Electrical Power and Energy Systems 72 (2015) 116–125

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

Stochastic optimal power flow based on conditional value at risk and distributional robustness

Tyler Summers [a,*], Joseph Warrington [b], Manfred Morari [a], John Lygeros [a]

[a] Automatic Control Laboratory, ETH Zurich, Switzerland
[b] Energy Advisory at Baringa Partners LLP, UK

ARTICLE INFO

Article history:
Received 1 February 2015
Accepted 16 February 2015
Available online 26 March 2015

Keywords:
Stochastic optimal power flow
Reserve policies
Convex reformulation of chance constraints
Renewable integration

ABSTRACT

We present a computationally-efficient approach for solving stochastic, multiperiod optimal power flow problems. The objective is to determine power schedules for controllable devices in a power network, such as generators, storage, and curtailable loads, which minimize expected short-term operating costs under various device and network constraints. These schedules are chosen in a multistage decision framework to include planned power output adjustments, or reserve policies, which track errors in the forecast of power requirements as they are revealed, and which may be time-coupled. Such an approach has previously been shown to be an attractive means of accommodating uncertainty arising from highly variable renewable energy sources. Given a probabilistic forecast describing the spatio-temporal variations and dependencies of forecast errors, we formulate a family of stochastic network and device constraints based on convex approximations of chance constraints, and show that these allow economic efficiency and system security to be traded off with varying levels of conservativeness. Our formulation indicates two broad approaches, based on conditional value and risk and distributional robustness, that provide alternatives to existing methods based on chance and robust constraints. The results are illustrated using a case study, in which conventional generators plan schedules around an uncertain but time-correlated wind power injection.

Introduction

The current widespread increase in penetration of intermittent renewable energy in power networks comes with an increase in uncertainty of supply. As penetration levels of such sources reach substantial fractions of total supplied power, current techniques for handling supply uncertainty become expensive, and the system is exposed to increasing operational risks. On the other hand, there is active ongoing research on obtaining forecasts of intermittent power supplied from renewable sources over various time scales [20]. The most sophisticated available forecasts are probabilistic, including not only point forecasts over a time horizon, but also information about probability distributions of forecast errors that describes spatiotemporal variations and dependencies. It is widely agreed that appropriate use of such forecasts and strategies for responding to forecast errors are required to make operational decisions that intelligently manage risks in the network to achieve a trade-off between economic efficiency and system security.

One of the fundamental decision problems in power networks is optimal power flow (OPF), where power schedules for controllable devices in a power system, such as generators, storage, and controllable loads, are determined to minimize operating cost under various device and network constraints. OPF is central to economic and secure operation and control of power systems and markets [27]. Future power networks will require the coordination of thousands of devices and joint optimization of millions of variables and increasingly the explicit incorporation of information about uncertainties. There are many OPF problem variations, including unit commitment, reserve scheduling, economic dispatch, security-constrained, DC approximations, full AC formulations and relaxations, and others. In this paper, we use a relatively simple but widely used linearized DC approximation to illustrate our results, although of course corresponding extensions and variations are interesting and necessary for the methods to be useful in practice.

Historically, many OPF formulations have only accommodated uncertainty in a rather rudimentary manner by choosing fixed reserve margins without using other known or estimated probabilistic information about forecast errors. More recent work that

* Corresponding author.
E-mail addresses: tsummers@control.ee.ethz.ch (T. Summers), joe.warrington@baringa.com (J. Warrington), morari@control.ee.ethz.ch (M. Morari), lygeros@control.ee.ethz.ch (J. Lygeros).

http://dx.doi.org/10.1016/j.ijepes.2015.02.024
0142-0615/

does explicitly handle uncertainty includes methods based on (1) chance constraints and the so-called scenario approach [19,34,31,9,30,22], in which decisions are made based on finite sampling of uncertain parameters from an assumed statistical model, or (2) robust optimization [33], in which knowledge of uncertainty bounds is assumed and device and network constraints are enforced for every possible uncertainty realization. Both approaches suffer from theoretical and practical drawbacks. Methods based on chance constraints involve incoherent risk measures that penalize frequency but not severity of constraint violations. Furthermore, the scenario approach in practice can be conservative as a result of drawing a large number of samples to get probabilistic constraint satisfaction guarantees. Methods based on robust optimization can be very conservative in enforcing constraints for every possible uncertainty realization, since some realizations deemed possible may in fact be extremely unlikely. Bienstock et al. have also recently considered power flow problems subject to chance constraints assuming that the uncertainties are Gaussian [3].

There are other risk measures, such as expected shortfall and conditional value at risk (CVaR), which are well known in finance [23], but have only received limited attention in the context of power networks. Some of these are discussed in the context of electricity markets in [11], which illustrates the use of these metrics in the context of individual power producers, retailers, and consumers managing financial risk in electricity markets. A recent related paper [35] also considers the use of CVaR for managing financial risk of wind power producers in a network.

Here we present an approach to solving a stochastic optimal power flow problem based on convex approximations of chance constraints [18]. We take the perspective of a transmission system operator managing operational risks across the network rather than that of an individual power supplier and use stochastic formulations of network and device constraints. The family of convex approximations of chance constraints we consider indicates two broad alternative approaches for handling uncertainty in optimal power flow problems that effectively interpolate between chance constraints and robust constraints. The first is a CVaR approach that limits both frequency and severity of constraint violations and requires uncertainty sampling. The second is a distributionally robust approach that allows probabilistic information to be incorporated into a single second order cone constraint and does not require any uncertainty sampling. We use a multistage stochastic programming formulation with affine reserve policies, which specify how controllable devices in the network should respond to forecast errors and can be computed tractably. This formulation allows strong temporal forecast error dependencies and time-coupled device costs and constraints to be incorporated. Our main contribution is to bring together these chance constraint alternatives with DC OPF and the use of reserve policies.

A preliminary version of this work appeared in [28]. Here, we extend [28] in several ways. We elaborate on the distributionally robust framework to handle multiple constraints jointly, to include information on the distribution support in addition to moment information, and to include a unimodality assumption on the distribution. We also raise the issue of pricing under probabilistic constraints and show that nodal prices arising from congestion in the network do not decompose in a straightforward way as they do in formulations with deterministic or robust constraints.

We demonstrate that this formulation can achieve alternative trade-offs between efficient and secure network operation while also reducing the conservatism of previous approaches. Based on network knowledge and a probabilistic model for forecast errors that accounts for spatiotemporal variations and correlations, risk can be intelligently distributed across the network. The results and various trade-offs are illustrated numerically on a simple two-bus example.

The rest of the paper is organized as follows. Section 'Network model and optimal power flow' describes the network model and formulates a stochastic optimal power flow problem that explicitly accounts for information about uncertainty. Section 'Convex approximations of chance constraints' describes a family of convex approximations of chance constraints and shows the approximations indicate two broad approaches for handling uncertainty in the optimal power flow problem. Section 'Elaborations on distributionally roBUST OPF' elaborates on the distributionally robust framework. Section 'Pricing under probabilistic constraints' illustrates the trade-offs with a simple numerical example. Section 'Numerical example' raises issues with pricing under probabilistic constraints, and Section 'Conclusions and outlook' gives concluding remarks and an outlook for future research.

Network model and optimal power flow

We consider the operation of N devices connected via a transmission network over a planning time horizon of T discrete time steps. The devices may include generators; fixed, deferrable, and curtailable loads; and storage devices such as batteries that can act as either generators or loads. We distinguish between two types of devices: those with fixed and (possibly) uncertain power flows that cannot be affected by decision variables (e.g., renewable infeeds or fixed loads), and those with controllable power flows that can be affected by decision variables (e.g., conventional thermal and hydro generation, deferrable/curtailable loads, or storage devices). The notation follows [33].

We do not consider here the unit commitment problem, which includes binary variables encoding whether generating units are on or off. We assume that this computation has already been done and is encoded into the power limits of controllable devices. Integrating unit commitment into our framework is left for future work.

Devices with fixed power flows

The fixed power flow for device i is given by $r_i + G_i\delta$ with positive values denoting net power injection into the network. The vector $r_i \in \mathbf{R}^T$ represents the nominal forecasted power over the planning horizon, and the linear function $G_i \in \mathbf{R}^{T\times N_\delta T}$ of the random vector $\delta \in \Delta \subseteq \mathbf{R}^{N_\delta T}$ represents the prediction error of the power injection or extraction for device i. If uncertainty of device i is not explicitly considered, we set $G_i = 0$.

We assume that information about the joint probability distribution of δ is known. The distribution captures spatial variations and dependencies among devices and temporal variations and dependencies across the horizon. This information can come in the form of knowledge of the full distribution, knowledge of certain moments such as the mean and variance, or a model of δ from which we can draw samples.

Devices with controllable power flows

The power flows of controllable devices are governed by given dynamics. Device i at time k has internal state $x_k^i \in \mathbf{R}^{n_i}$, where n_i is the state dimension. The dynamics of device i are assumed to be governed by the discrete-time linear dynamical system

$$x_{k+1}^i = \bar{A}_i x_k^i + \bar{B}_i u_k^i, \qquad (1)$$

where $\bar{A}_i \in \mathbf{R}^{n_i\times n_i}$ is the dynamics matrix, $\bar{B} \in \mathbf{R}^{n_i\times m_i}$ is the input matrix, and $u_i \in \mathbf{R}^{m_i}$ is an input that controls the net power injection. The first element $[x_k^i]_1$ of the state vector x_k^i represents the power injection of device i at time k into the network at a certain bus; other elements model internal dynamics such as state of charge of a battery or memory of previous states, which can be used to encode ramping constraints for thermal generation. For compact notation, we concatenate the states and inputs over the planning horizon: $\mathbf{x}^i = [(x_1^i)^T, \ldots, (x_T^i)^T] \in \mathbf{R}^{n_i T}$ and $\mathbf{u}^i = [(u_0^i)^T, \ldots, (u_{T-1}^i)^T] \in \mathbf{R}^{m_i T}$, which will be decision variables in the optimization problem we formulate in the following. Note that future states can be expressed as a linear function of the input sequence and the current state x_0^i according to the dynamics (1):

$$\mathbf{x}^i = A_i x_0^i + B_i \mathbf{u}^i, \tag{2}$$

where

$$A_i = \begin{bmatrix} \bar{A}_i \\ \bar{A}_i^2 \\ \vdots \\ \bar{A}_i^{T-1} \end{bmatrix}, \quad B_i = \begin{bmatrix} \bar{B}_i & 0 & \cdots & 0 \\ \bar{A}_i\bar{B}_i & \bar{B}_i & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \bar{A}_i^T\bar{B}_i & \cdots & \bar{A}_i\bar{B}_i & \bar{B}_i \end{bmatrix}.$$

Cost functions and constraints

We associated a cost function with each device $J_i : \mathbf{R}^{n_i T} \times \mathbf{R}^{\mathbf{m_i T}} \rightarrow \mathbf{R}$ that encodes the cost for the device to produce a given power schedule over the planning time horizon. The cost functions are assumed for mathematical convenience to be convex quadratic:

$$J_i(\mathbf{x}_i, \mathbf{u}_i) = f_{ix}^T \mathbf{x}_i + \frac{1}{2}\mathbf{x}_i^T H_{ix} \mathbf{x}_i + f_{iu}^T \mathbf{u}_i + \frac{1}{2}\mathbf{u}_i^T H_{iu} \mathbf{u}_i + c_i, \tag{3}$$

where H_{ix} and H_{iu} are positive semidefinite matrices.

There are three types of constraints: local device constraints, power balance constraints, and line flow constraints. The local constraints are linear inequalities of the form

$$T_i \mathbf{x}^i + U_i \mathbf{u}^i + V_i \delta \leqslant w_i, \tag{4}$$

where $T_i \in \mathbf{R}^{l_i \times n_i T}, U_i \in \mathbf{R}^{l_i \times m_i T}, V_i \in \mathbf{R}^{l_i \times N_\delta T}$. These can be used to encode a wide variety of constraints; for example, constraints on the allowable power injection range, or time coupling constraints on ramp rates of a generator.

The remaining two types of constraints are imposed by the network. In general, the steady-state active and reactive power flows in a network are related to the complex bus voltages in the network via nonlinear power flow equations. We consider a widely used approximation in which it is assumed that voltage phase angle differences between buses are small, bus voltage magnitudes are constant and close to 1 per unit, and lines are lossless. Under these assumptions, the reactive flows are neglected, and the active line flows are proportional to the phase differences between bus voltages. These assumptions amount to a linearization of a nonlinear AC OPF problem. For more details on modeling for linearized and decoupled OPF problems, see [26,1]. Although these approximations are widely used in practice, it should be recognized that they do not fully capture the full nonlinear flows and do not account for nodal voltage constraints.

The second type of constraint is a power balance constraint. The net power injection from all devices in the network must be zero for all times in the planning horizon, which can be encoded with the T linear equality constraints

$$\sum_{i=1}^{N} (r_i + G\delta + C_i \mathbf{x}_i) = 0. \tag{5}$$

Third, the power flow should also satisfy line rating constraints on all transmission lines in the network. If there are L transmission lines in the network and we consider constraints on lines flows in both directions, these can be encoded by the $2LT$ additional inequality constraints

$$\sum_{i=1}^{N} \Gamma_i (r_i + G_i \delta + C_i \mathbf{x}_i) \leqslant \bar{p}, \tag{6}$$

where $\Gamma_i \in \mathbf{R}^{2LT \times T}$. The matrices Γ_i map the power injections of each device to its contribution to each line flow and can be constructed from network line impedances (see [10]).

Reserve policies

In a standard OPF problem, uncertainty is ignored, e.g., by setting the prediction error vector δ to zero, and the device inputs $\mathbf{u}_i$ are chosen to minimize the sum of the device cost functions, which is a quadratic program. To explicitly account for uncertainty, in addition to computing such a nominal plan, we would also like to find an optimal strategy for responding to forecast errors. To do this, we allow the device inputs to depend on the uncertainty via a policy for each device $\mathbf{u}_i = \pi_i(\delta)$, where $\pi_i : \mathbf{R}^{N_\delta T} \rightarrow \mathbf{R}^{m_i T}$ is a *function* that belongs to a set of causal policies denoted by Π_c over which we would like to optimize.

Furthermore, the objective function and constraints both depend on the random variable δ, so these terms in the OPF problem need to be recast into stochastic forms. For the objective function, we consider optimizing the expected value of the sum of device cost functions. There are a variety of ways to recast the constraints. We will require the power balance equality constraints to hold for any possible uncertainty realization (after application of the policy). The inequality constraints could be enforced for any possible uncertainty realization based on assumed knowledge of uncertainty bounds, as in the robust setting of [33]. Alternatively, they could be "softened" and enforced in some weaker probabilistic sense based on assumed knowledge of the uncertainty probability distribution.

Here we enforce the equality constraints robustly and the inequality constraints probabilistically; however, one can consider other mixtures of robust and stochastic constraint formulations. For example, the power balance equality constraint could be formulated as two inequality constraints, and these could also be reformulated probabilistically. Further, some local device constraints, e.g., upper and lower generation limits, may be most appropriately modeled as robust constraints. In any case, when constraints are violated, further manual action will be required of the system operator, which is typical of current operational practice. The idea of including stochastic constraints is to limit the frequency and severity with which the operator must take such actions.

Substituting the policy, eliminating $\mathbf{x}_i$ using (2), and recasting the constraints leads to the following multistage stochastic programming formulation of the optimal power flow problem:

$$\underset{\pi_i \in \Pi_c}{\text{minimize}} \quad \mathbf{E}\sum_{i=1}^{N} J_i(A_i x_0^i + B_i \pi_i(\delta), \pi_i(\delta)), \tag{7a}$$

$$\text{subject to} \quad \sum_{i=1}^{N} (r_i + G_i\delta + C_i(A_i x_0^i + B_i \pi_i(\delta))) = 0, \quad \forall \delta, \tag{7b}$$

$$\mathcal{R}\left(\sum_{i=1}^{N} \Gamma_i (r_i + G_i\delta + C_i(A_i x_0^i + B_i \pi_i(\delta))) - \bar{p} \leqslant 0\right), \tag{7c}$$

$$\mathcal{R}\big(T_i(A_i x_0^i + B_i \pi_i(\delta)) + U_i \pi_i(\delta) - w_i \leqslant 0\big), \quad i = 1, \ldots, N. \tag{7d}$$

Here $\mathcal{R}$ denotes a generic transformation of the inequality constraints into stochastic versions, using probabilistic uncertainty information and possibly introducing auxiliary variables. The details of these transformations will be introduced in the next section.

The infinite-dimensional optimization over the set of admissible causal functions Π_c is intractable. Inspired by the formulation of [14] in the context of robust model predictive control, we restrict each function π_i to the class of causal affine policies,

$$\pi_i(\delta) = D_i\delta + e_i, \tag{8}$$

where each $D_i \in \mathbf{R}^{T\times N_\delta T}$ is block lower-triangular (to enforce causality) and represents a system of planned deviations with respect to a nominal plan $e_i \in \mathbf{R}^T$. Since the device cost functions are quadratic, the cost then becomes a linear function of the first and second moments of the distribution of δ. The robust equality constraints are equivalent to

$$\sum_{i=1}^{N}(r_i + C_i(A_i x_0^i + B_i e_i)) = 0, \quad \sum_{i=1}^{N}(G_i + C_i B_i D_i) = 0. \tag{9}$$

In the robust approach of [33], constraints (7c) and (7d) are required to hold for all uncertainty realizations. After reformulating they become linear in the decision variables D_i and e_i and some extra auxiliary matrix variables.

A linear variation of the solution in response to uncertainty also coincides with the way Automatic Generator Control (AGC) mechanisms often apply specific gains to different participating generators such that automatic changes in their output compensate a certain fraction of demand uncertainty for the whole system. The restriction (8) permits the convenient interpretation of diagonal entries of the matrices D_i as generator AGC parameters [33].

In the next section we discuss other reformulations that allow for the possibility of some degree of constraint violation in exchange for reduced solution cost.

Convex approximations of chance constraints

In this section, we discuss chance constraints and a family of convex approximations based on [18]. We highlight how two particular functions in the family of approximations indicate two broad approaches for handling uncertainty that lie in a sense in between chance and robust constraints that have been the focus of most power systems applications. The first function is piecewise linear and offers a conditional value at risk approach that limits both frequency and severity of constraint violation and requires sampling. The second function is quadratic and offers a distributionally robust approach that also effectively limits both frequency and severity of constraint violation and requires knowledge of only mean and covariance of the uncertainty. At the end of the section we formulate two prototypical stochastic OPF problems based on these approaches.

Chance constraints

Consider the chance-constrained optimization problem

$$\begin{aligned} &\text{minimize} \quad f_0(x), \\ &\text{subject to} \quad \mathbf{P}(f(x,\delta) \leqslant 0) \geqslant 1-\alpha, \end{aligned} \tag{10}$$

where $x \in \mathbf{R}^n$ is the decision variable, $\delta \in \mathbf{R}^d$ is a random variable with distribution $\mathbf{P}$, and, for the time being, $f(x,\delta) : \mathbf{R}^n \times \mathbf{R}^d \to \mathbf{R}$ is a single scalar constraint function that is convex in x for each δ; we discuss how to deal with multiple constraint functions later. We assume that $f_0(x)$ is a convex cost function, and α is a safety parameter specified by the modeler. This problem is convex in only a few special cases. For example, when $f(x,\delta) = a(x)^T\delta + b(x)$, with $a(x)$ and $b(x)$ affine in x, and δ is normally distributed, then the chance constraint can be expressed as a second-order cone constraint:

$$a(x)^T\bar{\delta} + b(x) + \Phi^{-1}(1-\alpha)\|\Xi^{1/2}a(x)\|_2 \leqslant 0, \tag{11}$$

where Φ^{-1} is the Gaussian quantile function, $\bar{\delta} = \mathbf{E}\delta$ is the mean of δ, and $\Xi = \mathbf{E}\delta\delta^T - \bar{\delta}\bar{\delta}^T$ is the covariance matrix of δ. In other cases, the random parameter is typically sampled from a distribution and a constraint is added for each of the sampled values, leading to a deterministic convex program. Recent research has focused on quantifying the probability and determining the required number of samples such that the solution of the sampled problem is feasible for the original problem [5,7]. Chance constraints are closely related to Value at Risk (VaR), a risk measure often used in finance [23].

Chance constraints have several drawbacks. They penalize frequency but not severity of constraint violation. Moreover, the associated VaR is not a "coherent" risk measure, i.e., it has some undesirable properties for certain types of uncertainty distributions [23]. Also, when using a sampling approach, the number of samples required to guarantee a certain probabilistic feasibility level can be large, which can make the sampled optimization problem difficult to solve. This can also potentially render the solution very conservative in practice, though there are sampling-and-discarding heuristics that can alleviate this in principle to some extent [8].

Convex approximation of chance constraints

One can obtain a family of related probabilistic constraints by making a conservative convex approximation. In particular, one can replace the constraint in problem (10) with another constraint whose feasible set is contained in the feasible set of problem (10).

First, note that for a random variable z and for any $t > 0, \mathbf{P}(tz \geqslant 0) = \mathbf{P}(z \geqslant 0) = \mathbf{E}[\mathbf{1}_{[0,\infty)}(tz)]$ where $\mathbf{1}_K(\cdot)$ is the indicator function over the set K. Now let $\psi : \mathbf{R} \to \mathbf{R}$ be a non-negative, convex function with $\psi(z) > \psi(0) = 1$ for all $z > 0$, which is called the *generating function* that will generate a family of convex approximations for the chance constraint. Since $\psi(tz) \geqslant \mathbf{1}(tz)$ $\forall tz \in \mathbf{R}$, it follows that $\mathbf{E}\psi(tz) \geqslant \mathbf{E}\mathbf{1}_{[0,\infty)}(tz) = \mathbf{P}(z \geqslant 0)$, i.e., the function $\mathbf{E}\psi(tz)$ is an upper bound on the probability that $z \geqslant 0$.

Replacing z with $f(x,\delta)$ and changing t to t^{-1} yields

$$\mathbf{E}[\psi(t^{-1}f(x,\delta))] \geqslant \mathbf{P}(f(x,\delta) > 0). \tag{12}$$

Thus, the constraint

$$\inf_{t>0}\left(t\mathbf{E}[\psi(t^{-1}f(x,\delta))] - t\alpha\right) \leqslant 0 \tag{13}$$

is a sufficient condition for the chance constraint in (10) to be satisfied. This constraint can be shown to be jointly convex in (t,x) [18].

There are several candidates for the generating function:

- **Markov:** $\psi(z) = [1+z]_+$.
- **Chebyshev:** $\psi(z) = ([1+z]_+)^2$.
- **Traditional Chebyshev:** $\psi(z) = (1+z)^2$.
- **Chernoff/Bernstein:** $\psi(z) = e^z$.

where $[\cdot]_+ = \max(\cdot, 0)$. Each function places a different penalty on the severity of constraint violation. We will focus on two functions, namely the Markov and traditional Chebyshev, that indicate two broad methods for handling uncertainty in constraints.

Markov generating function & conditional value at risk

The constraint obtained from the Markov generating function is closely related to the Conditional Value at Risk (CVaR), which is also a well known, coherent risk measure in finance [23]. The CVaR constraint can be written as

$$\mathbf{E}[f(x,\delta) + t]_+ \leqslant t\alpha, \tag{14}$$

with t an optimization variable. In contrast to the chance constraint in (10), this CVaR constraint limits both frequency and expected severity of constraint violation, which is arguably more appropriate for many types of constraints. Although this CVaR constraint is jointly convex in (t, x), the expectation cannot be evaluated explicitly in closed form due to non-smoothness of the Markov generating function. However, the constraint can be approximated using sample average approximation methods, which have received significant recent attention in the operations research literature, e.g., in [16,24,32], and have been shown to be effective in many practical problems [15]. In our numerical examples below we observe that accurate solutions can be obtained with a moderate number of samples.

Traditional Chebyshev generating function and distributional robustness

The traditional Chebyshev approximation provides a more conservative approximation of the chance constraint that has a distributionally robust interpretation. An important advantage is that in certain cases, the expectation can be evaluated analytically. To illustrate we turn to the case $f(x,\delta) = a(x)^T\delta + b(x)$, where $a(x)$ and $b(x)$ are affine in x, and δ has mean $\mathbf{E}\delta = \bar{\delta}$ and variance $\mathbf{E}\delta\delta^T - \bar{\delta}\bar{\delta}^T = \Xi$. The constraint obtained from the traditional Chebyshev generating function can be written as

$$a(x)^T\bar{\delta} + b(x) + \sqrt{\frac{1-\alpha}{\alpha}}\|\Xi^{1/2}a(x)\|_2 \leqslant 0, \tag{15}$$

which is a second-order cone constraint that depends only on the mean and variance of δ. Note that (15) has the same form as (11) but with a larger, more conservative parameter multiplying the second term. This is because (11) assumes a Gaussian distribution, whereas (15) is agnostic in the sense that it works for all distributions with the same mean and variance, as discussed below.

The constraint associated with the traditional Chebyshev generating function is an example of a distributionally robust constraint [6,12,13,36,25]. In particular, it can be shown that the constraint is equivalent to

$$\mathbf{P}(a(x)^T\delta + b(x) \leqslant 0) \geqslant 1-\alpha, \quad \forall \mathbf{P} \in \mathcal{P}(\bar{\delta}, \Xi),$$

where $\mathcal{P}(\bar{\delta}, \Xi)$ is the set of all distributions of δ with mean $\bar{\delta}$ and variance Ξ [6]. In other words, if (15) holds, then the corresponding chance constraint holds for any distribution of δ with the given mean and variance. In practice, probability distributions are not known and must be estimated from limited historical data. So the advantages are that it is typically easy to obtain good estimates for the mean and variance from data and that the reformulation as a single second-order cone constraint means that no sampling is required. A drawback is that such distributionally robust constraints can be very conservative because the corresponding worst case distributions are often unlikely to be encountered in practice. However, the conservatism can be reduced by assuming or estimating more about the distribution, e.g., the support [36], unimodality [25] and/or knowledge of higher moments [21].

CVaR OPF and distributionally robust OPF

Based on the above approximations, we can formulate the corresponding stochastic optimal power flow problems with associated stochastic versions of the local device and line flow constraints. Each line constraint k at each time step, associated with a row in (6), can be written in the form

$$f_k(D,e,\delta) = [\Phi(D)]_k\delta + [b(e)]_k, \tag{16}$$

where $D := (D_1, \ldots, D_N), e := (e_1, \ldots, e_N), [\cdot]_k$ denotes the kth row of a matrix or the kth element of a vector. The optimization variables D_i and e_i for $i = 1, \ldots, N$ enter linearly into Φ and b as follows:

$$\Phi(D) := \sum_{i=1}^{N} \Gamma_i C_i B_i D_i, \tag{17}$$

$$b(e) := -\bar{p} + \sum_{i=1}^{N} \Gamma_i (r_i + G_i\delta + C_i A_i x_0^i + C_i B_i e_i). \tag{18}$$

A similar form can be obtained for each local device constraint $j = 1, \ldots, l_i$ for each device i, which we denote individually as $g_{ij}(D, e, \delta)$.

Using the Markov generating function, we obtain the following CVaR stochastic optimal power flow problem

$$\begin{aligned}
&\underset{D,e,t}{\text{minimize}} \quad \mathbf{E}\sum_{i=1}^{N} J_i(A_i x_0^i + B_i(D_i\delta + e_i), D_i\delta + e_i)\\
&\text{subject to} \quad \sum_{i=1}^{N}(r_i + C_i(A_i x_0^i + B_i e_i)) = 0,\\
&\sum_{i=1}^{N}(G_i + C_i B_i D_i) = 0\\
&\mathbf{E}[f_k(D,e,\delta) + t_k]_+ \leqslant t_k\alpha_k, \quad k = 1, \ldots, 2LT\\
&\mathbf{E}[g_{ij}(D_i,e_i,\delta) + t_{ij}]_+ \leqslant t_{ij}\alpha_{ij}, \quad i = 1, \ldots, N, j = 1, \ldots, l_i, \quad \text{(CVaR OPF)}
\end{aligned}$$

where $t := (t_1, \ldots, t_{2LT}, t_{1,1}, \ldots, t_{N,l_N})$. This is a convex optimization problem in (D, e, t); the expected value in the last two sets of constraints can be approximated by a sample average. Since the cost functions are assumed to be convex quadratic, the expected value there can be explicitly evaluated in terms of the first and second moments of δ.

Here we have treated the individual chance constraint independently, and α are independent parameters associated with each constraint. In the following section, we discuss how a set of chance constraints can be considered jointly by choosing the α appropriately.

Using the traditional Chebyshev generating function, we obtain the following distributionally robust stochastic optimal power flow problem

$$\begin{aligned}
&\underset{D,e}{\text{minimize}} \quad \mathbf{E}\sum_{i=1}^{N} J_i(A_i x_0^i + B_i(D_i\delta + e_i), D_i\delta + e_i)\\
&\text{subject to} \quad \sum_{i=1}^{N}(r_i + C_i(A_i x_0^i + B_i e_i)) = 0,\\
&\sum_{i=1}^{N}(G_i + C_i B_i D_i) = 0\\
&f_k(D,e,\bar{\delta}) + \sqrt{\frac{1-\alpha_k}{\alpha_k}}\|\Xi^{1/2}[\Phi(D)]_k^T\|_2 \leqslant 0, \quad k = 1, \ldots, 2LT\\
&g_{ij}(D_i,e_i,\bar{\delta}) + \sqrt{\frac{1-\alpha_{ij}}{\alpha_{ij}}}\|\Xi^{1/2}[\Phi_g(D)]_{ij}^T\|_2 \leqslant 0, \quad i = 1, \ldots, N, \quad j = 1, \ldots, l_i.
\end{aligned}$$

(Distribution ally Robust OPF)

The last two sets of constraints are second order cone constraints that depend only on mean and the variance of forecast errors; once the mean and variance have been estimated from data, no sampling of the uncertainty is required to solve the optimization problem.

Elaborations on distributionally robust OPF

In this section, we elaborate on the distributionally robust framework in several ways that improve the basic setup from Section 'Convex approximations of chance constraints'. Specifically, we describe how multiple constraints can be handled simultaneously, how distribution support information can be incorporated in addition to the first and second moments, and how a unmodality assumption can reduce conservatism. These improvements are based on results in [36,25] and can easily be included into the stochastic OPF formulations.

Distributionally robust joint chance constraints

Consider the distributionally robust joint chance constrained problem

$$\begin{aligned} &\text{minimize} \quad f_0(x) \\ &\text{subject to} \quad \mathbf{P}(a_i(x)^T\delta + b_i(x) \leqslant 0, i = 1,\dots,m) \geqslant 1-\alpha, \\ &\qquad\qquad \forall \mathbf{P} \in \mathcal{P}(\bar{\delta}, \Xi), \end{aligned} \tag{19}$$

where $a_i(x)$ and $b_i(x)$ are affine in x. Here, the probability of *joint* satisfaction of the constraints is specified by the modeler to be at least $1-\alpha$ for any distribution of δ with mean $\bar{\delta}$ and variance Ξ.

A basic sufficient condition for joint chance constrained problems can be obtained using the Boole inequality (also known as the Bonferroni inequality). In particular, if each of the distributionally robust individual chance constraints are satisfied with safety parameter α_i and $\sum_{i=1}^m \alpha_i \leqslant \alpha$, then the joint chance constraint will also be satisfied with safety parameter α. However, this condition is not tight in general. Recently, Zymler et al. presented an improved condition formulation that can be shown to be "essentially" exact [36].

Associate a scaling parameter $\gamma_i,\ i = 1,\dots,m$ with each constraint, and let

$$\Omega = \begin{bmatrix} \Xi + \bar{\delta}\bar{\delta}^T & \bar{\delta} \\ \bar{\delta}^T & 1 \end{bmatrix}.$$

The main result from [36] is the following in a semidefinite programming reformulation of the distributionally robust joint chance constrained problem

$$\begin{aligned} &\text{minimize} \quad f_0(x) \\ &\text{subject to} \quad \beta + \frac{1}{\alpha}\mathbf{tr}(\Omega M) \leqslant 0, \quad M \succeq 0 \\ &\qquad M - \begin{bmatrix} 0 & \frac{\gamma_i}{2}a_i(x) \\ \frac{\gamma_i}{2}a_i(x)^T & \gamma_i b_i(x) - \beta \end{bmatrix} \succeq 0, \\ &\qquad i = 1,\dots,m, \end{aligned} \tag{20}$$

in variables $x \in \mathbf{R}^n, \beta \in \mathbf{R}$, and $M = M^T \in \mathbf{R}^{d+1\times d+1}$. The "essential" exactness means that there exist scaling parameters γ_i such that the reformulation is exact. Unfortunately, the problem is not jointly convex when the γ_i are included as decision variables, but heuristic sequential convex optimization methods can be used to improve them.

Including support information

The support of the uncertainty is typically a strict subset of $\mathbf{R}^d$ due to physical limitations. For example, any uncertain power injection of a wind farm is limited between zero and the maximum rated capacity of the farm. Disregarding this information when formulating stochastic constraints can render decisions unnecessarily conservative. Suppose the support of δ is given by a finite set of quadratic and linear inequalities

$$\Delta = \left\{ \delta \in \mathbf{R}^d \middle| \begin{bmatrix} \delta \\ 1 \end{bmatrix}^T W_i \begin{bmatrix} \delta \\ 1 \end{bmatrix} \leqslant 1,\ i = 1,\dots l \right\},$$

where W_i are positive semidefinite symmetric matrices. Then the distributionally robust joint chance constrained program can be expressed using the S-procedure [4] as

$$\begin{aligned} &\text{minimize} \quad f_0(x) \\ &\text{subject to} \quad \beta + \frac{1}{\alpha}\mathbf{tr}(\Omega M) \leqslant 0, \quad \tau_i \geqslant 0, \tau_0 \geqslant 0 \\ &\qquad M + \sum_{j=1}^{l} \tau_{0,j} W_j \succeq 0, \\ &\qquad M + \sum_{j=1}^{l} \tau_{i,j} W_j - \begin{bmatrix} 0 & \frac{\gamma_i}{2}a_i(x) \\ \frac{\gamma_i}{2}a_i(x)^T & \gamma_i b_i(x) - \beta \end{bmatrix} \succeq 0, \quad i = 1,\dots,m, \end{aligned} \tag{21}$$

which is again a semidefinite program in variables $x, \beta \in \mathbf{R}, M = M^T \in \mathbf{R}^{d+1\times d+1}$, and $\tau_i \in \mathbf{R}^l$ [36].

Distributional robustness with unimodal distributions

The worst case distributions associated with the distributionally robust constraint discussed so far are unlikely to be encountered in practice. Another way to reduce the conservativeness is to assume additional properties on the distribution beyond knowledge of the mean and variance. One natural such property is unimodality. Roughly speaking, unimodal distributions have a single mode with large deviations from the mode less likely than small deviations (see [29,25] for precise definitions). All of the distributionally robust formulations so far can also be extended to the unimodal case, based on recent work from Stellato [25] and Van Parys et al. on generalized Gauss inequalities [29].

Consider the distributionally robust problem

$$\begin{aligned} &\text{minimize} \quad f_0(x) \\ &\text{subject to} \quad \mathbf{P}(a(x)^T\delta + b(x) \leqslant 0) \geqslant 1-\alpha, \\ &\qquad\qquad \forall \mathbf{P} \in \mathcal{P}_{uni}(\bar{\delta}, \Xi), \end{aligned} \tag{22}$$

where $\mathcal{P}_{uni}(\bar{\delta}, \Xi)$ is the set of all unimodal distributions of δ with mean $\bar{\delta}$ and variance Ξ. The constraint can again be expressed as a single second-order cone constraint

$$a(x)^T\bar{\delta} + b(x) + f_{uni}(\alpha)\|\Xi^{1/2}a(x)\|_2 \leqslant 0, \tag{23}$$

where $f_{uni}(\alpha)$ is a safety factor function associated with unimodal distributions described in [25]. Note again that the constraint has the same form as both (11) and (15) but with an intermediate safety factor function, yielding a constraint that does not require assuming a normal distribution, which may underestimate risk of violation, and is less conservative than the constraint from the Chebyshev approximation.

Pricing under probabilistic constraints

Efficient pricing under hard constraints

Standard pricing theory in optimal power flow suggests that efficient nodal prices consist of a common base price plus a location-dependent modification arising from congestion on the network. When uncertainty is added to the OPF problem, the notion of efficient *reserve* prices can be defined. In terms of problem (7), efficient prices for power schedules e_i and reserve policies D_i are any prices λ_i and Π_i for which the optimal OPF solution is to be found among the minimizers of the following expression with respect to e_i and D_i for each participant i:

$$\mathbf{E}J_i(D_i, e_i, x_0^i) - \lambda_i^T e_i - \mathrm{tr}\{\Pi_i^T D_i\}. \tag{24}$$

Expression (24) corresponds to the profit accruing to participant i, and is the component of the Lagrangian of problem (7) corresponding to participant i's decision variables.

In [33] it was shown that efficient prices $\lambda_i^\star$ and $\Pi_i^\star$ for the hard-constrained version of (7) ([33], problem (10)) have the following form, for participant i. For real power output e_i the associated efficient price has the structure

$$\lambda_i^\star := -B_i^T C_i^T (\lambda^\star + \Gamma_i^T \nu^\star), \tag{25}$$

where $\lambda^\star \in \mathbf{R}^T$ is the multiplier at optimum corresponding to constraint (7b) and $\nu^\star \in \mathbf{R}^{2LT}$ is that corresponding to the line constraint. Similarly, reserve policy prices consist of a base component $\Pi^\star$ arising from the system-wide constraint (robust power match constraint) and a locational component $\Gamma_i^T \Psi^\star$ arising from the line constraints:

$$\Pi_i^\star := -B_i^T C_i^T (\Pi^\star + \Gamma_i^T \Psi^\star). \tag{26}$$

As noted in [33], product $B_i^T C_i^T$ is for most practical modeling cases an identity matrix, and the minus sign arises from the sign convention used to write the problem's constraints.

Efficient pricing under distributionally robust constraints

When hard constraints are replaced with the probabilistic constraints used in the present paper, the way in which the efficient prices $\lambda_i^\star$ and $\Pi_i^\star$ are constructed from the optimal dual variables of problem (7) changes. Consider the version of (7) in which the Chebyshev generating function has been used – problem (Distributionally Robust OPF). A partial Lagrangian of (Distributionally Robust OPF) can be formed by relaxing the first three constraints. Assigning multipliers λ and Π to the first two constraints as in [33], and μ_k to each element k of the third constraint, the partial Lagrangian takes the form

$$\begin{aligned}\mathcal{L}(D,e,\lambda,\Pi,\mu) = {} & \mathbf{E}\sum_{i=1}^{N} J_i(A_i x_0^i + B_i(D_i\delta + e_i), D_i\delta + e_i) \\ & + \lambda^T \sum_{i=1}^{N} (r_i + C_i(A_i x_0^i + B_i e_i)) \\ & + \mathrm{tr}\left\{ \Pi^T \sum_{i=1}^{N} (G_i + C_i B_i D_i) \right\} \\ & + \sum_{k=1}^{2LT} \mu_k \left[f_k(D,e,\bar{\delta}) + \sqrt{\frac{1-\alpha_k}{\alpha_k}} \|\Xi^{1/2} [\Phi(D)]_{k\cdot}^T\|_2 \right].\end{aligned} \tag{27}$$

Assuming the prediction error is zero-mean, $\bar{\delta} = 0$, we have

$$f_k(D,e,\bar{\delta}) = [b(e)]_k \tag{28}$$

and

$$\|\Xi^{1/2}[\Phi(D)]_{k\cdot}^T\|_2 = \|D_1\gamma_1^k + \cdots + D_N\gamma_N^k\|_2 \tag{29}$$

for some coefficients $\gamma_1^k, \ldots, \gamma_N^k$ for each k. Lagrangian (27) is not, in contrast to Lagrangian (11) in [33], separable between participants i, since the matrices D_i appear nonlinearly. This means the simple pricing interpretation described above for the hard-constrained version of (7) is not available for this version. However it is possible to write a comparable version of (24) to account for the distributionally-robust constraints:

$$\begin{aligned}& \mathbf{E}J_i(D_i, e_i, x_0^i) - \lambda_i^T e_i \\ & - \mathrm{tr}\left\{ \left(-B_i^T C_i^T \Pi + \sum_{k=1}^{2LT} \mu_k \sqrt{\frac{1-\alpha_k}{\alpha_k}} \Upsilon_{ik} \right)^T D_i \right\}.\end{aligned} \tag{30}$$

The price λ_i has the same form as (25) with ν replaced by $\mu = [\mu_1, \ldots, \mu_{2LT}]^T$. The price associated with matrix D_i has a similar structure except that the locational element differs. Matrix coefficients Υ_{ik} are defined as the matrix derivative of the distributionally-robust constraint function with respect to matrix D_i, evaluated at the optimal solution $(e_i^\star, D_i^\star)$, to problem (Distributionally Robust OPF):

$$\Upsilon_{ik} := \frac{\partial}{\partial D_i}\left[\|\Xi^{1/2}[\Phi(D)]_{k\cdot}^T\|_2 \right]\Big|_{D=D^\star}. \tag{31}$$

Pricing under other probabilistic constraint formulations

Section *B* above assumed a traditional Chebyshev generating function for the probabilistic constraints; the pricing interpretation differs when other generating functions are used. For example, efficient prices under the Markov generating function, i.e., corresponding to the solution to (CVaR OPF), are a function of the disturbance samples that are used to generate the relevant constraints. Unfortunately, the analysis presented above is difficult to repeat in this case, since the left hand side term $\mathbf{E}[f_k(D,e,\delta) + t_k]_+$ is difficult to characterize in terms of a series of realizations of δ.

Operational implications

It is common for many system operators to derive nodal prices from the multipliers returned from an optimal power flow problem. These nodal prices diverge when transmission line constraints are binding. However, as shown in subsections *A* and *B* above, the nature of this price divergence depends on the way the system operator has modeled the constraint. This has implications on the fairness of the prices that result. It is possible that biases in the prices can arise from a failure of the system operator to model the line constraint faithfully, since congestion rents will not be split in a manner that reflects the way participants' actions contribute to hitting the physical constraint. This subject may warrant further study.

Numerical example

This section illustrates the results via a numerical example. We consider the two-bus network shown in Fig. 1. A wind farm and a relatively inexpensive thermal generation unit are connected to bus 1, and a relatively expensive thermal generation unit and a fixed load are connected to bus 2. Table 1 shows the network and device parameters in terms of the notation in Section 'Network model and optimal power flow'. Section 'Static case study' illustrates the relative performance of different approximations described above, and Section 'Dynamic case study' illustrates that the benefit of using time-coupled reserve policies depends on whether the line constraint is treated robustly or probabilistically.

Our numerical example is intentionally simple so that the approaches described above can be compared as clearly as possible. Computational scaling of our methods is of course important. However, a strong indication of the scaling can be found in recent work that solves similar problems, including [30,22] for OPF problems with uncertainty sampling and [3] for OPF problems with second-order cone constraints.

Static case study

We first consider a single-stage stochastic optimal power flow problem to illustrate the basic trade-off between cost and network security in terms of frequency and severity of constraint violation. The wind farm has maximum capacity 700 MW and the forecasted output for the next time step is 500 MW. For illustrative purposes, the forecast errors are drawn from a Gaussian probability distribution with zero mean and a standard deviation of 37.5 MW. The fixed load at bus 2 is 1000 MW. The transmission line from

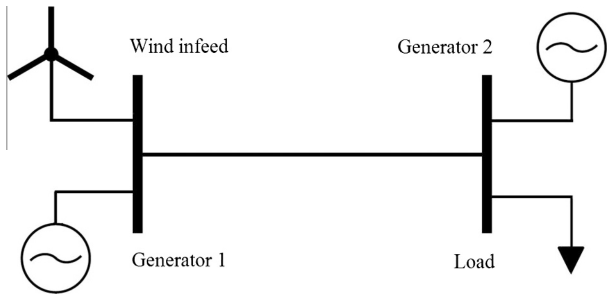

Fig. 1. Two bus power network.

Table 1
Network parameters (generator models as in [33]).

Device	Description
Generator 1	Linear fuel cost \$30/MWh Quadratic fuel cost \$0.05/(MWh)^2 Quadratic ramping cost \$1/(MWh)^2
Generator 2	Linear fuel cost \$60/MWh Quadratic fuel cost \$0.10/(MWh)^2
Wind infeed	See description in Table 2
Load	Fixed at 1000 MW, no uncertainty
Transmission line	Maximum rating 950 MW

bus 1 to bus 2 has a maximum rating of 950 MW. In this example, we consider only this line constraint; there are no other local device constraints.

In this example, there is a trade-off between cost and system security. To minimize cost, one would like to use a larger share of the less expensive thermal generation unit at bus 1, but committing too much from this generator may overload the line if the wind output is much higher than expected. In particular, if the line constraint is ignored, then the optimal affine policies are

$$e_1 = 433, \quad D_1 = -0.67, \quad e_2 = 67, \quad D_2 = -0.33, \tag{32}$$

which means that the nominal injections from generators 1 and 2 are 433 MW and 67 MW, respectively, and that generators 1 and 2 agree to adjust their injections in the event of wind power excess or shortage by 67% and 33% of forecast error, respectively. Under this policy, the line constraint is violated with a frequency of about 9% by about 6.5 MW on average from integrating the Gaussian distribution.

If the constraints are enforced robustly as in [33] based on an assumption that the forecast error is upper bounded by 200 MW, then the optimal affine policies are

$$e_1 = 431.6, \quad D_1 = -0.91, \quad e_2 = 68.4, \quad D_2 = -0.09. \tag{33}$$

The nominal injections are almost the same, but more of the excess wind power is absorbed by reducing the output of the cheaper generator 1 in order to robustly satisfy the constraint, leading to increased cost. Under this policy, the line constraint is (almost) never violated. The optimal cost associated with the reserve policies, on the other hand, is 26% higher than the case in which the line constraint is ignored.

The line constraint can be softened to reduce costs by allowing limited violation, with a limit on the frequency of violation and a penalty on the severity of violation. The trade-off can be explicitly adjusted by changing the parameter α which governs the allowable frequency of violation and by choosing the type of constraint reformulation, and one can effectively interpolate between ignoring the line constraint and enforcing it robustly.[1]

Fig. 2 shows how the optimal cost varies with the constraint violation parameter α in relation to the no constraint and robust cases for four different stochastic reformulations of the line constraint: a chance constraint assuming that the forecast error is Gaussian using (11), the Markov approximation using (14) and evaluating the expectation with 1000 samples, the traditional Chebyshev approximation using (15), and a chance constraint using the scenario approach.[2] The specified constraint violation

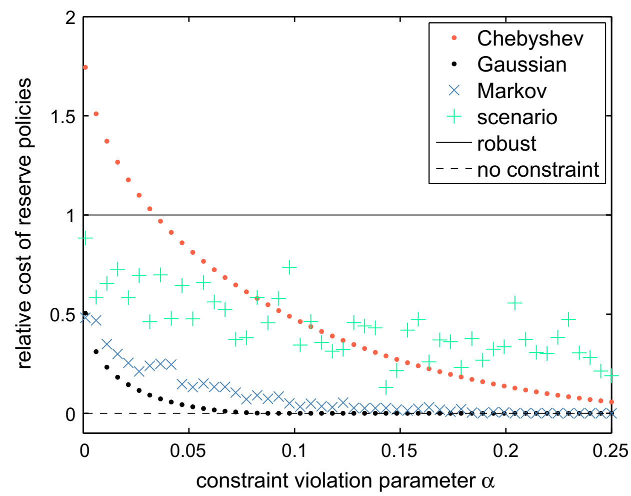

Fig. 2. Optimal cost of operating reserves vs. the constraint violation parameter for different stochastic reformulations of the line constraint.

level for the Gaussian chance constraint matches the actual violation level since the uncertainty used in this example is Gaussian (for $\alpha = 0.09$, the cost is the same as ignoring the constraint); however, if the uncertainty is not Gaussian, then this constraint can underestimate the risk. As expected, the Markov approximation is more conservative than the Gaussian chance constraint because it includes a penalty on severity of violation, but it is only marginally more conservative. It is less smooth than the Gaussian and Chebyshev cases due to the sample estimation of the expectation. The Chebyshev approximation is more conservative still and can even be more conservative than the robust case for small values of α since uncertainty bounds are not explicitly accounted for. The chance constraint with scenario approach is also less smooth than the Chebyshev and Gaussian cases and in this case is more conservative than the Markov approximation due to sampling.

Each type of constraint reformulation gives a different cost and different probabilistic guarantees and penalties on constraint violations. The most appropriate reformulation depends on many factors, and it is possible to mix and match different constraint types for different constraints.

Dynamic case study

A synergy between the use of time-coupled reserve decisions and the use of stochastic constraints is revealed when the two are combined. We demonstrate this by considering a dynamic case study in which expected short-run operating costs are minimized over a limited time horizon.

As demonstrated in [33], it is instructive to compare the cost outcomes under two different structural restrictions on the

[1] An alternative method to achieve this interpolation is to combine robust and chance constraint approaches by enforcing robustness to a subset of the support rather than the full support [2,17]. If the subset is chosen appropriately, one effectively enforces a chance constraint with a certain amount of probability mass cut out by the subset. The interpolation is achieved by adjusting the size of the subset.

[2] We used a standard scenario approach with confidence parameter 10^{-6} described in [5]. There are more sophisticated variations that can be used to reduce conservatism by over-sampling and strategically removing samples.

Table 2
Specification of uncertain wind infeed.

Static case study, Section 'Static case study':
Nominal infeed 500 MW.
Stochastic case: $\mathbf{E}[\delta] = 0, \Xi = 37.5^2$ MW^2.
Robust case: $\delta \leqslant 200$ MW
Dynamic case study, Section 'Dynamic case study':
Nom. infeed [500.0 584.1 590.9 514.1 424.3 404.1 472.1 565.7] MW.
Stochastic case: $\mathbf{E}[\delta] = 0, \Xi$ (in units 10^3 MW^2):

$$\begin{bmatrix} 1.05 & 1.02 & 1.04 & 1.06 & 1.07 & 1.02 & 1.06 & 1.07 \\ 1.02 & 2.01 & 2.00 & 2.02 & 2.04 & 1.97 & 2.02 & 1.99 \\ 1.04 & 2.00 & 3.07 & 3.17 & 3.18 & 3.10 & 3.14 & 3.09 \\ 1.06 & 2.02 & 3.17 & 4.34 & 4.36 & 4.25 & 4.25 & 4.20 \\ 1.07 & 2.04 & 3.18 & 4.36 & 5.49 & 5.36 & 5.36 & 5.28 \\ 1.02 & 1.97 & 3.10 & 4.25 & 5.36 & 6.26 & 6.24 & 6.19 \\ 1.06 & 2.02 & 3.14 & 4.25 & 5.36 & 6.24 & 7.31 & 7.25 \\ 1.07 & 1.99 & 3.09 & 4.20 & 5.28 & 6.19 & 7.25 & 8.25 \end{bmatrix}$$

Robust case: $-3 \cdot [\Xi^{1/2}]_{k,k} \leqslant \delta_k \leqslant 3 \cdot [\Xi^{1/2}]_{k,k}$ MW, $\forall k$.

Table 3
Costs for dynamic case study.

Test	Robust	Gaussian chance constraint, $\alpha = 0.09$
Full LT policy	\$64,489	\$49,437
Diagonal policy	\$65,183	\$49,991
Cost increase	+0.45%	+1.01%

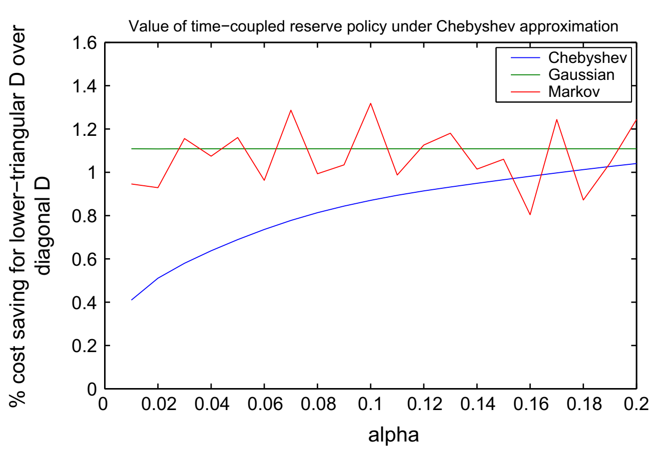

Fig. 3. Relative benefit of time-coupled response to uncertainty observed under different approximate treatments of the chance constraint, for different risk parameters α.

matrices D_i (note that these matrices were scalars in the static case study above):

(1) Diagonal-only: $[D_i]_{j,k} = 0$ for $j \neq k$. This represents the best possible linear, time-decoupled response to uncertainty.
(2) Full lower-triangular: $[D_i]_{j,k} = 0$ for $k > j$. This represents the best causal, linear, time-coupled response to uncertainty.

The expected operating costs were minimized given a current operating point of 250 MW for both generators, and the nominal wind infeed forecast and uncertainty statistics given in Table 2 (Ξ was generated with the Monte Carlo model used in [33]). The following cases were compared: (1) the line flow constraint is enforced robustly assuming the uncertainty δ is restricted to the set Δ described in Table 2; (2) the line flow constraint is enforced in a probabilistic sense, using a conditional value at risk constraint based on the Markov function, a distributionally robust constraint based on the traditional Chebyshev function, and a Gaussian-assumption chance constraint.

The results are shown in Table 3. The benefit of allowing full lower-triangular decision rules was 0.45% in the robust case, and 1.01% in the Gaussian chance constraint case. In other words, the benefit of using a time-coupled response to uncertainty was greater when the constraint was treated probabilistically as opposed to robustly.

Results for the different approximations of the chance constraint are shown in Fig. 3. While the Markov and Gaussian-assumption approaches report a consistent benefit for time-coupled responses to uncertainty, the Chebyshev approximation brings about a lower relative benefit from time-coupled policies at lower risk levels.

Conclusions and outlook

A stochastic optimal power flow problem was formulated, for which a family of convex approximations can be used to trade off cost against security in different ways. We highlighted two broad approaches that emerge from our analysis and occupy the space in between chance constrained and robust approaches that have until now been the main focus in power systems applications. In a conditional value at risk approach, a limit is placed on both frequency and severity of constraint violation. This approach offers a good approximation but at a potentially higher computational cost because sampling is required. In a distributionally robust approach, stochastic constraints can be expressed as single second-order cone constraints and require knowledge of only point forecasts and the variance of forecast errors. This approach offers relative computational simplicity but can lead to conservative results. We also described some variations on this approach that can reduce conservatism. A simple numerical example illustrated the basic trade-offs of economic efficiency and system security. The dynamic case study demonstrated that the apparent benefit of planning a time-coupled response to uncertainty depends strongly on how the problem's constraints are approximated.

The use of probabilistic constraints also has interesting implications on the way in which nodal prices are derived. The differences in nodal prices arising from different line constraint characterizations could be studied in more detail.

In future work the stochastic formulation of network and device constraints considered in this paper could be explored in OPF variations such as unit commitment, security constrained, or relaxations of full AC models. Of significant practical interest would also be to develop distributed optimization algorithms for large scale networks that can be implemented on modern distributed computing platforms.

Acknowledgment

T. Summers is partially supported by the ETH Zurich Postdoctoral Fellowship Program.

References

[1] Alsac O, Bright J, Prais M, Stott B. Further developments in lp-based optimal power flow. IEEE Trans Power Syst 1990;5(3):697–711.
[2] Bertsimas D, Brown DB. Constructing uncertainty sets for robust linear optimization. Oper Res 2009;57(6):1483–95.
[3] Bienstock D, Chertkov M, Harnett S. Chance-constrained optimal power flow: risk-aware network control under uncertainty. SIAM Rev 2014;56(3):461–95.
[4] Boyd S, Vandenberghe L. Convex optimization. Cambridge University Press; 2004.
[5] Calafiore G, Campi MC. Uncertain convex programs: randomized solutions and confidence levels. Math Program 2005;102(1):25–46.
[6] Calafiore GC, El Ghaoui L. On distributionally robust chance-constrained linear programs. J Optim Theory Appl 2006;130(1):1–22.
[7] Campi MC, Garatti G. The exact feasibility of randomized solutions of uncertain convex programs. SIAM J Optim 2008;19(3):1211–30.

[8] Campi MC, Garatti S. A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality. J Optim Theory Appl 2011;148(2):257–80.
[9] Capitanescu F, Fliscounakis S, Panciatici P, Wehenkel L. Cautious operation planning under uncertainties. IEEE Trans Power Syst 2012;27(4):1859–69.
[10] Christie RD, Wollenberg BF, Wangensteen I. Transmission management in the deregulated environment. Proc IEEE 2000;88(2):170–95.
[11] Conejo AJ, Carrion M, Morales JM. Electricity markets. In: Decision making under uncertainty in electricity markets. International series in operations research & management science, vol. 153. USA: Springer; 2010. p. 1–26.
[12] Delage E, Ye Y. Distributionally robust optimization under moment uncertainty with application to data-driven problems. Oper Res 2010;58(3):595–612.
[13] Goh J, Sim M. Distributionally robust optimization and its tractable approximations. Oper Res 2010;58(4-part-1):902–17.
[14] Goulart PJ, Kerrigan EC, Maciejowski JM. Optimization over state feedback policies for robust control with constraints. Automatica 2006;42(4):523–33.
[15] Linderoth J, Shapiro A, Wright S. The empirical behavior of sampling methods for stochastic programming. Ann Oper Res 2006;142(1):215–41.
[16] Mak W-K, Morton DP, Wood RK. Monte carlo bounding techniques for determining solution quality in stochastic programs. Oper Res Lett 1999;24(1):47–56.
[17] Margellos K, Goulart P, Lygeros J. On the road between robust optimization and the scenario approach for chance constrained optimization problems. IEEE Trans Autom Control 2014;59(8):2258–63.
[18] Nemirovski A, Shapiro A. Convex approximations of chance constrained programs. SIAM J Optim 2006;17(4):969–96.
[19] Ozturk UA, Mazumdar M, Norman BA. A solution to the stochastic unit commitment problem using chance constrained programming. IEEE Trans Power Syst 2004;19(3):1589–98.
[20] Pinson P. Wind energy: forecasting challenges for its operational management. Stat Sci 2013;28(4):564–85.
[21] Popescu I. A semidefinite programming approach to optimal-moment bounds for convex classes of distributions. Math Oper Res 2005;30(3):632–57.
[22] Roald L, Vrakapoulou M, Oldewurtel F, Andersson G. Risk-constrained optimal power flow with probabilistic guarantees. In: 18th power systems computation conference, Wroclaw, Poland; 2014.
[23] Rockafellar RT, Uryasev S. Optimization of conditional value-at-risk. J Risk 2000;2:21–42.
[24] Shapiro A. Monte carlo sampling methods. Handbooks Oper Res Manage Sci 2003;10:353–425.
[25] Stellato B. Data-driven chance constrained optimization [Master's thesis]. ETH Zürich; 2014.
[26] Stott B, Alsac O. Fast decoupled load flow. IEEE Trans Power Apparatus Syst 1974;PAS-93(3):859–69.
[27] Stott B, Alsac O. Optimal power flow – basic requirements for real-life problems and their solutions. White paper, Arizona; 2012.
[28] Summers TH, Warrington J, Morari M, Lygeros J. Stochastic optimal power flow based on convex relaxations of chance constraints. In: 18th power systems computation conference, Wroclaw, Poland; 2014.
[29] Van Parys BPG, Goulart PJ, Kuhn D. Generalized Gauss inequalities via semidefinite programming. Technical report, ETH Zürich; 2014.
[30] Vrakopoulou M, Margellos K, Lygeros J, Andersson G. A probabilistic framework for reserve scheduling and N-1 security assessment of systems with high wind power penetration. IEEE Trans Power Syst 2013;28(4):3885–96.
[31] Wang Q, Guan Y, Wang J. A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output. IEEE Trans Power Syst 2012;27(1):206–15.
[32] Wang W, Ahmed S. Sample average approximation of expected value constrained stochastic programs. Oper Res Lett 2008;36(5):515–9.
[33] Warrington J, Goulart PJ, Mariethoz S, Morari M. Policy-based reserves for power systems. IEEE Trans Power Syst 2013;28(4):4427–37.
[34] Zhang H, Li P. Chance constrained programming for optimal power flow under uncertainty. IEEE Trans Power Syst 2011;26(4):2417–24.
[35] Zhang Y, Giannikis GB. Robust optimal power flow with wind integration using conditional value-at-risk; 2013.
[36] Zymler S, Kuhn D, Rustem B. Distributionally robust joint chance constraints with second-order moment information. Math Program 2013;137(1–2):167–98.