synchronicity is probably the most acute contingency in power systems, and its prediction is a difficult nonlinear and dynamic problem. As shown recently in [20], one can formulate an accurate linear and static necessary condition for the loss of synchronicity. A chance-constrained version of this formulation can be incorporated seamlessly into our CC-OPF framework.

Finally, we see many opportunities in utilizing the CC-OPF (possibly modified) as an ingredient of even more complex problems, such as combined unit commitment (scheduling large power plants normally days, weeks, or even months ahead) [51] with CC-OPF, planning grid expansion [6] while accounting for cost operation under uncertainty, or incorporating CC-OPF in mitigating emergency of possible cascades of outages [17, 40, 44, 43, 11, 12, 22, 8]. In this context, it would be advantageous to speed up our already very efficient CC-OPF even further. See, for example, [13, 10]. A different methodology, relying on distributed algorithms, can be found in [35].

## 5. Appendix.

**5.1. Data-Robust Algorithm.** Here we provide a description of our data-robust approach for the CC-OPF problem discussed in section 2.4.3. Our first task is to incorporate constraint (2.59) into an efficient solution procedure. To motivate our approach, we will first show that a methodology based on traditional robust optimization techniques will (likely) not succeed at this task. We will then describe the method we rely on.

Constraint (2.59) can be incorporated into a formulation, provided we can appropriately represent the two maxima. Here, recall that the quantity $R_{ij}$,

$$R_{ij} = \max_{r \in \mathcal{M}} \left\{ e_{ij}^T \breve{\mathcal{B}} r \right\} \tag{5.1}$$

(previously defined in (2.60)), is independent of all variables and can be solved beforehand for all lines $\{i, j\}$; when $\mathcal{M}$ is of the form $U(\gamma, \Gamma)$ given above this is a linear programming problem, and when $\mathcal{M} = E(A, b)$ the task amounts to finding a point in the boundary of an ellipsoid with normal parallel to a given vector and thus requires solving a linear system of equations. We will likewise define a quantity $R_{ji}$ using $e_{ji}$ instead of $e_{ij}$.

It is the second maximization that presents some challenges.

LEMMA 5.1. *Suppose $\mathcal{S} = U(\gamma, \Gamma)$, and suppose a vector $\delta \in R^n$ is given. Then*

$$\max_{v \in \mathcal{S}} \left\{ \sum_{k \in \mathcal{W}} v_k (\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2 \right\} = \begin{array}{ll} \min & \Gamma a + \sum_{k \in \mathcal{W}} \gamma_k b_k \\ s.t. & \frac{1}{\gamma_k} a + b_k \geq (\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2 \quad \forall\, k \in \mathcal{W}, \\ & b_k \geq 0 \quad \forall\, k \in \mathcal{W}; \quad a \geq 0. \end{array}$$

*Proof.* Since $(\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2 \geq 0$ without loss of generality in the maximum we will have that $v_k \geq 0$ for all $k$. Thus, we can rewrite the maximum as

$$\begin{array}{lll} & \max & \sum_{k \in \mathcal{W}} v_k (\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2 \\ (5.2) & \text{s.t.} & \sum_{k \in \mathcal{W}} \frac{1}{\gamma_k} v_k \leq \Gamma \\ (5.3) & & v_k \leq \gamma_k \quad \forall\, k \in \mathcal{W}, \\ (5.4) & & 0 \leq v. \end{array}$$

This is a linear program; denoting by $a$ the dual variable for (5.2) and by $b_k$ that for (5.3), strong linear programming duality now gives the result. □

The use of linear programming duality as in Lemma 5.1 is key in the context of sets of the form $U(\gamma, \Gamma)$. In the case of an ellipsoidal set $E(A, b)$ as in (2.51) there is an analogue to Lemma 5.1 that instead uses the S-lemma [15, 25]. In the standard robust optimization approach, Lemma 5.1 would be leveraged to produce a result of the following type.

LEMMA 5.2. *Suppose $\mathcal{S} = U(\gamma, \Gamma)$. The data-robust chance-constrained problem is obtained by replacing for each line $\{i, j\}$ constraints* (2.34), (2.35) *of the nominal formulation with*

$$\beta_{ij}(\bar{\theta}(\bar{\mu})_i - \bar{\theta}(\bar{\mu})_j) + \beta_{ij} R_{ij} + \beta_{ij} \eta(\epsilon_{ij})\, s_{ij} \leq f_{ij}^{max}, \tag{5.5}$$

$$\beta_{ij}(\bar{\theta}(\bar{\mu})_j - \bar{\theta}(\bar{\mu})_i) + \beta_{ij} R_{ji} + \beta_{ij} \eta(\epsilon_{ij})\, s_{ij} \leq f_{ij}^{max}, \tag{5.6}$$

$$\left[ \sum_{k \in \mathcal{W}} \bar{\sigma}_k^2 (\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2 + \Gamma a^{\{i,j\}} + \sum_{k \in \mathcal{W}} b_k^{\{i,j\}} \right]^{1/2} \leq s_{ij}, \tag{5.7}$$

$$(\pi_{ik} - \pi_{jk} - \delta_i + \delta_j)^2 - \frac{1}{\gamma_k} a^{\{i,j\}} - b_k^{\{i,j\}} \leq 0 \quad \forall k \in \mathcal{W}, \tag{5.8}$$

$$b_k^{\{i,j\}} \geq 0 \quad \forall k \in \mathcal{W}; \quad a^{\{i,j\}} \geq 0. \tag{5.9}$$

*Here, $s_{i,j}$, $a^{\{i,j\}}$, and $b_k^{\{i,j\}}$ $(k \in \mathcal{W})$ are additional variables.*

*Proof sketch.* Without loss of generality at the optimum the $a^{\{i,j\}}$ and $b_k^{\{i,j\}}$ are chosen so as to minimize the left-hand side of (5.7) subject to all other variables held fixed, thereby obtaining the "min" in Lemma 5.1. □

Lemma 5.2 points out the difficulty that the standard robust optimization approach would engender in the context of our problem. First, the number of constraints (5.7) and (5.8) is *large*: it equals $|\mathcal{E}|(1 + |\mathcal{W}|)$, and thus in the case of a large transmission system it could approach many tens of thousands (or more). Thus, even though we obtain a *compact* formulation (i.e., of polynomial size) it is likely to be

proven too large for present-day solvers. But there is a second and more fundamental problem: constraint (5.7) *is not convex*. This is a significant methodological difficulty. A similar set of hurdles arises when using uncertainty sets $E(A, b)$.

**5.2. Cutting-Plane Algorithm for the Data-Robust Problem.** Here we describe a cutting-plane algorithm for the data-robust version of the CC-OPF problem, akin to Algorithm 2.6 used in section 2.3 to solve the nominal (known data) case.

In the following description for brevity we assume that only line-limit chance constraints are incorporated (no generator-related chance constraints).

ALGORITHM 5.3. *Cutting-Plane Algorithm for Data-Robust Problem.*

*Initialization:* The linear "master" system $A(\bar{p}, \alpha, \delta, \theta, s)^T \geq b$ is defined to include constraints (2.30)–(2.34). For each line $\{i, j\}$ compute the $R_{ij}$ parameters defined in (2.60).

*Iterate:*

(1) Solve $\min\{F(\bar{p}, \alpha) \ : \ A(\bar{p}, \alpha, \delta, \theta, s)^T \geq b\}$. Let $(\bar{p}^*, \alpha^*, \delta^*, \theta^*, s^*)$ be an optimal solution.

(2) For each line $\{i, j\}$ compute

$$V_{ij}^* \doteq \max_{v \in \mathcal{S}} \left\{ \sum_{k \in \mathcal{W}} v_k (\pi_{ik} - \pi_{jk} - \delta_i^* + \delta_j^*)^2 \right\},$$

with solution $\bar{v} = \bar{v}(\delta^*) \in \mathcal{S}$.

(3) If, for each line $\{i, j\}$, $\left[V_{ij}^*\right]^{1/2} \leq s_{ij}^*$ (within tolerance) **Stop.**

(4) Otherwise, add to the master system inequality (2.64).

Clearly the algorithm always maintains a valid (linear) relaxation to the data-robust optimization problem—this follows because (2.64) is simply an outer approximation (tangent hyperplane) to one of the conic sets defined by (2.63). Thus, in addition, the stopping criterion in (3) is valid since it proves that the current solution $(\bar{p}^*, \alpha^*, \delta^*, \theta^*, s^*)$ is in fact feasible. And, moreover, when the stopping criterion does not apply, then clearly the cutting plane added in (4) does cut off the current vector $\delta^*$. Finally, we point out that the optimization problem in (2) is a linear program in the case of the polyhedral uncertainty set, and a linear objective, ellipsoidally constrained program in the other case, both efficiently solvable.

**5.3. Additional Experiments.**

**5.3.1. Changing Locations for Wind Farms.** In this example we consider the effect of changing locations of wind farms. We take the MATPOWER 30-bus case with line capacities scaled by .75 and add three wind farms with average power output in a ratio of 2/3/4 and standard deviations at 30% of the average. Two choices of locations are shown in Figure 5.1. The first remains feasible for penetration up to 10%, while the second can withstand up to 55% penetration. This experiment shows that choosing location of the wind farms is critical for achieving the ambitious goal of high renewable penetration.

**5.3.2. Reversal of Line Flows.** Here we consider the 9-bus case with two wind sources and 25% average penetration and standard deviations set to 30% of the av-

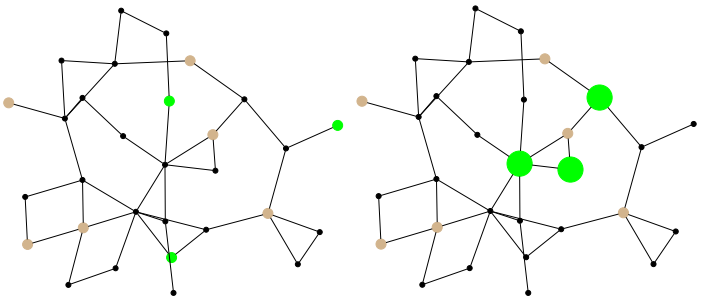

**Fig. 5.1** *30-bus case with three wind farms. The case on the left supports only up to* 10% *penetration before becoming infeasible, while the one on the right is feasible for up to* 55% *penetration.*

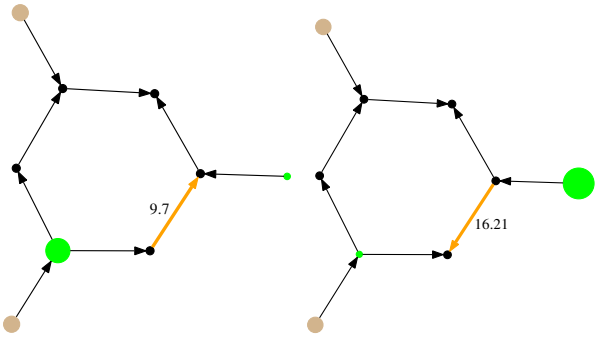


**Fig. 5.2** *9-bus case,* 25% *average penetration from two wind sources. With shifting winds, the flow on the orange line changes direction with a large absolute difference.*

erage case and analyze the following two somewhat rare but still admissible wind configurations: (1) wind source (a) produces its average amount of power and source (b) three standard deviations *below* average; (2) the reverse of the case (1). This results in a substantial reversal of flow on a particular line shown in Figure 5.2. This example suggests that when evaluating and planning for grids with high penetration of renewables one needs to be aware of potentially fast and significant structural rearrangements of power flows. Flow reversals and other qualitative changes of power flows, which are extremely rare in the grid of today, will become significantly much more frequent (typical) in the grid of tomorrow.

**Acknowledgments.** The authors are thankful to S. Backhaus and R. Bent for their comments, and to K. Dvijotham for help with BPA data.

They would like to acknowledge two recent preprints [47, 46] discussing CC-OPF ideas related in general but different in details (in particular, in what concern formulation and efficient implementation). The authors became aware of these preprints after posting the first version of this manuscript online.

## REFERENCES

[1] 20% *of Wind Energy by* 2030: *Increasing Wind Energy's Contribution to U.S. Electricity Supply*, DOE/GO-102008-2567, Department of Energy (DOE), Energy Efficiency and Renewable Energy, 2008.

[2] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, 1965.
[3] D. Apostolopoulou, P. W. Sauer, and A. D. Dominguez-Garcia, *Automatic generation control and its implementation in real time*, in Proceedings of the 47th Hawaii International Conference on System Sciences (HICSS), 2014, pp. 2444–2452.
[4] S. Baghsorkhi and I. Hiskens, *Analysis tools for assessing the impact of wind power on weak grids*, in Proceedings of the 6th Annual IEEE Systems Conference (SysCon), 2012, pp. 1–8.
[5] R. Ben-Tal and A. Nemirovski, *Robust solutions of linear programming problems contaminated with uncertain data*, Math. Programming, 88 (2000), pp. 411–424.
[6] R. Bent, G. Toole, and A. Berscheid, *Transmission network expansion planning with complex power flow models*, IEEE Trans. Power Systems, 27 (2012), pp. 904–912.
[7] A. Bergen and V. Vittal, *Power System Analysis*, Prentice-Hall, 2000.
[8] A. Bernstein, D. Hay, M. Uzunoglu, and G. Zussman, *Power Grid Vulnerability to Geographically Correlated Failures: Analysis and Control Implications*, preprint, 2012; available online from http://arxiv.org/abs/1206.1099.
[9] D. Bertsimas and M. Sim, *Price of robustness*, Oper. Res., 52 (2004), pp. 35–53.
[10] D. Bienstock, *Potential Function Methods for Approximately Solving Linear Programming Problems, Theory and Practice*, Kluwer Academic, 2002.
[11] D. Bienstock, *Adaptive online control of cascading blackouts*, in Proceedings of the 2011 IEEE Power and Energy Society General Meeting, IEEE, 2011, pp. 1–8.
[12] D. Bienstock, *Optimal control of cascading power grid failures*, in Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC), 2011, pp. 2166–2173.
[13] D. Bienstock and G. Iyengar, *Approximating fractional packings and coverings in* $0(1/\epsilon)$ *iterations*, SIAM J. Comput., 35 (2006), pp. 825–854.
[14] H. Bludszuweit, J. Dominguez-Navarro, and A. Llombart, *Statistical analysis of wind power forecast error*, IEEE Trans. Power Systems, 23 (2008), pp. 983–991.
[15] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
[16] A. Charnes, W. Cooper, and G. Symonds, *Cost horizons and certainty equivalents: An approach to stochastic programming of heating oil*, Management Sci., 4 (1958), pp. 235–263.
[17] J. Chen, J. Thorp, and I. Dobson, *Cascading dynamics and mitigation assessment in power system disturbances via a hidden failure model*, Internat. J. Electrical Power Energy Systems, 27 (2005), pp. 318–326.
[18] M. Chertkov, F. Pan, and M. Stepanov, *Predicting failures in power grids: The case of static overloads*, IEEE Trans. Smart Grids, 2 (2010), p. 150.
[19] M. Chertkov, M. Stepanov, F. Pan, and R. Baldick, *Exact and efficient algorithm to discover extreme stochastic events in wind generation over transmission power grids*, in Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC), 2011, pp. 2174–2180.
[20] F. Dörfler, M. Chertkov, and F. Bullo, *Synchronization in complex oscillator networks and smart grids*, Proc. Natl. Acad. Sci. USA, 10.1073/pnas.1212134110 (2013).
[21] *Energy Management Planning for the Integration of Wind Energy into the Grid in Germany, Onshore and Offshore by* 2020, Deutsche Energie-Agentur GmbH (DENA), 2005.
[22] M. Eppstein and P. Hines, *A "random chemistry" algorithm for identifying collections of multiple contingencies that initiate cascading failure*, IEEE Trans. Power Systems, 27 (2012), pp. 1698–1705.
[23] E. Erdogan and G. Iyengar, *Ambiguous chance constrained problems and robust optimization*, Math. Programming, 107 (2007), pp. 37–61.
[24] D. Goldfarb and F. Alizadeh, *Second-order cone programming*, Math. Programming, 95 (2003), pp. 3–51.
[25] D. Goldfarb and G. Iyengar, *Robust portfolio selection problems*, Math. Programming, 28 (2001), pp. 1–38.
[26] G. Gonzalez et al., *Experience integrating and operating wind power in the peninsular Spanish power system. Point of view of the transmission system operator and a wind power producer*, CIGRE, 2006.
[27] *Grid Integration of Wind Generation*, Technical brochure 450, International Conference on Large High Voltage Electric Systems, CIGRE, 2009.
[28] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag, 1998.
[29] *Gurobi Optimizer*, http://www.gurobi.com/, 2012.
[30] B.-M. Hodge and M. Milligan, *Wind power forecasting error distributions over multiple timescales*, in Detroit Power Engineering Society Meeting, 2011, pp. 1–8.

[31] M. Huneault and F. D. Galiana, *A survey of the optimal power flow literature*, IEEE Trans. Power Systems, 6 (1991), pp. 762–770.
[32] *ILOG CPLEX Optimizer*, http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/, IBM, 2012.
[33] *Integration of Renewable Resources: Transmission and Operating Issues and Recommendations for Integrating Renewable Resources on the California ISO-Controlled Grid*, http://www.caiso.com/1ca5/1ca5a7a026270.pdf, CAISO, 2007.
[34] J. E. Kelley, *The cutting-plane method for solving convex programs*, J. Soc. Indust. Appl. Math., 8 (1960), pp. 703–712.
[35] M. Kraning, E. Chu, J. Lavaei, and S. Boyd, *Message Passing for Dynamic Network Energy Management*, http://www.stanford.edu/~boyd/papers/decen_dyn_opt.html, 2012.
[36] P. Kundur, *Power System Stability and Control*, McGraw-Hill, New York, 1994.
[37] Y. Makarov, C. Loutan, J. Ma, and P. de Mello, *Operational impacts of wind generation on California power systems*, IEEE Trans. Power Systems, 24 (2009), pp. 1039–1050.
[38] L. Miller and H. Wagner, *Chance-constrained programming with joint constraints*, Oper. Res., 13 (1965), pp. 930–945.
[39] *MOSEK Optimizer*, http://www.mosek.com/, 2012.
[40] D. Nedic, I. Dobson, D. Kirschen, B. Carreras, and V. E. Lynch, *Criticality in a cascading failure blackout model*, Internat. J. Electrical Power Energy Systems, 28 (2006), pp. 627–633.
[41] A. Nemirovski and A. Shapiro, *Convex approximations of chance constrained programs*, SIAM J. Optim., 17 (2006), pp. 969–996.
[42] U. Ozturk and M. Mazumdar, *A solution to the stochastic unit commitment problem using chance constrained programming*, IEEE Trans. Power Systems, 19 (2004), pp. 1589–1598.
[43] R. Pfitzner, K. Turitsyn, and M. Chertkov, *Controlled Tripping of Overheated Lines Mitigates Power Outages*, preprint, arxiv:1104.4558, 2011.
[44] R. Pfitzner, K. Turitsyn, and M. Chertkov, *Statistical classification of cascading failures in power grids*, in Proceedings of the 2011 IEEE Power and Energy Society General Meeting, 2011, pp. 1–8.
[45] A. Prékopa, *On probabilistic constrained programming*, in Proceedings of the Princeton Symposium on Mathematical Programming, 1970, pp. 113–138.
[46] L. Roald, F. Oldewurtel, T. Krause, and G. Andersson, *Analytical reformulation of security constrained optimal power flow with probabilistic constraints*, in Proceedings of the Grenoble PowerTech, Grenoble, France, 2013.
[47] A. Sjodin, D. Gayme, and U. Topcu, *Risk-Mitigated Optimal Power Flow with High Wind Penetration*, http://smart.caltech.edu/papers/riskoptimal.pdf, 2012.
[48] *U.S.-Canada Power System Outage Task Force. Report on the August* 14, 2003 *Blackout in the United States and Canada: Causes and Recommendations*, Tech. report, https://reports.energy.gov, Department of Energy (DOE), 2004.
[49] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, *A probabilistic framework for security constrained reserve scheduling of networks with wind power generation*, in Proceedings of the 2012 IEEE International Energy Conference and Exhibition (ENERGYCON), 2012, pp. 452–457.
[50] G. Wang, M. Negrete-Pincetic, A. Kowli, E. Shafieepoorfard, S. Meyn, and U. Shanbhag, *Dynamic competitive equilibria in electricity markets*, in Control and Optimization Theory for Electric Smart Grids, A. Chakrabortty and M. Illic, eds., Springer, 2011.
[51] J. Wang, M. Shahidehpour, and Z. Li, *Security-constrained unit commitment with volatile wind power generation*, IEEE Trans. Power Systems, 23 (2008), pp. 1319–1327.
[52] *Wind Generation & Total Load in the BPA Balancing Authority*, http://transmission.bpa.gov/business/operations/wind/, Bonneville Power Administration, 2012.
[53] A. J. Wood and B. Wollenberg, *Power Generation, Operation and Control*, John Wiley and Sons, 1983.
[54] H. Zhang, *Chance Constrained Programming for Optimal Power Flow under Uncertainty*, IEEE Trans. Power Systems, 26 (2011), pp. 2417–2424.
[55] S. Zhang, Y. Song, Z. Hu, and L. Yao, *Robust optimization method based on scenario analysis for unit commitment considering wind uncertainties*, in Proceedings of the 2011 IEEE Power and Energy Society General Meeting, 2011, pp. 1–7.
[56] R. Zimmerman, C. Murillo-Sanchez, and R. Thomas, *MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education*, IEEE Trans. Power Systems, 26 (2011), pp. 12–19.