Stochastic optimal power flow based on convex approximations of chance constraints

Tyler Summers, Joseph Warrington, Manfred Morari, and John Lygeros

Abstract— **This paper presents 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 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 relaxations of chance constraints, and show that these allow economic efficiency and system security to be traded off with varying levels of conservativeness. The results are illustrated using a simple case study, in which conventional generators plan schedules around an uncertain but time-correlated wind power injection.**

I. Introduction

The current widespread increase in penetration of intermittent renewable energy in power networks such as wind and solar 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 prohibitively 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 [6]. 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 tradeoff between economic efficiency and system security.

One of the fundamental decision problems in power networks is optimal power flow (OPF), in which power schedules are determined for controllable devices in a power system, such as generators, storage, and controllable loads, which minimize an operating cost function under various device and network constraints. OPF is central to economic and secure operation and control of power systems and markets [9]. 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 does explicitly handle uncertainty includes methods based on (1) chance constraints and the so-called scenario approach [10], in which decisions are made based on finite sampling of uncertain parameters from an assumed statistical model, or (2) robust optimization [11], in which knowledge of uncertainty bounds is assumed and device and network constraints are enforced for every possible uncertainty realization. However, these approaches have some 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 sufficiently 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.

There are other risk measures, such as expected shortfall and conditional value at risk (CVaR), which are well known in finance [8], but have only received limited attention in the context of power networks. Some of these are discussed in the context of electricity markets in [4]. Specifically, this work mainly 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 [12] also considers the use of CVaR for managing financial risk of wind power producers in a network. Bienstock et al. have also recently considered power

The authors are with the Automatic Control Laboratory, ETH Zurich, Switzerland, email: {tsummers,warrington,morari,lygeros}@control.ee.ethz.ch. T. Summers is partially supported by the ETH Zurich Postdoctoral Fellowship Program.

flow problems subject to chance constraints assuming that the uncertainties are Gaussian [1].

In this paper, we present an approach to solving a stochastic optimal power flow problem based on convex approximations of chance constraints [5]. Three main features taken together distinguish our work from previous work. First, we take the perspective of a transmission system operator managing operational risks across the network rather than individual power supplier and use stochastic formulations of network and device constraints. Second, the family of convex relations of chance constraints we consider, which includes the coherent risk measure CVaR as a special case, allows various penalties to be given to the severity of constraint violations and can interpolate between chance constraints and robust constraints. Another special case, relating to Chebyshev bounds, allows probabilistic information to be incorporated into a single second order cone constraint and does not require any uncertainty sampling. Third, 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.

We demonstrate that this formulation can achieve a trade-off 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 II describes the network model and formulates a stochastic optimal power flow problem that explicitly accounts for information about uncertainty. Section III describes the convex approximations of chance constraints and shows how a family of such approximations can be utilized in the optimal power flow problem. Section IV presents numerical results, and Section V gives concluding remarks and an outlook for future research.

II. 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., thermal generation, deferrable/curtailable loads, or storage devices). The notation follows [11].

A. 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 prediction 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, which captures spatial variations and dependencies among devices and temporal variations and dependencies across the horizon. In particular, we assume either knowledge of the full distribution, knowledge of certain moments such as the mean and variance, or that we have a model of δ from which we can draw samples.

B. 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

$$x_{k+1}^i = \bar{A}_i x_k^i + \bar{B}_i u_k^i \tag{1}$$

where $\bar{A}_i \in \mathbf{R}^{n_i\times n_i}$ is the dynamics matrix, and $\bar{B} \in \mathbf{R}^{n_i\times m_i}$ 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 charge in 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: $\mathbf{x}^i = [(x_1^i)^T, ..., (x_T^i)^T] \in \mathbf{R}^{n_i T}$ and $\mathbf{u}^i = [(u_0^i)^T, ..., (u_{T-1}^i)^T] \in \mathbf{R}^{m_i T}$, 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

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

C. Cost functions and constraints

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

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

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, one can include 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 can be neglected, and the active line flows are proportional to the phase differences between bus voltages.

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

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

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 [3]).

D. 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 $\pi_i : \mathbf{R}^{N_\delta T} \to \mathbf{R}^{m_i T}$ 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 [11]. Alternatively, they could be "softened" and enforced in some weaker probabilistic sense based on assumed knowledge of the uncertainty probability distribution.

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:

$$\begin{aligned}
&\underset{\pi_i \in \Pi_c}{\text{minimize}} \quad \mathbf{E} \sum_{i=1}^{N} J_i(A_i x_0^i + B_i \pi_i(\delta), \pi_i(\delta)) \\
&\text{subject to} \quad \sum_{i=1}^{N} (r_i + G_i \delta + C_i(A_i x_0^i + B_i \pi_i(\delta))) = 0, \ \forall \delta \\
&\mathcal{R}\left(\sum_{i=1}^{N} \Gamma_i (r_i + G_i \delta + C_i(A_i x_0^i + B_i \pi_i(\delta))) - \bar{p} \leq 0 \right) \\
&\mathcal{R}\left(T_i(A_i x_0^i + B_i \pi_i(\delta)) + U_i \pi_i(\delta) - w_i \leq 0 \right), \quad i = 1, ..., N
\end{aligned} \tag{7}$$

where $\mathcal{R}$ denotes a generic transformation of the inequality constraints into stochastic versions, using probabilistic uncertainty information and possibly introducing auxiliary variables, which will be described in the next section.

The infinite-dimensional optimization over the set of admissible causal functions Π_c is intractable. Therefore, 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 considered in [11], the robust constraints are recast to become linear in the decision variables D_i and e_i and some extra auxiliary matrix variables.

The following section will now discuss other reformulations that allow for the possibility of some degree of constraint violation in exchange for reduced solution cost.

III. Convex Approximations of Chance Constraints

In this section, we discuss chance constraints and a family of convex relaxations based on the the results in [5]. For more details, see [5], [8]. At the end of the section we formulate a stochastic OPF problem based on these constraints.

A. Chance constraints

Consider the chance-constrained optimization problem

$$\begin{array}{ll}\text{minimize} & \mathbf{E} f_0(x,\delta) \\ \text{subject to} & \mathbf{Prob}(f(x,\delta) \le 0) \ge 1-\alpha\end{array} \tag{10}$$

where $x \in \mathbf{R}^n$ is the decision variable, $\delta \in \mathbf{R}^d$ is a random variable, and $f(x,\delta) : \mathbf{R}^n \times \mathbf{R}^d \to \mathbf{R}$ is a single scalar constraint function (we discuss how to deal with multiple constraint functions later). This problem is convex in some cases. For example, when $f_i(x,\delta) = a^T x + b$ with $\delta = [a^T \ b]^T$ and δ is normally distributed, then the chance constraint can be expressed as a second-order cone constraint:

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

where Φ^{-1} is the Gaussian quantile function, $\tilde{x} = [x^T \ 1]^T$, and Ξ is the covariance matrix of δ. In other cases, the random parameter is typically sampled from a distribution and a corresponding deterministic optimization problem is solved based on the sampled values. 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 [2]. Chance constraints are closely related to Value at Risk (VaR), a risk measure often used in finance [8].

Chance constraints have several drawbacks. They penalize frequency but not severity of constraint violation. Moreover, the associated VaR is not a "coherent" risk measure in the sense that it has some undesirable properties for certain types of uncertainty distributions [8]. Also, when using a sampling approach, the number of samples required to guarantee a certain probabilistic feasibility level can be large, making the sampled optimization problem difficult to solve and potentially rendering the solution very conservative in practice.

B. Convex approximation of chance constraints and conditional value at risk

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 contains the feasible set of problem (10).

First, note that for a random variable z and for any $t > 0$, $\mathbf{Prob}(tz \ge 0) = \mathbf{Prob}(z \ge 0) = \mathbf{E}[\mathbf{1}_{[0,\infty)}(tz)]$ 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 $\psi(tz) \ge \mathbf{1}(tz)\ \forall tz \in \mathbf{R}$, it follows that $\mathbf{E}\psi(tz) \ge \mathbf{E}\mathbf{1}_{[0,\infty)}(tz) = \mathbf{Prob}(z \ge 0)$, i.e., the function $\mathbf{E}\psi(tz)$ is an upper bound on the probability that $z \ge 0$.

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

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

Thus, the constraint

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

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) [5].

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. The best approximation is given by the generating function that is closest to the indicator function; accordingly, it can be shown that the Markov generating function gives the best approximation for a single scalar chance constraint. The constraint obtained from the Markov generating function is closely related to the Conditional Value at Risk (CVaR), which is also a well known risk measure in finance that penalizes both frequency and severity of constraint violation and is coherent [8]. It can be written as

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

The other candidates give more conservative approximations. However, one advantage of using a smooth generating function, such as the traditional Chebyshev or the Chernoff/Bernstein, is that in some cases we can explicitly evaluate the expression in (13); otherwise, one must resort to an uncertainty sampling method.

1) Example: Traditional Chebyshev approximation for an affine inequality: To illustrate a case in which the constraint can be expressed analytically, we consider here a specific example of a single affine inequality and a traditional Chebyshev approximation of a corresponding chance constraint. Let $f(x,\delta) = a^T x + b$, where $\delta = [a^T \ b]^T$ is random with mean $\mathbf{E}\delta = \bar{\delta}$ and variance $\mathbf{E}\delta\delta^T - \bar{\delta}\bar{\delta}^T := \Xi$. Let $\tilde{x} = [x^T \ 1]^T$. The constraint obtained from the traditional Chebyshev generating function can be written as

$$\bar{\delta}^T \tilde{x} + \sqrt{\frac{1-\alpha}{\alpha}}\|\Xi^{1/2}\tilde{x}\|_2 \le 0, \tag{15}$$

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. The constraint associated with the traditional Chebyshev generating function is "distributionally robust" in the sense that the constraint will hold for any distribution of δ with the given mean and variance. This can be very conservative, but the conservatism can be reduced by assuming more about the distribution, e.g., unimodality and/or knowledge of higher moments [7].

2) Multiple constraints: The preceding discussion was for a single scalar inequality. There are several ways that the approximation could be extended to handle multiple inequalities $f_i(x,\delta) \le 0$, $i = 1, ..., m$. One could construct a single constraint from the set, for example: $\max_i f_i(x,\delta) \le$

0. It is also possible to define multivariable generating functions that can be directly used with multiple constraints. Finally, one can treat each constraint separately and use the development above, obtaining the transformed constraints $\inf_{t>0}\left(t\mathbf{E}[\psi(t^{-1}f_i(x,\delta))] - t\alpha_i\right) \le 0$ for a set of risk levels α_i associated with each constraint [5].

C. Stochastic OPF with convex approximations of chance constraints

Based on the above approximations, we can formulate 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

$$f_k(x,\delta) = [\Phi]_{k\cdot}\delta + [b]_k \tag{16}$$

where $[\cdot]_{k\cdot}$ denotes the k^{th} row of a matrix and $[\cdot]_k$ the k^{th} element of a vector. The optimization variables D_i and e_i for $i = 1, \ldots, N$ enter linearly into x and b as follows:

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

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

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 $g_{ij}(x,\delta)$.

Using, for example, the Markov generating function, this leads to the following stochastic optimal power flow problem

$$\begin{aligned}
&\underset{D,e,t}{\text{minimize}} \quad \mathbf{E}\sum_{i=1}^{N} J_i(A_i x_0^i + B_i(D_i\delta + e_i), D_i\delta + e_i)\\
&\text{subject to} \sum_{i=1}^{N}(r_i + C_i(A_i x_0^i + B_i e_i)) = 0,\\
&\sum_{i=1}^{N}(G_i + C_i B_i D_i) = 0\\
&\mathbf{E}[f_k(D,e,\delta) + t_k]_+ \le t_k\alpha_k\,,\ \ k = 1, \ldots, 2LT\\
&\mathbf{E}[g_{ij}(D_i,e_i,\delta) + t_{ij}]_+ \le t_{ij}\alpha_{ij}\,,\\
&\qquad\qquad i = 1, \ldots, N, j = 1, \ldots, l_i.
\end{aligned} \tag{19}$$

where $D := (D_1, \ldots, D_N)$, $e := (e_1, \ldots, e_N)$, and $t := (t_1, \ldots, t_{2LT}, t_{1,1}, \ldots, t_{N,l_N})$, which is a convex optimization problem in (D, e, t). In this case, the expected value in the last two sets of constraints would be approximated by a sample average. Other generating functions can be used to obtain similar formulations. If the traditional Chebyshev generating function is used, the last two sets of constraints would each be a second order cone constraint and no sampling of the uncertainty would be required.

D. Interpretation of probabilistic constraints

The constraint reformulations above are best applied to situations where the occasional violation of a constraint can (a) be accepted, and (b) makes physical sense. This is the case for the transmission line limits modeled by $f_k(D,e,\delta)$ above, since the true operating capabilities of many such lines depend on their temperature, which can cause the line to sag, rather than a hard current rating.

The same consideration must be employed more selectively when considering the operation of devices such as generators and storage units connected to the network, since these constraints may combine hard limits (for example a zero power output bound) and soft limits (for example a generator ramp rate limit). These considerations can be accommodated by choosing various values for α_{ij} when adapting constraints $g_{ij}(D_i, e_i, \delta)$; a lower α-value corresponds to a lower tolerance for constraint violation.

IV. Numerical Example

This section illustrates the method 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 I shows the network and device parameters in terms of the notation in Section II. Subsection IV-A illustrates the relative performance of different approximations described above, and Subsection IV-B illustrates that the benefit of using time-coupled reserve policies depends on whether the line constraint is treated robustly or probabilistically.

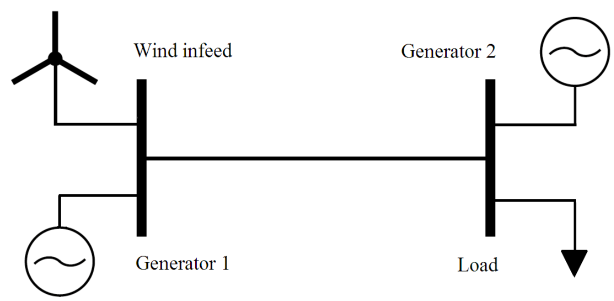

Fig. 1. Two bus power network.

TABLE I
NETWORK PARAMETERS (GENERATOR MODELS AS IN [11]).

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 II
Load	Fixed at 1,000 MW, no uncertainty
Transmission line	Maximum rating 950 MW

A. Static case study

We first consider a single-stage stochastic optimal power flow problem to illustrate the basic tradeoff 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 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

$$e_1 = 433, \; D_1 = -0.67, \quad e_2 = 67, \; D_2 = -0.33, \tag{20}$$

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.

If the constraints are enforced robustly as in [11] based on an assumption that the forecast error is upper bounded by 200 MW, then the optimal affine policies are

$$e_1 = 431.6, \; D_1 = -0.91, \quad e_2 = 68.4, \; D_2 = -0.09. \tag{21}$$

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 never violated, but the optimal cost associated with the reserve policies is increased by 26% over the case in which the line constraint is ignored.

As explained previously, the line constraint can be softened to reduce costs by allowing limited violation in a specific probabilistic sense, 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.

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[1]. The specified constraint violation 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 due to sampling and in this case is more conservative than the Markov case even though it gives weaker probabilistic guarantees.

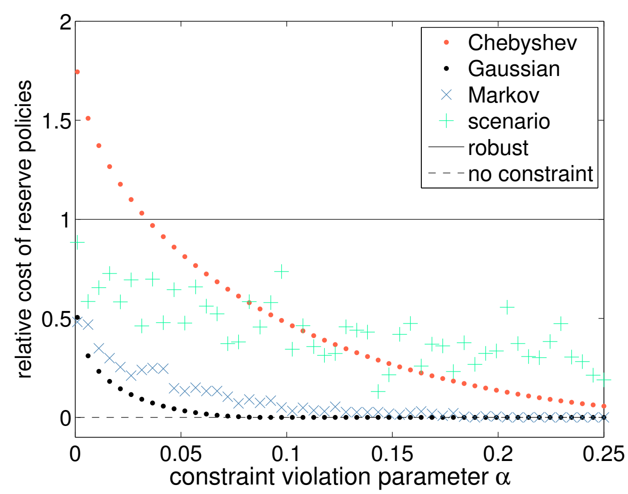

Fig. 2. Optimal cost of operating reserves vs. the constraint violation parameter for different stochastic reformulations of the line constraint.

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.

B. Dynamic case study

A synergy between the use of time-coupled reserve decisions and the use of approximated chance 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 [11], it is instructive to compare the cost outcomes under two different structural restrictions on the 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 that could be planned.

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

TABLE II

SPECIFICATION OF UNCERTAIN WIND INFEED

Static case study, Section IV-A:
Nominal infeed 500 MW.
Stochastic case: $\mathbf{E}[\delta] = 0$, $\Xi = 37.5^2$ MW^2.
Robust case: $\delta \leq 200$ MW

Dynamic case study, Section IV-B:
Nom. infeed [500.0 584.1 590.9 514.1 424.3 404.1 472.1 565.7] MW.
Stochastic case: $\mathbf{E}[\delta] = 0$, Ξ (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: $-3 \cdot [\Xi^{1/2}]_{k,k} \leq \delta_k \leq 3 \cdot [\Xi^{1/2}]_{k,k}$ MW, $\forall k$.

TABLE III

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%

2) Full lower-triangular: $[D_i]_{j,k} = 0$ for $k > j$. This represents the best causal, linear, time-coupled response to uncertainty that could be planned.

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 II (Ξ was generated with the Monte Carlo model used in [11]). 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 II; (2) the line flow constraint is enforced in a probabilistic set, using the Gaussian-assumption chance constraint.

The results are shown in Table III. 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.

V. CONCLUSIONS AND OUTLOOK

A chance-constrained stochastic optimal power flow problem was formulated, for which a family of convex approximations can be used in order to trade off cost against security. It was shown that the Chebyshev CVaR approximation often leads to conservative results but has computational advantages because it can be expressed as a single second order cone constraint. In contrast, the restriction to a Gaussian assumption on the uncertainty leads to lower-cost solutions at the expense of realism (the uncertainty may be more heavy-tailed than a Gaussian and therefore constraints may be violated more frequently). The sampling-based Markov approach offers a good approximation of the constraint but at a potentially high computational cost. The dynamic case study demonstrated that the apparent benefit of planning a time-coupled response to uncertainty depends rather strongly on how the problem's constraints are approximated.

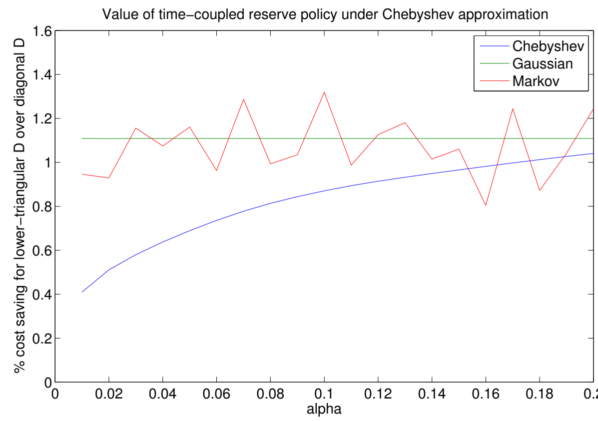

Fig. 3. Relative benefit of time-coupled response to uncertainty observed under different approximate treatments of the chance constraint, for different risk parameters α.

REFERENCES

[1] D. Bienstock, M. Chertkov, and S. Harnett. Chance constrained optimal power flow: Risk-aware network control under uncertainty. *arXiv preprint arXiv:1209.5779*, 2012.

[2] G. Calafiore and M. C. Campi. Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming*, 102(1):25–46, January 2005.

[3] R.D. Christie, B.F. Wollenberg, and I. Wangensteen. Transmission management in the deregulated environment. *Proc. of the IEEE*, 88(2):170–195, 2000.

[4] A. J. Conejo, M. Carrion, and J. M. Morales. Electricity markets. In *Decision Making Under Uncertainty in Electricity Markets*, number 153 in International Series in Operations Research & Management Science, pages 1–26. Springer US, January 2010.

[5] A. Nemirovski and A. Shapiro. Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969996, 2006.

[6] P. Pinson. Wind energy: Forecasting challenges for its operational management. *Preprint*, 2013.

[7] I. Popescu. A semidefinite programming approach to optimal-moment bounds for convex classes of distributions. *Mathematics of Operations Research*, 30(3):632–657, 2005.

[8] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of risk*, 2:2142, 2000.

[9] B. Stott and O. Alsac. Optimal power flow - basic requirements for real-life problems and their solutions. White paper, Arizona, 2012.

[10] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson. A probabilistic framework for security constrained reserve scheduling of networks with wind power generation. In *2nd IEEE Conference on Energy and Exhibition*, Florence, Italy, 2012.

[11] J. Warrington, P.J. Goulart, S. Mariethoz, and M. Morari. Policy-based reserves for power systems. *IEEE Transactions on Power Systems*, 28(4):4427–4437, 2013.

[12] Y. Zhang and G.B. Giannikis. Robust optimal power flow with wind integration using conditional value-at-risk. 2013.