Stochastic DC Optimal Power Flow With Reserve Saturation

Rohit Kannan, James R. Luedtke, Line A. Roald
University of Wisconsin-Madison, Madison, WI, USA
E-mail: {rohit.kannan, jim.luedtke}@wisc.edu, Corresponding Author: roald@wisc.edu

***Abstract*—We propose an optimization framework for stochastic optimal power flow with uncertain loads and renewable generator capacity. Our model follows previous work in assuming that generator outputs respond to load imbalances according to an affine control policy, but introduces a model of *saturation* of generator reserves by assuming that when a generator's target level hits its limit, it abandons the affine policy and produces at that limit. This is a particularly interesting feature in models where wind power plants, which have uncertain upper generation limits, are scheduled to provide reserves to balance load fluctuations. The resulting model is a nonsmooth nonconvex two-stage stochastic program, and we use a stochastic approximation method to find stationary solutions to a smooth approximation. Computational results on 6-bus and 118-bus test instances demonstrate that by considering the effects of saturation, our model can yield solutions with lower expected generation costs (at the same target line violation probability level) than those obtained from a model that enforces the affine policy to stay within generator limits with high probability.**

***Index Terms*—Optimal power flow, renewables integration, generation limits, stochastic programming.**

I. Introduction

Large shares of renewable energy increases the variability and uncertainty in power grid operations, and frequently leads to a larger demand for balancing energy through generation reserves. Understanding and counteracting potentially adverse effects of this uncertainty requires models that accurately capture its impact on the network. Ignoring the effect of uncertainties while making dispatching decisions can result in unsafe operations [1], whereas considering them can significantly improve system security while simultaneously enabling economic efficiency [2]. Many approaches to stochastic optimal power flow (OPF) problems have typically relied on affine generation control policies to balance fluctuating power demands, mimicking the actions of the automatic generation control [1–5]. These policies require traditional generators to provide a determined fraction of the necessary reserves. The feasibility of the affine control policy is typically enforced using conservative chance-constrained approximations [1, 3, 5], robust constraints [6], or by constraining the expected exceedance of determined reserves [2, 4]. A key limitation of the affine control policy is that it does not adequately model the behavior of the generators as they reach their upper or lower generation limits [2, 4]. When the system faces large demand fluctuations, some generators are likely to hit their limits if the affine policy is used, in which case a realistic generator will simply stop providing reserves and maintain a fixed power output. Failing to model this behavior may result in conservative results with economically inferior dispatching decisions because requiring feasibility of the affine policy forces each generator to maintain a too large reserve capacity. This drawback becomes more pronounced when considering reserves from uncertain resources, such as reserves provided by renewable generators themselves [2] or demand response resources [5].

To address this drawback, we introduce a new optimization model that includes a more realistic and flexible representation of reserve activation and captures the impact of upper and lower generator limits, which we call *reserve saturation*. While this reserve saturation model is an accurate reflection of current system operations, it has, to the best of our knowledge, never before been considered in the context of stochastic OPF. Related work resets the affine control policy through activation of manual reserves [4], imposes hard limitations on wind power generation [2], or use multi-parametric programming as a preprocessing step [7]. However, all these methods require the user to pre-specify important aspects of the piecewise-affine policies, leading to potentially sub-optimal solutions. In contrast, we introduce a two-stage stochastic formulation for the DC OPF problem that includes the reserve saturation model in the second-stage, and thus inherently incorporates and enforces power generation limits in conventional and wind generators. By explicitly enforcing the piecewise control policy as a second-stage constraint, the optimization problem is able to identify the optimal generation and reserve allocation considering this behavior, without any pre-specified (and potentially sub-optimal) input.

After introducing the model, we investigate conditions under which it is feasible and derive a stochastic approximation method for solving a smooth version of the model to local optimality. Finally, we demonstrate the practical benefits of our modeling framework on case studies based on a small 6-bus system and the IEEE 118-bus system. In particular, we assess the economic and environmental impact of allowing wind power plants to provide significant reserves, and demonstrate empirically that our approach finds solutions that satisfy

This research is supported by the Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Applied Mathematics program under Contract Number DE-AC02-06CH11347.

physical limits with high probability through out-of-sample testing.

This paper is organized as follows. Section II outlines our reserve saturation model within a two-stage stochastic programming framework, and Section III presents a stochastic approximation method for solving an approximation. Section IV briefly discusses alternative modeling approaches for determining candidate first-stage solutions. Computational results are reported in Section V, and we conclude in Section VI.

Notation. We denote vectors by lower case letters and their components using subscripts. We let $\mathrm{int}(S)$ denote the interior of a set S, write $(\cdot)_+$ and $(\cdot)_-$ to denote $\max\{\cdot, 0\}$ and $\min\{\cdot, 0\}$, $\mathrm{logspace}(a, b, n)$ to denote a vector of n logarithmically-spaced points between 10^a and 10^b (both inclusive), and write $\mathbb{E}[\cdot]$ and $\sigma[\cdot]$ to denote expectation and standard deviation operators. We do not make a notational distinction between random variables and their realizations.

II. Optimal power flow with reserve saturation

We introduce a two-stage stochastic programming model for determining power generation and reserve levels in a power system facing random loads and wind generation uncertainty. In the first stage, the nominal generation levels, reserve capacities and reserve participation factors for each generator are determined. These decisions are taken in advance of observing the random demand and wind generation capacity. The second stage models the system response to the observed load and wind generation. This response, which requires the generators to activate reserves to balance the system, is determined by the reserve participation factors from the first-stage of the optimization model. In our model, the random loads are uncertain and non-dispatchable, representing a combination of standard load and non-dispatchable renewable generation. We assume that wind power plants are fully dispatchable, except that their output is capped by the random available capacity.

A novel feature of our model is that we explicitly model *generator saturation* in the second-stage formulation, which occurs when the output of a generator, as determined by its nominal generation level, participation factor, load imbalance, and the control policy reaches its upper or lower generation limit. A generator that reaches its lower/upper limit continues to produce at that limit, and any additional balancing energy must be provided by the remaining generators that have not yet reached saturation. In this model, generators are allowed to exceed their scheduled reserve capacity, but we assume the system operator pays a higher price for doing so. While generation limits are satisfied by virtue of our modeling framework, we use a penalty on the the expected violation of line limits to obtain a solution that satisfies the line limits with high probability. The objective in our model is hence to minimize the expected generation costs while keeping the expected violation of the line limits small.

A. Network representation

We model the network as an undirected connected graph $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ denotes the set of nodes/buses and $\mathcal{E}$ denotes the set of edges/transmission lines. The set of wind generators, regular generators, and loads are denoted by $\mathcal{W}$, $\mathcal{R}$, and $\mathcal{D}$, respectively, and $\mathcal{G} := \mathcal{R} \cup \mathcal{W}$ denotes the set of all generators. The demand at node $i \in \mathcal{D}$ is a random variable having expected value d_i. We assume for notational convenience that each node in the network houses a load *and* either a wind generator, or a regular generator. It is straightforward to extend the model to include nodes with no/multiple generators/loads.

B. First-stage decisions and constraints

The first-stage decisions include, for each generator $i \in \mathcal{G}$, the nominal generation levels p^0_i, the scheduled up- and down-reserve levels r^+_i and r^-_i, and the generator participation factors for reserves, α_i. Note that the set of dispatchable generators includes the wind generators. To ensure consistency of our control policy (cf. [1, 2]), we require that the nominal generation levels satisfy a power balance constraint for the expected value of the demands and lie within pre-specified bounds ($p^{0,\mathrm{L}}$ and $p^{0,\mathrm{U}}$):

$$\sum_{i \in \mathcal{G}} p^0_i = \sum_{j \in \mathcal{D}} d_j, \quad p^{0,\mathrm{L}} \leq p^0 \leq p^{0,\mathrm{U}}. \tag{1}$$

For regular generators, $p^{0,\mathrm{L}} = p^{\min}$ and $p^{0,\mathrm{U}} = p^{\max}$ represent the upper and lower generation limits. For wind power plants, $p^{0,\mathrm{L}}$ and $p^{0,\mathrm{U}}$ represent the maximum and minimum generation that the operator is willing to schedule from that plant.

The up- and down-reserve levels r^+ and r^- are constrained to lie within pre-specified limits ($r^{+,\max}, r^{-,\max}$), and comply with capacity limits for regular generators:

$$0 \leq r^+ \leq r^{+,\max}, \quad 0 \leq r^- \leq r^{-,\max}, \tag{2}$$

$$p^0_i + r^+_i \leq p^{\max}_i, \quad p^0_i - r^-_i \geq p^{\min}_i, \quad \forall i \in \mathcal{R}. \tag{3}$$

We assume that the reserve activation follows through the automatic generation control (AGC), where the contribution of each generator is determined through a participation factor [8]. The participation factors α are required to sum to one, and a subset of generators $\mathcal{G}^{\mathrm{res}} \subset \mathcal{G}$ are required to provide reserves with a minimum participation factor ε:

$$\alpha \geq 0, \quad \sum_{i \in \mathcal{G}} \alpha_i = 1, \quad \alpha_i \geq \varepsilon, \ \forall i \in \mathcal{G}^{\mathrm{res}}. \tag{4}$$

Let $\mathcal{R}es(\alpha) := \{i \in \mathcal{G} : \alpha_i > 0\}$ denote the set of generators with positive participation factors for any choice of α satisfying constraint (4), and note that $\mathcal{G}^{\mathrm{res}} \subset \mathcal{R}es(\alpha)$.

C. Uncertain parameters and the recourse policy

We let ω denote the underlying random variables, and assume that we can generate iid samples from its probability distribution. Let $\tilde{d}_i(\omega)$ represent the random fluctuations in the power demands for $i \in \mathcal{D}$, and $p^{\min}_i(\omega)$ and $p^{\max}_i(\omega)$, $i \in \mathcal{W}$, represent the minimum and maximum wind generator power outputs. For notational simplicity, we also define $p^{\max}_i(\omega) \equiv p^{\max}_i$ and $p^{\min}_i(\omega) \equiv p^{\min}_i$ for $i \in \mathcal{R}$. Let $\Sigma_d(\omega) := \sum_{i \in \mathcal{D}} \tilde{d}_i(\omega)$ denote the net demand fluctuation.

A common assumption in power system modeling [1, 2, 5, 7, 8] is that the AGC reserve activation to balance the load fluctuation $\Sigma_d(\omega)$ can be modelled as an affine control policy. This affine policy adjusts the generation of the regular and wind generators as

$$p_i(\omega) = p_i^0 + \alpha_i \Sigma_d(\omega), \quad \forall i \in \mathcal{G}, \tag{5}$$

where $p_i(\omega)$ denotes the power output of generator $i \in \mathcal{G}$ for a realization of the random variables ω. While the affine policy satisfies the total power balance constraint by virtue of Eqns. (1) and (4), the generation levels $p_i(\omega)$ determined by this policy could exceed the generation limits $p_i^{\min}(\omega)$ and $p_i^{\max}(\omega)$ if the magnitude of the deviation $\Sigma_d(\omega)$ is large. To avoid such violations, existing approaches [1, 2, 5, 7] impose tight constraints on the probability or expected magnitude of generation limit violations, leading to conservative nominal generation levels p^0 and allocation of the participation factors α, and preventing optimal use of generation capacity.

We propose a more realistic and physically accurate model that includes *reserve saturation*. This model allows generators to provide reserves with a determined participation factor *only* until they hit their generation limits, after which other non-saturated generators are required to contribute additional reserves according to their relative participation factors. To represent this model, we first define *target generation levels* $p_i^{\mathrm{T}}(\omega)$ (which may violate generation limits) as follows:

$$p_i^{\mathrm{T}}(\omega) = p_i^0 + \alpha_i \Sigma_d(\omega) + \alpha_i s(\omega), \quad \forall i \in \mathcal{G}, \tag{6}$$

where $s(\omega)$ is a slack reserves variable which represents the imbalance incurred by generators that have reached their bounds and are no longer contributing reserves. Note that in this model, $s(\omega)$ will be nonzero only if at least one generator is saturated before the load is balanced.

For each generator $i \in \mathcal{G}$, we now determine the *actual generation level* $p_i(\omega)$ (that honors generation limits) using the piecewise-affine policy

$$p_i(\omega) = \begin{cases} p_i^{\min}(\omega), & \text{if } p_i^{\mathrm{T}}(\omega) < p_i^{\min}(\omega) \\ p_i^{\mathrm{T}}(\omega), & \text{if } p_i^{\min}(\omega) \le p_i^{\mathrm{T}}(\omega) \le p_i^{\max}(\omega) \\ p_i^{\max}(\omega), & \text{if } p_i^{\mathrm{T}}(\omega) > p_i^{\max}(\omega). \end{cases} \tag{7}$$

By including reserve saturation, the generators follow their target generation as long as $p_i^{\min}(\omega) \le p_i^{\mathrm{T}}(\omega) \le p_i^{\max}(\omega)$, but remain at their upper or lower bound if the limits are exceeded.

While the first stage only requires the nominal generation and load to be balanced, the second stage includes the DC power flow constraints for each node $i \in \mathcal{V}$:

$$\sum_{j\,:\,(i,j)\in\mathcal{E}} \beta_{ij}\left[\theta_i(\omega) - \theta_j(\omega)\right] = p_i(\omega) - d_i - \tilde{d}_i(\omega), \tag{8}$$

where $\theta_i(\omega)$ denotes the phase angle at bus $i \in \mathcal{V}$ and β_{ij} ($= \beta_{ji}$) denotes the susceptance in the line (i, j). Summing Eqn. (8) yields the following total power balance constraint:

$$\sum_{i\in\mathcal{G}} p_i(\omega) = \sum_{j\in\mathcal{D}} \left(d_j + \tilde{d}_j(\omega)\right). \tag{9}$$

For any given values of the generator levels $p_i(\omega)$, there is a one-dimensional affine space of solutions $\theta_i(\omega)$ to Eqn. (8). We assume without loss of generality that the first node is chosen as the reference bus with $\theta_1(\omega) \equiv 0$, which, along with Eqn. (8), implies that there is a unique solution to the phase angles $\theta_i(\omega)$, $i \in \mathcal{V}$, e.g., see Lemma 1.1 of [1]. Line flows $[\beta_{ij}(\theta_i(\omega) - \theta_j(\omega))]$ are encouraged to obey line limits by using penalty terms in the objective function.

D. Solution to the second-stage problem

We now characterize conditions under which the system of equations (6), (7), and (9) has a (unique) solution for the second-stage variables $\left(p_i^{\mathrm{T}}(\omega), p_i(\omega), s(\omega)\right)$ given fixed values for the first-stage variables (p^0, r^+, r^-, α).

Theorem 1. *For each value of the first-stage variables satisfying Eqns.* (1) *to* (4)*, the system of equations* (6)*,* (7)*, and* (9) *is feasible exactly when* $\Sigma_d(\omega) \in D_F(p^0, \alpha, \omega)$*, where*

$$\begin{aligned} D_F(p^0,\alpha,\omega) := \Big[&\sum_{i\in\mathcal{R}es(\alpha)} p_i^{\min}(\omega) + \sum_{j\notin\mathcal{R}es(\alpha)} \bar{p}_j^0(\omega) - \sum_{k\in\mathcal{D}} d_k, \\ &\sum_{i\in\mathcal{R}es(\alpha)} p_i^{\max}(\omega) + \sum_{j\notin\mathcal{R}es(\alpha)} \bar{p}_j^0(\omega) - \sum_{k\in\mathcal{D}} d_k \Big]. \end{aligned}$$

Here, $j \notin \mathcal{R}es(\alpha)$ *is shorthand for* $j \in \mathcal{G}\backslash\mathcal{R}es(\alpha)$*, and*

$$\bar{p}_j^0(\omega) = \text{median}\left(p_j^0, p_j^{\min}(\omega), p_j^{\max}(\omega)\right), \quad \forall j \notin \mathcal{R}es(\alpha).$$

Furthermore, the solution for the $p_i(\omega)$ *variables is always unique, whereas the solution for* $\left(p_i^{\mathrm{T}}(\omega), s(\omega)\right)$ *is unique iff* $\Sigma_d(\omega) \in \text{int}(D_F(p^0, \alpha, \omega))$.

The proof for Theorem 1 can be found in Appendix A. We henceforth assume that $\Sigma_d(\omega) \in \text{int}(D_F(p^0, \alpha, \omega))$ for a.e. realization of ω for each value of p^0 and α satisfying Eqns. (1) to (4). Theorem 1 and its proof then implies that given first-stage decisions p^0 and α and a realization of the random variables ω, computing *the* recourse solution reduces to a one-dimensional search for the slack reserves $s(\omega)$.

E. Two-stage stochastic programming model

We propose the following two-stage stochastic DC-OPF model with reserve saturation:

$$\begin{aligned} \min_{\substack{p^0, r^+, \\ r^-, \alpha}} \; &\sum_{i\in\mathcal{G}} \left[f_{1,i}(p_i^0) + f_{2,i}(r_i^+) + f_{3,i}(r_i^-) \right] + Q(p^0, r^+, r^-, \alpha) \\ \text{s.t. } &\text{Constraints (1) to (4)}, \end{aligned} \tag{P}$$

where $Q(p^0, r^+, r^-, \alpha) = \mathbb{E}_\omega\big[q(p^0, r^+, r^-, \alpha, \omega)\big]$ denotes the expected second-stage costs with $q(p^0, r^+, r^-, \alpha, \omega) :=$

$$\begin{aligned} \min_{\substack{p(\omega), p^{\mathrm{T}}(\omega), \\ s(\omega), \theta(\omega)}} &\sum_{i\in\mathcal{G}} \Big[q_{1,i}\left(p_i(\omega) - p_i^0\right) + q_{2,i}\left(p_i(\omega) - (p_i^0 + r_i^+)\right) + \\ &\qquad q_{3,i}\left(p_i(\omega) - (p_i^0 - r_i^-)\right) \Big] + \\ &\sum_{(i,j)\in\mathcal{E}} q_{4,ij}\left(\beta_{ij}\left[\theta_i(\omega) - \theta_j(\omega)\right]\right) \\ \text{s.t.} \quad &\text{Constraints (6) to (8).} \end{aligned} \tag{R}$$

The functions f_1, f_2, and f_3 in the first-stage objective quantify the cost of nominal power generation and the cost of up- and down-reserve capacities, respectively. In the second-stage problem, the term involving the function q_1 quantifies the cost of deviating from the generation level decided in the first-stage, representing e.g., a mileage payment to generators, whereas the terms involving the functions q_2 and q_3 correspond to the penalties for using up- and down-reserves beyond the scheduled reserve limits. Finally, the function q_4 penalizes 'large line flows', with the penalty coefficient chosen to trade-off between the cost of power generation and the line flow violation probability. For simplicity, we use a linear weighting approach for balancing the expected generation cost and the expected violation cost. Alternatively, a constraint on the expected violation penalty could be imposed. We assume that functions f_1 to f_3 and q_1 to q_4 are continuously differentiable with Lipschitz continuous gradients. For a practical example of how the functions f_1, f_2, f_3, q_1, q_2, q_3, and q_4 can be defined, we refer to Section V.

III. Solution approach

Modeling reserve saturation introduces bilinear terms (in the expression for the target generation levels (6)) and nonsmooth nonconvex functions (in the saturation policy (7)). Thus, Problem (P) is a nonsmooth nonconvex two-stage stochastic program, which is in general a challenging problem class to solve even to local optimality. Theorem 1, however, indicates that the recourse problem can be solved efficiently given any candidate first-stage decision, as long as we have an efficient approach to compute the unique recourse solution for a given first-stage decision. In this section, we further show that by replacing the saturation Eqn. (7) with a suitable smooth approximation, partial derivatives of *the* recourse solution with respect to the first-stage decisions can be computed by solving a linear system. Therefore, Theorem 1 suggests a sampling-based decomposition approach (i.e., an approach that works in the space of the first-stage variables) for solving an approximation of Problem (P) to obtain stationary solutions. The remainder of this section proposes a smooth approximation of Problem (P) and a stochastic approximation-based [9, 10] decomposition approach for solving it. We refer the reader to [11, 12] for an overview of stochastic approximation methods

A. Smooth approximation of Problem (P)

We propose a smooth approximation of Problem (P) to obtain a formulation in which all functions are continuously differentiable. The nonsmooth saturation function in (7) is approximated by the continuously differentiable function

$$p_i(\omega) = g_{\tau_{sat}}(p_i^{\mathrm{T}}(\omega); p_i^{\min}(\omega), p_i^{\max}(\omega)), \tag{10}$$

where $g_\tau(x; x^{\mathrm{L}}, x^{\mathrm{U}}) :=$

$$\begin{cases} x^{\mathrm{L}}, & \text{if } x < x^{\mathrm{L}} - \tau \\ x^{\mathrm{L}} + \left(x - (x^{\mathrm{L}} - \tau)\right)^2 / 4\tau, & \text{if } x^{\mathrm{L}} - \tau \le x \le x^{\mathrm{L}} + \tau \\ x, & \text{if } x^{\mathrm{L}} + \tau < x < x^{\mathrm{U}} - \tau \\ x^{\mathrm{U}} - \left(x - (x^{\mathrm{U}} + \tau)\right)^2 / 4\tau, & \text{if } x^{\mathrm{U}} - \tau \le x \le x^{\mathrm{U}} + \tau \\ x^{\mathrm{U}}, & \text{if } x > x^{\mathrm{U}} + \tau \end{cases},$$

and $\tau_{sat} > 0$ is a parameter that controls the approximation quality. We call the approximation of Problem (P) resulting from this modification '*the smooth approximation*'. Smaller values of τ_{sat} yield more accurate, but numerically worse-scaled formulations[1]. Although we use a smooth approximation of the piece-wise linear reserve activation function, the actual generation level $p_i(\omega)$ remains feasible as it never exceeds the lower and upper generation bounds.

B. Solving the recourse problem of the smooth approximation

The analysis of Theorem 1 carries over to the smooth approximation because the function g_τ is monotonically non-decreasing. Therefore, given values of the first-stage variables and a realization of ω, the *unique* recourse solution can be computed by solving the one-dimensional equation for the slack reserves $s(\omega)$ that results from substituting Eqns. (6) and (10) into Eqn. (9). We solve this equation by bisection.

Denote the (target) power generation levels obtained from Eqns. (6) and (10) with $s(\omega) := 0$ by $\hat{p}_i^{\mathrm{T}}(\omega) := p_i^0 + \alpha_i \Sigma_d(\omega)$ and

$$\hat{p}_i(\omega) := g_{\tau_{sat}}(\hat{p}_i^{\mathrm{T}}(\omega); p_i^{\min}(\omega), p_i^{\max}(\omega)).$$

Let the residual power imbalance of Eqn. (9) at these generation levels be denoted by

$$\delta d(\omega) := \sum_{i \in \mathcal{G}} \hat{p}_i(\omega) - \sum_{j \in \mathcal{D}} \left(d_j + \tilde{d}_j(\omega) \right).$$

If $\delta d(\omega)$ is negative, we need to increase generation by providing up-reserves, whereas if $\delta d(\omega)$ is positive, we need to decrease generation by providing down-reserves to balance the overall load for the smooth approximation.

To determine the lower and upper bounds (s^{L}, s^{U}) for the bisection procedure, we consider two cases. If $\delta d(\omega) < 0$, we use $s^{\mathrm{L}}(\omega) = -\delta d(\omega)$ and

$$s^{\mathrm{U}}(\omega) = \max_{i \in \mathcal{R}es(\alpha)} \left\{ \alpha_i^{-1}(p_i^{\max}(\omega) + \tau_{sat} - p_i^0) - \Sigma_d(\omega) \right\},$$

as lower and upper bounds (s^{L}, s^{U}), whereas if $\delta d(\omega) > 0$, we use $s^{\mathrm{U}}(\omega) = -\delta d(\omega)$ and

$$s^{\mathrm{L}}(\omega) = \min_{i \in \mathcal{R}es(\alpha)} \left\{ \alpha_i^{-1}(p_i^{\min}(\omega) - \tau_{sat} - p_i^0) - \Sigma_d(\omega) \right\}.$$

C. Computing stochastic gradients for the approximation

We describe how stochastic gradients of the objective function of Problem (P) are estimated given values of the first-stage variables and a realization of ω. Given partial derivatives of the (unique) recourse solution with respect to the first-stage decisions, we can compute stochastic gradients of the objective function of the approximation using the chain rule under mild conditions (see Theorem 7.44 of [13]).

The partial derivatives of the recourse solution with respect to the reserves r^+ and r^- are identically zero. Partial derivatives of the solution to the generation levels $p_i(\omega)$ with respect to the variables p^0 and α are computed by differentiating Eqns. (6), (9), and (10) and solving the resulting

[1] In the case study, we set $\tau_{sat} = 10^{-4}(p_i^{\max}(\omega) - p_i^{\min}(\omega))$.

linear system of sensitivities. Partial derivatives of the phase angle solution $\theta_i(\omega)$ with respect to p^0 and α are computed by differentiating Eqns. (8) and solving the resulting linear system. We summarize these relationships in Appendix B.

D. Solving the smooth approximation using PSG

We use the projected stochastic gradient (PSG) method of [9, 10] to obtain stationary solutions to the smooth approximation[2]. We argue below that the assumptions of [10] hold. Assumption (A2) of [10] holds since we assume that the conditions of Theorem 1 hold. Furthermore, the smooth saturation function g_τ is continuously differentiable with Lipschitz continuous gradient. The sensitivities of the recourse solutions with respect to the first-stage variables are also Lipschitz continuous. Hence, the objective function of our smooth approximation is continuously differentiable with Lipschitz continuous gradient. Because the first-stage feasible region is compact, assumption $\overline{(A3)}$ of [10] also holds and the PSG method is guaranteed to converge to stationary solutions.

Algorithm 1 presents a basic version of our PSG workflow. In this algorithm, we use $x := (p^0, r^+, r^-, \alpha)$ as the set of all first-stage variables, and define X to be the set of x that satisfy Eqns. (1) to (4). The operator $\mathrm{Proj}_X(y)$ returns the point in X that has smallest Euclidean distance to y. For simplicity the algorithm is described with a fixed step length γ, but we use a variation of AdaGrad [14] for determining step lengths.

Algorithm 1 PSG algorithm for solving the smooth approx.

```
Input: Initial guess x_1 ∈ X, number of iterations T ∈ ℕ,
  mini-batch size K ∈ ℕ, and step length γ > 0.
for t = 1, ..., T do
    for k = 1, ..., K do
        Let ω_k be a random observation of ω.
        Solve Eqns. (6), (8), and (10) to obtain s(ω_k) and
          p^T(ω_k) for the given x_t and ω_k.
        Solve Eqns. (11)-(14) with ω_k, s(ω_k), p^T(ω_k) and
          x_t and use the chain rule to get a stochastic
          gradient ĝ_k of the objective of (P) at iterate x_t.
    end for
    Let x_{t+1} = Proj_X( x_t − γ (1/K) Σ_{k=1}^K ĝ_k ).
    Estimate the objective of Problem (P) using an inde-
      pendent sample of ω, and check termination criteria.
end for
Output: Iterate with the smallest estimated objective.
```

IV. Alternative Models

We compare the solution of the smooth approximation (SA) with the solutions from two alternative models that determine candidate first-stage decisions using the affine policy in Eqn. (5) instead of the saturation model in Eqns. (6) and (10).

[2] An alternative is to use sample average approximation (SAA) to solve the smooth approximation, which can also exploit its decomposable structure

1) Conservative Affine Policy (CAP) Model: The first model we compare against is inspired by [1, 5, 7]. This model enforces individual generator limits using chance constraints with maximum violation allowances $\varepsilon_{gen} \ll 1$, thereby avoiding the need to consider saturation effects for the affine policy, while still using penalty terms to limit line violations.

2) Generator Penalty (GP) Model: The second alternative we consider does not directly include a constraint on the violation probability, but rather penalizes the expected violation of the generator limits by the generation levels determined by the affine policy (cf. [2]) using the terms $\gamma_{gen} \max\{0, p_i(\omega) - p_i^{\max}(\omega), p_i^{\min}(\omega) - p_i(\omega)\}^2$, $i \in \mathcal{G}$, in the recourse objective for a penalty coefficient $\gamma_{gen} > 0$. In our computational experiments, we investigate whether it is possible to choose γ_{gen} such that the GP model yields good solutions to the true Problem (P).

Both of these alternative models are solved using sample average approximation [13] with a nonlinear programming solver. Appendix C presents these models in greater detail.

V. Computational experiments

A. Modeling and implementation details

For the generators, we set the lower bound to zero $p_i^{\min}(\omega) \equiv 0$, $\forall i \in \mathcal{G}$. For the regular generators, $p^{0,\mathrm{L}} = 0$, $p_i^{0,\mathrm{U}} = p_i^{\max}$, $\forall i \in \mathcal{R}$, while wind generators have $p_i^{0,\mathrm{U}} = \mathbb{E}\left[p_i^{\max}(\omega)\right] + 5\sigma\left[p_i^{\max}(\omega)\right]$, $\forall i \in \mathcal{W}$. For the reserve bounds, we set $r^{+,\max} = r^{-,\max} = p^{0,\mathrm{U}}$. We assume that all generators are required to have a positive participation factor such that $\mathcal{G}^{\mathrm{res}} = \mathcal{G}$, with $\varepsilon = \min\left\{0.001, \frac{0.01}{|\mathcal{G}|}\right\}$, and we choose $\tau_{sat} = 10^{-4}(p_i^{\max}(\omega) - p_i^{\min}(\omega))$, $i \in \mathcal{G}$. We assume that the generation limits are wide enough for relatively complete recourse to hold.

The cost functions for all regular generators $\forall i \in \mathcal{R}$ are specified as $f_{1,i}(z) = c_i z$, $f_{2,i}(z) = f_{3,i}(z) = c_i c_{res} z$, where c_i represents the generation cost and $c_{res} = 1.5$ is a reserve cost factor. For the wind generators, we assume that the marginal cost is zero, and use $f_{1,i}(z) \equiv 0$ for the generation cost functions. For the reserve cost functions, we set $f_{2,i}(z) = f_{3,i}(z) = (\min_{k \in \mathcal{R}} c_k)\, c_{wind} c_{res} z$, $\forall i \in \mathcal{W}$, where $c_{wind} = 0.1$ is the relative cost factor for wind reserves. The penalty functions for the second stage are specified as $q_{1,i}(z) \equiv 0$, i.e. the generators are allowed to deviate from the first-stage generation without a penalty. For the penalties representing the exceedance of the scheduled reserve capacities, we set $q_{2,i}(z) = \gamma_{res} f_{2,i}(g^+_{\tau_{pos}}(z))$, and $q_{3,i}(z) = \gamma_{res} f_{3,i}(-g^-_{\tau_{pos}}(z))$, $\forall i \in \mathcal{G}$, where $\gamma_{res} = 10$ is the penalty for demanding reserves beyond the scheduled capacity. The value $\tau_{pos} = 10^{-4}$ is a smoothing parameter, and $g^+_{\tau_{pos}}$ is the smooth approximation to the $(\cdot)_+$ function defined by:

$$g^+_{\tau_{pos}}(z) := \tau_{pos} \log\Big(1 + \exp\big(\frac{z}{\tau_{pos}}\big)\Big),$$

and $g^-_{\tau_{pos}}(z) := -g^+_{\tau_{pos}}(-z)$ is the smooth approximation to $(z)_-$. Finally, the line flow penalty is defined as

$$q_{4,ij}(z) = \gamma_{line} \max\{0, |z| - \delta_{ij} f_{ij}^{\max}\}^2,$$

TABLE I: Comparison of the lowest cost solutions with joint line violation probability $\leq 0.5\%$ for the 6-bus model `case1`.

Model	Expected total cost	First-stage cost Gen.	Res.	Gen. viol.	Wind util. %
SA	3049	2713	115	-	70
CAP 10^{-3}	3505	3188	226	<0.002	45.2
CAP 10^{-2}	3186	2991	126	<0.01	55.9
GP	3043	2689	110	0.069	71.4

TABLE II: Comparison of the lowest cost solutions with joint line violation probability $\leq 0.5\%$ for the 6-bus model `case2`.

Model	Expected total cost	First-stage cost Gen.	Res.	Gen. viol.	Wind util. %
SA	3546	1514	1675	-	100
CAP 10^{-3}	4446	2797	1127	<0.001	69.1
CAP 10^{-2}	4268	2614	1122	<0.01	76.7
GP	3904	2133	1186	0.567	96.8

where $f_{ij}^{\max}$ is the $(i,j)^{\text{th}}$ line flow limit, $\delta_{ij} \equiv 0.95$, and $\gamma_{line} > 0$ is the line flow penalty coefficient that is varied.

We solve SAAs of the comparison models CAP and GP with 500 scenarios (which can be expressed as convex quadratic programs) to determine candidate first-stage solutions. We use a solution from the GP model with $\gamma_{gen} = 20$ as the initial guess x_1 for our smooth approximation model. The quality of the solutions obtained using all approaches are evaluated on the true model (P) (i.e., including reserve saturation) using a common independent Monte Carlo sample of size 10^5.

The code and data of the test instances are available at https://github.com/rohitkannan/DCOPF-reserve-saturation. Our codes are written in Julia 0.6.2 [15], and use Gurobi 7.5.2 [16] to solve convex programs through the JuMP 0.18.2 interface [17]. We use IPOPT 3.12.8 [18] in situations where Gurobi encounters numerical difficulties. All computational tests were conducted on a Surface Book 2 laptop running Windows 10 Pro with a 1.90 GHz four core Intel i7 CPU, 16 GB of RAM.

B. Case Study I: 6-bus system

Our 6-bus example (with $|\mathcal{G}| = 3$) is based on http://motor.ece.iit.edu/data/6bus_Data_ES.pdf. We recast 'generator G2' as a wind generator, and consider normally distributed loads and wind generator capacities with average wind output equal to half the average load. We consider three cases:
`case1`: wind generators *can* provide reserves,
`case2`: wind generators *do not* provide reserves (but are allowed to spill wind without cost, with $\alpha_{wind} = 0.1\varepsilon$), and
`case3`: wind generators are non-dispatchable (i.e., they act like negative loads).
We assume that the standard deviation of the wind output and the loads are 10% of the average for the first two cases, but only 5% of the average for the third case to ensure relatively complete recourse. For these three cases, we compare the solution obtained with the smooth approximation, as well as the solutions from the CAP and GP models.

1) Pareto plots for the three algorithms: To provide a complete picture of the performance of solutions that can be obtained from each algorithm, we present Pareto plots to display the quality of solutions obtained across different parameter values. To generate the Pareto plot, we do a parameter sweep for the tuning parameters of each model. The line violation penalty parameters are changed between $\gamma_{line} = \text{logspace}(1,5,17)$ for our smooth approximation (except for `case2`, where we use 21 values of γ_{line} between 10^1 and 10^5), and between $\gamma_{line} = \text{logspace}(1,5,9)$ for the GP and CAP models. For the GP model we also do a parameter sweep for the generator violation penalties with $\gamma_{gen} = \text{logspace}(0,5,16)$. For the CAP model, we consider two different violation probabilities for generator chance constraints, viz., $\varepsilon_{gen} = [10^{-3}, 10^{-2}]$. Since the solutions depend on the samples and are therefore random, we create five replications for each parameter combination. For each solution, we calculate the expected cost of power generation (including cost of reserves and reserve penalties) and the joint probability that *any* line flow limit is violated by evaluating the system behavior based on the true policy (which includes reserve saturation) on an independent sample. Note that we do not include any assessment of the generator violation probability since this probability is zero in the true model.

Fig. 1 shows the Pareto plots for the three different algorithms and the three different cases, with expected generation cost plotted against the expected joint violation probability for the line flows. We plot solutions for the smooth approximation (blue squares), the GP model (red dots) and the CAP model with two different values for ε_{gen} (black circles and crosses).

We observe from the Pareto plots that our smooth approximation always provides nearly non-dominated solutions for all three cases. The solutions obtained with the GP model are not concentrated along the Pareto front, as generation violation penalties that are either too large or too small lead to larger-than-necessary cost. Solutions from the CAP model provide a different Pareto front with a larger cost than the smooth approximations, though the solutions coincide with the smooth approximation for smaller values of the violation probability.

Beyond these general behaviors, the models compare differently between different cases. In `case1`, careful parameter tuning allows the GP model to find points along the Pareto curve. In `case2`, there is a gap between the lowest cost solutions obtained with the GP model and the Pareto curve found with the smooth approximation. The smooth approximation is able to find lower cost solutions when the probability of line violations is not set too low. The lack of data points in the Pareto curve for the smooth approximation model in `case2` between line violation probabilities of 6×10^{-4} and 2×10^{-3} is a result of using a linear weighting approach for balancing the expected generation and violation costs for the nonconvex Problem (P), see Chapter 3 of [19]. In `case3`, the solutions of all three algorithms cluster along the Pareto front.

2) Detailed comparison of differences: To analyze the cause of these differences, we investigate some of the solutions in more detail. For each algorithm and cases 1 and 2, we

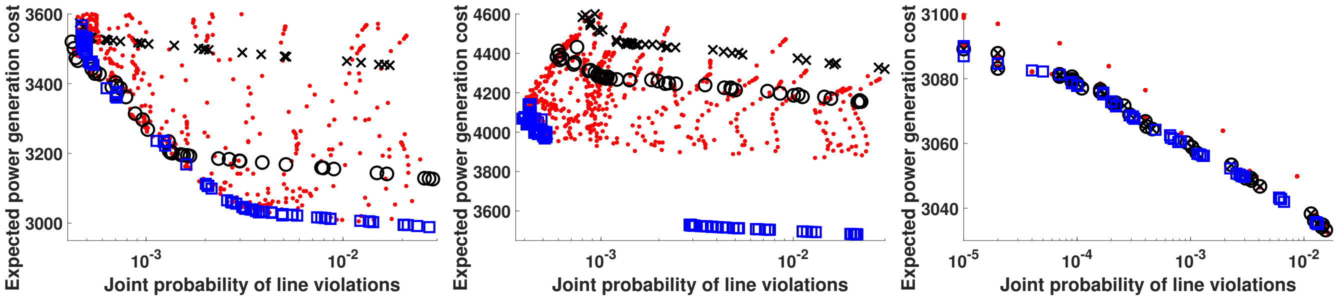

Fig. 1: (Left to right) Pareto plots for the 6-bus system generated using five replicates for `case1`, `case2`, and `case3`.
Blue squares: Smooth approximation **Red dots:** GP solution **Black crosses:** CAP 10^{-3} **Black circles:** CAP 10^{-2}

list the results for the lowest cost solution with a joint line violation probability $\leq$ 0.5% in Tables I and II. We list the total expected cost, the first-stage scheduled generation and reserve capacity costs, the out-of-sample joint violation probabilities of the lines, and the joint violation probability of the generators if we would have considered an affine control policy (only calculated for the CAP and GP models). We also list the expected utilization of wind energy, as a percentage of total available wind power.

In `case1` we observe that the solutions from GP and the smooth approximation have very similar total costs and utilization of the wind energy. The smooth approximation invests more in both generation and reserve capacity in the first stage, which is balanced by paying lower penalties in the second stage. Interestingly, the optimal choice of tuning parameters for the GP solution leads to a relatively high violation probability for the generators at 6.9%. If we enforce a lower violation probability, as is done in the CAP model and has typically been done in literature (see e.g. [1]), the total expected cost increases significantly and the utilization of wind energy drops.

For `case2`, where wind generators are not allowed to provide reserves, we observe that the total expected cost is significantly lower for the smooth approximation than for both the GP and CAP models. The smooth approximation has a lower first-stage generation cost (indicating high dispatch levels for the cheap wind power), but invests more in procuring reserves (that can make up for overestimates in the wind generation). This leads to full utilization of the available wind power. In comparison, the GP model schedules less wind power in the first stage, leading to a higher cost and lower wind utilization. Interestingly, the affine policy in the best GP solution violates the generator limits with more than 50% probability. This also explains why the CAP solutions, where the generation violation probability is explicitly limited, leads to much higher total expected cost (and much lower wind power utilization) than the other two models.

Finally, we do not compare the solutions in `case3` as they are very similar for all three algorithms. In this case, the solutions balance the cost of scheduling more power from the less expensive generator with paying penalties for violating the line constraints. The reserve activation is happening at the more expensive generator, which is far away from saturation. The generators never violate their limits even with an affine control policy and the models are therefore the same.

C. Case Study II: 118-bus system

In the second part of our case study, we consider the more realistic test case based on the IEEE 118-bus system from http://motor.ece.iit.edu/data/JEAS_IEEE118.doc with modifications suggested in [2], including the addition of 25 wind generators to the 54 regular generators, increasing the average demand by 50%, and reducing the line flow limits by 25%. We again consider normally distributed loads and wind generator capacities, and consider five different levels of wind penetration: average wind output = 25%, 50%, 75%, 100%, or 125% of the average system load. To obtain appropriate parameter values for the algorithms, we run a similar parameter sweep as for the 6-bus test case with modified $\gamma_{gen} = \text{logspace}(0, 4, 9)$ and $\varepsilon_{gen} = [10^{-5}, 10^{-4}]$. It takes roughly 1.5 minutes on average to solve the GP and CAP models and roughly 7 minutes on average to solve the smooth approximation model for one instance. We then pick the solution with lowest expected total cost and joint line flow violation probability $\leq 0.5\%$ for each algorithm and each wind level penetration. The expected total cost, expected wind utilization and expected fraction of total system load served by wind power is calculated using a Monte Carlo simulation with 10^5 samples. The results are plotted in Fig. 2, and are also described in Tables III to VII.

As can be seen from Fig. 2, the smooth approximation outperforms the other methods in all aspects, at all wind penetration levels. The cost is lower, the wind utilization is higher and more load is served by the wind generators. The solution obtained by a properly tuned affine GP model achieves results that are quite similar to the smooth approximation, although the relative cost difference is quite high. At 125% wind penetration levels, the expected cost is 30% higher for the GP solution. Both of the CAP methods perform significantly worse than the smooth approximation.

VI. Conclusion and Future Work

We propose a stochastic DC optimal power flow model with reserve saturation. Specifically, our model assumes that generators follow an affine control policy until they reach a generation limit, at which point they operate at that limit. The model is a two-stage stochastic program with nonconvex,

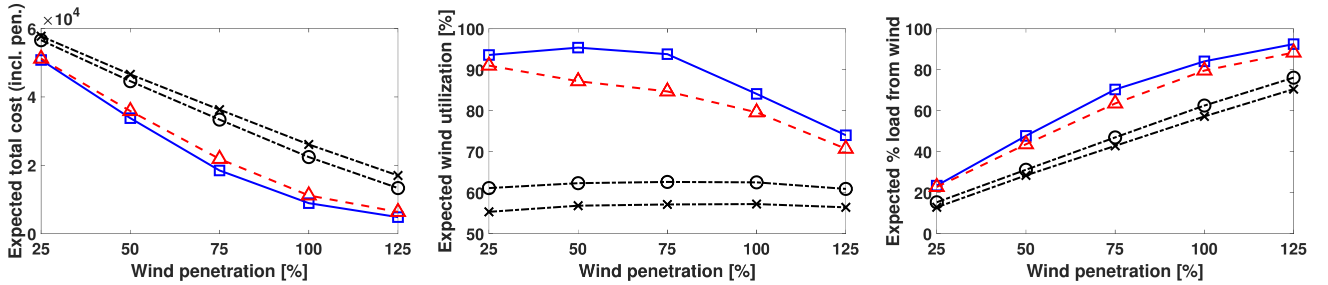

Fig. 2: Summary of solutions for the different models in the 118-bus case with varying wind penetration levels.
Blue squares: Smooth approximation **Red triangles:** GP solution **Black crosses:** CAP 10^{-5} **Black circles:** CAP 10^{-4}

TABLE III: Comparison of the lowest cost solutions with joint line viol. prob. $\leq 0.5\%$ for the 118-bus model with 25% wind penetration.

Model	Expected total cost	First-stage cost Gen.	Res.	Gen. viol.	Wind util. %
SA	50803	50000	364	-	93.6
CAP 10^{-5}	57760	57487	164	<0.01	55.3
CAP 10^{-4}	56587	56304	164	<0.01	61.1
GP	51199	50470	161	0.997	91.0

TABLE IV: Comparison of the lowest cost solutions with joint line viol. prob. $\leq 0.5\%$ for the 118-bus model with 50% wind penetration.

Model	Expected total cost	First-stage cost Gen.	Res.	Gen. viol.	Wind util. %
SA	33797	32567	622	-	95.4
CAP 10^{-5}	46594	46342	164	<0.01	56.8
CAP 10^{-4}	44610	44358	164	<0.01	62.3
GP	35872	35444	180	0.96	87.2

TABLE V: Comparison of the lowest cost solutions with joint line viol. prob. $\leq 0.5\%$ for the 118-bus model with 75% wind penetration.

Model	Expected total cost	First-stage cost Gen.	Res.	Gen. viol.	Wind util. %
SA	18503	16947	817	-	93.8
CAP 10^{-5}	36288	36029	164	<0.01	57.1
CAP 10^{-4}	33367	33103	164	<0.02	62.6
GP	21882	21432	177	0.89	84.7

TABLE VI: Comparison of the lowest cost solutions with joint line viol. prob. $\leq 0.5\%$ for the 118-bus model with 100% wind pen.

Model	Expected total cost	First-stage cost Gen.	Res.	Gen. viol.	Wind util. %
SA	8946	7865	458	-	84.1
CAP 10^{-5}	26104	25840	164	<0.02	57.2
CAP 10^{-4}	22422	22158	164	<0.02	62.5
GP	11205	10712	161	0.82	79.6

TABLE VII: Comparison of the lowest cost solutions with joint line viol. prob. $\leq 0.5\%$ for the 118-bus model with 125% wind pen.

Model	Expected total cost	First-stage cost Gen.	Res.	Gen. viol.	Wind util. %
SA	4863	3886	420	-	74.0
CAP 10^{-5}	17028	16740	164	<0.01	56.4
CAP 10^{-4}	13354	13053	164	<0.01	60.9
GP	6378	5840	161	0.83	70.7

nonsmooth second stage constraints, and we propose a stochastic approximation method to solve a smooth approximation. We empirically observe that our model yields solutions that outperform those obtained from a model that constrains the affine control policy to rarely violate generation limits. On the other hand, using a model that penalizes expected violation of generator limits can sometimes yield competitive solutions with a well-tuned choice of the penalty parameter.

Extensions to our model that would be interesting to investigate in future work include constraining the probability or expected violation of line limits rather than penalizing violation of line limits in the objective, and using an AC power flow model in place of the DC model.

References

[1] D. Bienstock, M. Chertkov, and S. Harnett. Chance-constrained optimal power flow: Risk-aware network control under uncertainty. *SIAM Review*, 56(3):461–495, 2014.

[2] L. Roald, G. Andersson, S. Misra, M. Chertkov, and S. Backhaus. Optimal power flow with wind power control and limited expected risk of overloads. In *Power Systems Comp. Conf. (PSCC)*, 2016.

[3] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson. A Probabilistic Framework for Reserve Scheduling and N-1 Security Assessment of Systems With High Wind Power Penetration. *IEEE Transactions on Power Systems*, 28(4):3885–3896, 2013.

[4] L. Roald, S. Misra, M. Chertkov, and G. Andersson. Optimal power flow with weighted chance constraints and general policies for generation control. In *Proc. IEEE Conf. Decis. Control*, pages 6927–6933, 2015.

[5] Y. Zhang, S. Shen, and J. Mathieu. Distributionally robust chance-constrained OPF with uncertain renewables and uncertain reserves provided by loads. *IEEE Trans. Power Syst.*, 32(2):1378–1388, 2016.

[6] 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.

[7] M. Vrakopoulou and I.A. Hiskens. Optimal control policies for reserve deployment with probabilistic performance guarantees. In *IEEE Conf. on Decision and Control (CDC)*, pages 4470–4475. IEEE, 2017.

[8] A J Wood and B F Wollenberg. *Power generation, operation, and control*. John Wiley & Sons, 2012.

[9] S. Ghadimi, G. Lan, and H. Zhang. Mini-batch stochastic approximation methods for nonconvex stochastic composite optimization. *Mathematical Programming*, 155(1-2):267–305, 2016.

[10] D. Davis and D. Drusvyatskiy. Stochastic subgradient method converges at the rate $O\left(k^{-1/4}\right)$ on weakly convex functions. *arXiv preprint arXiv:1802.02988*, 2018.

[11] Léon Bottou, Frank E Curtis, and Jorge Nocedal. Optimization methods for large-scale machine learning. *Siam Review*, 60(2):223–311, 2018.

[12] G. Lan. *Lectures on Optimization Methods for Machine Learning*. 2019. URL http://pwp.gatech.edu/guanghui-lan/wp-content/uploads/sites/330/2019/08/LectureOPTML.pdf.

[13] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on stochastic programming: modeling and theory*. SIAM, 2009.

[14] J. Duchi, E. Hazan, and Y. Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(Jul):2121–2159, 2011.
[15] J. Bezanson, A. Edelman, S. Karpinski, and V. Shah. Julia: a fresh approach to numerical computing. *SIAM Review*, 59(1):65–98, 2017.
[16] Gurobi Optimizer Reference Manual. URL http://www.gurobi.com.
[17] I. Dunning, J. Huchette, and M. Lubin. JuMP: A modeling language for mathematical optimization. *SIAM Review*, 59(2):295–320, 2017.
[18] A. Wächter and L. T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical programming*, 106(1):25–57, 2006.
[19] M. Ehrgott. *Multicriteria optimization*. Springer, 2006.

Appendix A
Proof of Theorem 1

If $\Sigma_d(\omega) \notin D_F(p^0, \alpha, \omega)$, then Eqns. (6), (7), and (9) are inconsistent since Eqns. (6) and (7) together imply

$$\begin{aligned}&\sum_{i\in\mathcal{R}es(\alpha)} p_i^{\min}(\omega) + \sum_{j\notin\mathcal{R}es(\alpha)} \bar{p}_j^0(\omega) - \sum_{k\in\mathcal{D}} d_k \\ \leq &\sum_{i\in\mathcal{G}} p_i(\omega) - \sum_{k\in\mathcal{D}} d_k \\ \leq &\sum_{i\in\mathcal{R}es(\alpha)} p_i^{\max}(\omega) + \sum_{j\notin\mathcal{R}es(\alpha)} \bar{p}_j^0(\omega) - \sum_{k\in\mathcal{D}} d_k,\end{aligned}$$

which makes the satisfaction of Eqn. (9) impossible.

We now show that the system of equations (6), (7), and (9) has a solution whenever $\Sigma_d(\omega) \in D_F(p^0, \alpha, \omega)$. For a chosen value of the slack reserves $s(\omega)$, denote the value of $p_i(\omega)$ obtained using Eqns. (6) and (7) by $\hat{p}_i(s)$ (we omit dependence on the first-stage variables and ω for simplicity). Note that $\hat{p}_i(s)$ is a monotonically nondecreasing continuous function of s for each $i \in \mathcal{G}$, which implies that $\sum_{i\in\mathcal{G}} \hat{p}_i(s)$ is a monotonically nondecreasing continuous function of s. Furthermore, we have from Eqns. (6) and (7) that generators with a nonzero participation factor will eventually hit their bounds for small/large enough chosen values of s, i.e., there exists $M > 0$ large enough for which $\hat{p}_i(-M) = p_i^{\min}(\omega)$ and $\hat{p}_i(M) = p_i^{\max}(\omega)$, $\forall i \in \mathcal{R}es(\alpha)$, which implies $\sum_{i\in\mathcal{G}} \hat{p}_i(-M) = \sum_{i\in\mathcal{R}es(\alpha)} p_i^{\min}(\omega) + \sum_{j\notin\mathcal{R}es(\alpha)} \bar{p}_j^0(\omega)$ and $\sum_{i\in\mathcal{G}} \hat{p}_i(M) = \sum_{i\in\mathcal{R}es(\alpha)} p_i^{\max}(\omega) + \sum_{j\notin\mathcal{R}es(\alpha)} \bar{p}_j^0(\omega)$. From the intermediate value theorem applied to $\sum_{i\in\mathcal{G}} \hat{p}_i(\cdot)$, there exists $\hat{s} \in [-M, M]$ such that $\sum_{i\in\mathcal{G}} \hat{p}_i(\hat{s}) = \sum_{j\in\mathcal{D}} \left[d_j + \tilde{d}_j(\omega)\right]$ for any $\Sigma_d(\omega) \in D_F(p^0, \alpha, \omega)$. Therefore, the system of equations (6), (7), and (9) has a solution for the variables $\left(p_i^{\mathrm{T}}(\omega), p_i(\omega), s(\omega)\right)$ whenever $\Sigma_d(\omega) \in D_F(p^0, \alpha, \omega)$.

When $\Sigma_d(\omega) \in \mathrm{int}(D_F(p^0, \alpha, \omega))$, we have from Eqn. (9) that there exists a generator that has not hit its generation limits, i.e., $\exists j \in \mathcal{R}es(\alpha)$ such that $p_j^{\min}(\omega) < p_j(\omega) < p_j^{\max}(\omega)$ at a solution to Eqns. (6), (7), and (9). This implies that the sum $\sum_{i\in\mathcal{G}} \hat{p}_i(\cdot)$ is monotonically (strictly) increasing in a neighborhood of $s(\omega)$ around the above solution, which establishes its uniqueness since $\sum_{i\in\mathcal{G}} \hat{p}_i(s)$ is a monotonically nondecreasing continuous function of s. The argument for the 'only if' part is similar. □

Appendix B
Calculation of Partial Derivatives

The linear system below is solved to obtain the partial derivatives of the recourse solution with respect to the first-stage decision variable q, where q is a placeholder for either p_l^0 or α_l, $l \in \mathcal{G}$:

$$\frac{\partial p_i^{\mathrm{T}}}{\partial q}(\omega) = \frac{\partial p_i^0}{\partial q} + (s(\omega) + \Sigma_d(\omega))\frac{\partial \alpha_i}{\partial q} + \alpha_i \frac{\partial s}{\partial q}(\omega), \quad (11)$$

$$\frac{\partial p_i}{\partial q}(\omega) = \frac{\partial g_{\tau_{sat}}}{\partial p_i^{\mathrm{T}}}(p_i^{\mathrm{T}}(\omega); p_i^{\min}(\omega), p_i^{\max}(\omega))\frac{\partial p_i^{\mathrm{T}}}{\partial q}(\omega), \quad (12)$$

$$\sum_{k\in\mathcal{G}} \frac{\partial p_k}{\partial q}(\omega) = 0, \quad \frac{\partial \theta_1}{\partial q}(\omega) = 0 \quad (13)$$

$$\sum_{j\,:\,(i,j)\in\mathcal{E}} \beta_{ij}\left[\frac{\partial \theta_i}{\partial q}(\omega) - \frac{\partial \theta_j}{\partial q}(\omega)\right] = \frac{\partial p_i}{\partial q}(\omega). \quad (14)$$

Appendix C
Details of the Formulations

We explicitly write out the three formulations considered for the computational experiments below (parameter settings are listed in Sec. V).

Smooth approximation (SA):

$$\begin{aligned}\min_{\substack{p^0, r^+,\\ r^-, \alpha}} \ &\sum_{i\in\mathcal{G}} \left[c_i p_i^0 + \bar{c}_i\left(r_i^+ + r_i^-\right)\right] + Q_1(p^0, r^+, r^-, \alpha) \\ \text{s.t.} \ &\text{Constraints (1) to (4)},\end{aligned}$$

where $\bar{c}_i = c_i c_{res}$, for $i \in \mathcal{R}$, and $\bar{c}_i = c_{wind} c_{res}(\min_{k\in\mathcal{R}} c_k)$, for $i \in \mathcal{W}$, $c_i = 0$, $\forall i \in \mathcal{W}$, $Q_1(p^0, r^+, r^-, \alpha) = \mathbb{E}_\omega\big[q_1(p^0, r^+, r^-, \alpha, \omega)\big]$ denotes the expected second-stage costs with $q_1(p^0, r^+, r^-, \alpha, \omega) :=$

$$\begin{aligned}\min_{\substack{p(\omega), p^{\mathrm{T}}(\omega),\\ s(\omega), \theta(\omega)}} \ &\sum_{i\in\mathcal{G}} \gamma_{res}\bar{c}_i g_{\tau_{pos}}^+(p_i(\omega) - p_i^0 - r_i^+) + \\ &\sum_{i\in\mathcal{G}} \gamma_{res}\bar{c}_i g_{\tau_{pos}}^+(p_i^0 - p_i(\omega) - r_i^-) + \\ &\sum_{(i,j)\in\mathcal{E}} \max\left\{0, |\beta_{ij}\left[\theta_i(\omega) - \theta_j(\omega)\right]| - \delta_{ij} f_{ij}^{\max}\right\}^2 \\ \text{s.t.} \ &\text{Constraints (6), (8), and (10).}\end{aligned}$$

Conservative affine policy (CAP) model

$$\begin{aligned}\min_{\substack{p^0, r^+,\\ r^-, \alpha}} \ &\sum_{i\in\mathcal{G}} \left[c_i p_i^0 + \bar{c}_i\left(r_i^+ + r_i^-\right)\right] + Q_2(p^0, r^+, r^-, \alpha) \\ \text{s.t.} \ &\text{Constraints (1) to (4)}, \\ &\mathbb{P}\left\{p_i^0 + \alpha_i \Sigma_d(\omega) \geq p_i^{\min}(\omega)\right\} \geq 1 - \varepsilon_{gen}, \quad \forall i \in \mathcal{G}, \\ &\mathbb{P}\left\{p_i^0 + \alpha_i \Sigma_d(\omega) \leq p_i^{\max}(\omega)\right\} \geq 1 - \varepsilon_{gen}, \quad \forall i \in \mathcal{G},\end{aligned}$$

where $Q_2(p^0, r^+, r^-, \alpha) = \mathbb{E}_\omega\big[q_2(p^0, r^+, r^-, \alpha, \omega)\big]$ denotes the expected second-stage costs with $q_2(p^0, r^+, r^-, \alpha, \omega) :=$

$$\min_{p(\omega),\theta(\omega)} \sum_{i\in\mathcal{G}} \gamma_{res}\bar{c}_i \left(\left(p_i(\omega) - p_i^0 - r_i^+\right)_+ + \left(p_i^0 - p_i(\omega) - r_i^-\right)_+ \right) + \sum_{(i,j)\in\mathcal{E}} \max\left\{0, |\beta_{ij}\left[\theta_i(\omega) - \theta_j(\omega)\right]| - \delta_{ij} f_{ij}^{\max}\right\}^2$$

s.t. Constraints (5) and (8).

Generator penalty (GP) model

$$\min_{\substack{p^0,r^+,\\ r^-,\alpha}} \sum_{i\in\mathcal{G}} \left[c_i p_i^0 + \bar{c}_i \left(r_i^+ + r_i^-\right)\right] + Q_3(p^0, r^+, r^-, \alpha)$$

s.t. Constraints (1) to (4),

where $Q_3(p^0, r^+, r^-, \alpha) = \mathbb{E}_\omega\left[q_3(p^0, r^+, r^-, \alpha, \omega)\right]$ denotes the expected second-stage costs with $q_3(p^0, r^+, r^-, \alpha, \omega) :=$

$$\min_{p(\omega),\theta(\omega)} \sum_{i\in\mathcal{G}} \gamma_{res}\bar{c}_i \left(\left(p_i(\omega) - p_i^0 - r_i^+\right)_+ + \left(p_i^0 - p_i(\omega) - r_i^-\right)_+ \right) + \sum_{(i,j)\in\mathcal{E}} \max\left\{0, |\beta_{ij}\left[\theta_i(\omega) - \theta_j(\omega)\right]| - \delta_{ij} f_{ij}^{\max}\right\}^2 + \sum_{i\in\mathcal{G}} \gamma_{gen} \max\left\{0, p_i(\omega) - p_i^{\max}(\omega), p_i^{\min}(\omega) - p_i(\omega)\right\}^2$$

s.t. Constraints (5) and (8).