Variance-Aware Optimal Power Flow: addressing the trade-off between cost, security and variability

Daniel Bienstock and Apurv Shukla Columbia University, NY, United States
{dano, apurv.shukla}@columbia.edu

***Abstract*—Recent work has studied modifications to DC-OPF computations so as to better account for risk arising as a result of stochastic variation in the output of renewable sources. Typically such modifications rely on mathematical constructs such as chance-constraints that can still yield convex formulations. However, numerical simulations show that the computed policies can translate into power flow patterns with high variance. We introduce a number of convex variants of OPF that trade-off variance and cost minimization, describe practical algorithms for the solution of such problems, and present numerical experiments.**

***Index Terms*—Chance-constraints, stochastic generation, OPF.**

I. Introduction

A number of trends in power engineering practice have resulted in increasing exposure of transmission systems to risk in the form of variability and more broadly unpredictable behavior. Renewable penetration has been an early driver in this direction, to which one should probably add 'smart' loads, the use of distributed photovoltaics used in response to market signals, and other factors.

One impact of such stochasticity is the potential for increased equipment overloads. Prior research has addressed this problem by developing alternative formulations for the Optimal Power Flow (OPF) problem that also include modifications to balancing mechanisms. An example is provided by chance-constrained formulations for OPF (see [1]–[11]), which yield both generator output values but also participation factors, so that the probability that any given line exceeds its operational limit is upper bounded by a small value, under some assumptions on the stochasticity of uncertain sources.

Even though such schemes are *risk-aware* they may nevertheless fail to capture another (and important) facet of risk. In particular, numerical experiments such as those discussed herein, show that dispatch and participation decisions computed by the chance-constrained formulation may yield real-time power flows with high variability; specifically with the feature that some transmission lines have high *variance* of power flow. Such a feature is, likely, undesirable from a real-time operational perspective. See [12], which discusses negative impact of power flow variability on voltage profile and transformer operation. In general, high variance will likely hamper system understanding and control, and pricing policies (see [13], [14], [15] for work on pricing under uncertainty). [16] describes an optimization method for selecting a mix of renewable generation (under Gaussian stochastics) so as to attain a desirable mean-variance trade-off in flows on tie lines of a balancing area. Also see [17].

This paper continues the work initiated in [18]. We describe modifications to the OPF problem which we term *variance-aware OPF*. In these modifications we explicitly trade-off variance in power-flow related quantities to operational cost. We provide computational experiments with several variance-aware schemes, and present an analysis of the variance-cost trade-off that borrows ideas from the classical "Sharpe-ratio" perspective in economics.

In Sections II-III we review prior work and introduce some basic concepts and formulations. Section IV presents a simple example of the impact of variance. Section V presents our main variance-aware formulations, and Section V-D describes numerical experiments with larger systems. Section VII discusses a different convex optimization formulation for addressing the trade-off between OPF cost and solution variance.

II. Notation and basic formulations

In this paper we will focus on the linearized, or DC model for power flows. We will use the following selected nomenclature, with additional terms described later:

$\mathcal{B}$ = set of buses, $n = |\mathcal{B}|$; B = bus susceptance matrix.
$\hat{B} = B$, with the last row and column removed
$\breve{B} = \begin{bmatrix} \hat{B}^{-1} & 0 \\ 0 & 0 \end{bmatrix}$; $\breve{B}_i = i^{th}$ row of $\breve{B}$.
θ_i = phase angle at bus i, $\bar{\theta}_i = \mathbf{E}(\theta_i)$.
$\mathcal{E}$ = set of lines; $m = |\mathcal{E}|$.
$f_{ij}^{\max}$, b_{ij} = power flow limit, and susceptance, for line $ij \in \mathcal{E}$.
f_{ij} = power flow on line ij, $\bar{f}_{ij} = \mathbf{E}(f_{ij}) = b_{ij}(\bar{\theta}_i - \bar{\theta}_j)$
$\pi_{ij} \doteq \breve{B}_i^T - \breve{B}_j^T$ for each line $ij \in \mathcal{E}$
$\mathcal{G}$ = set of generator buses; we assume at most one generator per bus
p_i^{min}, p_i^{max} = minimum and maximum output of $i \in \mathcal{G}$.
$\mathcal{S}$ = set of stochastic injection buses, $\mu_k + \omega_k$ = stochastic injection at bus $k \in \mathcal{B}$,

- μ_k = constant, ω_k = zero-mean random variable.
- $\mu_k = \omega_k = 0$ if $k \notin S$,
- Ω = covariance of ω, an $n \times n$ matrix,
- $\mathcal{W}$ = support of distribution for ω,

$\mathcal{R}$ = set of buses participating in balancing
$\mathcal{A}$ = matrix of participation factors; α_{ij} = (i,j)-entry of $\mathcal{A}$,

- $\mathcal{A}$ is $n \times n$

- $\alpha_{ij} = 0$ if $i \notin \mathcal{R}$ or $j \notin \mathcal{S}$
- $\mathbf{V}(\mathcal{A})$ = vector with entries $b_{ij}^2 \pi_{ij}^T (I - \mathcal{A}) \Omega (I - \mathcal{A}^T) \pi_{ij}$ (variances) for $ij \in \mathcal{E}$

$\mathcal{K}$ = convex set of allowable participation factor matrices
p = vector of generation amounts, extended to all $b \in \mathcal{B}$ (forcing $p_b = 0$ when $b \notin \mathcal{B}$).
$c(p)$ = (convex) cost of generation vector p.
d= (fixed) vector of loads.

The standard DC-OPF formulation, using this notation, is as follows:

$$\min_p c(p) \tag{1a}$$
$$\text{s.t.} \quad B\theta = p - d \tag{1b}$$
$$\forall ij \in \mathcal{E}: \; b_{ij}|\theta_i - \theta_j| \leq f_{ij}^{\max} \tag{1c}$$
$$\forall i \in \mathcal{G}: \; p_i^{min} \leq p_i \leq p_i^{max} \tag{1d}$$

A well-known simplification to constraints (1b)-(1c) relies on a pseudo-inverse for the B matrix (see, for example [1], [3]). For a line ij, et π_{ij}^T denote the row of the pseudo-inverse corresponding to ij. Then the power flow on ij equals $b_{ij}\pi_{ij}^T(p - d)$. Using this equation we reduce (1b)-(1c) to the system

$$\forall ij \in \mathcal{E} : b_{ij}|\pi_{ij}^T(p - d)| \leq f_{ij}^{max} \tag{2}$$

III. Security-constrained formulations

Previous research [1], [2], [3], [4], [5], [7], [9], [19] has focused on alterations to DC-OPF models which take into account uncertain injections so as to adjust generation decisions the use of AGC (Automatic Generation Control).

The methodology proposed in [1], [3], [9], modifies the DC-OPF computation so as to output both a vector $\bar{p}$ of *controllable generation* amounts and (as in [9]) an $n \times n$ matrix $\mathcal{A}$ used to model controllable generator participation in balancing. In this scheme, a vector $\omega \in \mathcal{W}$ of output variation, is balanced by generation changes given by $\mathcal{A}\omega$, with the result that the total generator output at bus i is

$$p_i = p_i(\omega) = \bar{p}_i - [\mathcal{A}\omega]_i = \bar{p}_i - \sum_{j \in \mathcal{B}} \alpha_{ij}\omega_j. \tag{3}$$

For simplicity, when i is not a controllable generator bus, we write $\alpha_{ij} = 0$ for all j. As a result the vector of net injections equals $\bar{p} - d + \mu + \omega - \mathcal{A}\omega$. For the system to be balanced we need the stochastic condition

$$\forall w \in \mathcal{W} : \quad \sum_{i \in \mathcal{B}} (\bar{p} - d + \mu + \omega - \mathcal{A}\omega)_i = 0. \tag{4}$$

Suppose we also require $\sum_{i \in \mathcal{B}}(\bar{p} - d + \mu)_i = 0$, i.e. absent stochastics the system is balanced. Then (4) is the same as

$$\forall \omega \in \mathcal{W} : \quad \omega_i - \sum_{j \in \mathcal{B}} \alpha_{ij}\omega_j = 0, \quad \forall\, i \in \mathcal{B}. \tag{5}$$

This condition will be satisfied if we impose [9]

$$1 = \sum_{j \in \mathcal{B}} \alpha_{ij} \quad \forall\, i \in \mathcal{B}, \tag{6}$$

which is an equation in ω-space (thus removing one degree of freedom in ω). If $\mathcal{W}$ is full-dimensional then (6) is in fact required for (4) [20]. Of course, we may impose additional requirements on $\mathcal{A}$, for example by having bounds on individual α_{ij}. We may also impose a "global" policy [9], namely that for every bus $i \in \mathcal{R}$ [1], $\alpha_{ij} = \alpha_{ik}$ for every $j \neq k$.
Notation. We will denote by $\mathcal{A} \in \mathcal{K}$ a generic set of admissible participation matrices $\mathcal{A}$, typically described by convex constraints.

In summary, where f_{ij} denotes flow on line ij, we have

$$\forall ij \in \mathcal{E} : \quad f_{ij} = b_{ij}\pi_{ij}^T(\bar{p} - d + \mu + \omega - \mathcal{A}\omega), \tag{7a}$$
$$\mathbf{E}(f_{ij}) = b_{ij}\pi_{ij}^T(\bar{p} - d + \mu). \tag{7b}$$

Likewise by construction

$$\mathbf{V}(\mathcal{A})_{ij} \doteq \mathbf{Var}(f_{ij}) = b_{ij}^2 \mathbf{Var}(\pi_{ij}^T(I - \mathcal{A})\omega) = b_{ij}^2 \pi_{ij}^T (I - \mathcal{A}) \Omega (I - \mathcal{A}^T) \pi_{ij} \tag{8}$$

Depending on how randomness is modeled, or whether the matrix $\mathcal{A}$ is fixed or subject to optimization, one obtains a number of variants of the problem (1).

The use of *chance constraints* has emerged as a central idea in the modeling of safe operation. Consider a given line ij. A chance constraint on this line is of the form

$$\mathbf{P}(|f_{ij}| > f_{ij}^{\max}) < \epsilon \tag{9}$$

where $0 < \epsilon < 1$ reflects the planner's tolerance for risk (or perhaps, perception on the impact of risk on the system). A relaxation of this requirement is that

$$\mathbf{P}(f_{ij} > f_{ij}^{\max}) < \epsilon \quad \text{and} \quad \mathbf{P}(f_{ij} < -f_{ij}^{\max}) < \epsilon. \tag{10}$$

When ω is Gaussian, (10) can be summarized as

$$|\mathbf{E}(f_{ij})| + \Phi^{-1}(1 - \epsilon)\mathbf{Std}(f_{ij}) \leq f_{ij}^{\max} \tag{11}$$

where $\Phi^{-1}(1-\epsilon)$ is the ϵ-quantile for a normal distribution and $\mathbf{Std}$ is standard deviation. System (11) is SOCP representable [1], [3].

A. Modifications used in this paper

We substitute (11) with

$$|\mathbf{E}(f_{ij})| + \nu_{ij}\, \mathbf{Std}(f_{ij}) \leq f_{ij}^{\max}. \tag{12}$$

Here, ν_{ij} is a *safety* parameter. As outlined above in the Gaussian case (12) is equivalent to (9) when $\nu_{ij} = \Phi^{-1}(1-\epsilon)$. However, the Gaussian case is not the only one where such an equivalence holds; see the discussion in [18]. In each case one needs an appropriately constructed parameter $\nu_{ij} = \nu_{ij}(\epsilon)$. Additionally (12) can be used to tightly approximate distributionally robust (or "ambiguous") chance constraints. Refer e.g. to [21]. In any case, one could argue that even when the stochastics of ω is poorly understood or nontrivial with the result that it is difficult to fully justify a particular choice for the safety parameters, in general we will still be able to compute Ω, or in the worst case estimate it from data. The safety-parameter approach would still be practicable

[1] In the sequel, a participating bus.

and relevant even though it is equipped with a satisfactory mathematical guarantee. We will term (12) a *safety* constraint.

Using matrix Ω we can provide a different expression for the variance of a line flow f_{ij}. Let us write

$$D \doteq \breve{B}\mathcal{A}, \quad \text{and} \tag{13a}$$
$$\forall\, k \in \mathcal{S}: \; \gamma_{ij,k} \doteq \breve{B}_{ik} - \breve{B}_{jk} - D_{ik} + D_{jk}. \tag{13b}$$

Lemma 1. *For any line ij, the variance of flow on ij under scheme (3) is given by*

$$\mathbf{Var}(f_{ij}) = b_{ij}^2\, \gamma_{ij}\, \Omega\, \gamma_{ij}^T \tag{14}$$

A proof of this result can be obtained by substituting (13) into (8) (also see prior work: [1], [3], [9]). It is also important to note that (14) holds for all probability distributions.

Similar developments apply to generator output. Let generation cost at a bus i be given by $c_i(p) \doteq c_{i0}p^2 + c_{i1}p + c_{i2a}$, and let the i^{th} row of $\mathcal{A}$ be denoted by $\mathcal{A}_i$. Then one easily shows that:

Lemma 2. *Let i be a bus. Then $\mathbf{Var}(p_i) = \mathcal{A}_i^T\Omega\mathcal{A}_i$, and $\mathbf{E}(c_i(p_i)) = c_{i0}(\bar{p}_i^2 + \mathcal{A}_i^T\Omega\mathcal{A}_i) + c_{i1}\bar{p}_i + c_{i2}$.*

Next we provide our initial formulation, which extends the chance-constrained formulation in [3]. Let $n = |\mathcal{B}|$ and also $m = |\mathcal{E}|$ (see Section II). As inputs to the formulation we have safety parameters ν_{ij} and ν_i for each line ij and generator i, respectively ($\nu_i = 0$ at non-generator buses). The formulation uses variables $\bar{p}, \bar{\theta}$ (n-vectors), $\bar{f}$ (m-vector), $\mathcal{A}, D$ ($|\mathcal{S}| \times |\mathcal{S}|$ and $n \times |\mathcal{S}|$ matrices, respectively), and γ and s (an $m \times |\mathcal{S}|$ matrix and m-vector, respectively). As above, we use $\mathcal{A} \in \mathcal{K}$ to denote a given set of convex constraints on $\mathcal{A}$.

$$\min \sum_{i \in \mathcal{G}} \mathbf{E}(c_i(p_i)) \tag{15a}$$
$$\text{s.t. } \mathcal{A} \in \mathcal{K} \tag{15b}$$
$$B\bar{\theta} = \bar{p} + \mu - d \tag{15c}$$
$$\bar{f}_{ij} = b_{ij}(\theta_i - \bar{\theta}_j) \tag{15d}$$
$$b_{ij}|\bar{f}_{ij}| + \nu_{ij}\, s_{ij} \le f_{ij}^{\max} \quad \forall ij \in \mathcal{E}, \tag{15e}$$
$$\breve{B}\mathcal{A} = D \tag{15f}$$
$$\gamma_{ij,k} = \breve{B}_{i,k} - \breve{B}_{j,k} - D_{i,k} + D_{j,k}, \quad \forall ij \in \mathcal{E},\, k \in \mathcal{B} \tag{15g}$$
$$s_{ij} \ge b_{ij}\sqrt{\gamma_{ij}\Omega\gamma_{ij}^T} \quad \forall ij \in \mathcal{E} \tag{15h}$$
$$\forall i \in \mathcal{G}:$$
$$p_i^{min} + \nu_i\sqrt{\mathcal{A}_i^T\Omega\mathcal{A}_i} \le \bar{p}_i \le p_i^{max} - \nu_i\sqrt{\mathcal{A}_i^T\Omega\mathcal{A}_i}. \tag{15i}$$

It is straightforward to verify that this formulation captures the desired features. This is a second-order cone program which is amenable to solution by many current optimization packages. On the minus side, [9], [3], [19] have highlighted numerical challenges arising when attempting solutions of models like (15), on large transmission systems, by appealing to black-box solvers. A practical alternative is that of relying to simple but effective cutting-plane methods. We will return to this point later on.

To close this section, we point out some alternative modeling perspectives. See e.g. [22], [2]). An "ambiguous" model, that is to say, a distributionally robust model under Gaussianity is studied in [4].

IV. An example of the variance-cost trade-off

To motivate the forthcoming discussion, in this section we present a small example of problem (15) where there is a strong trade-off between cost and a system variance measure. Consider Figure 1.

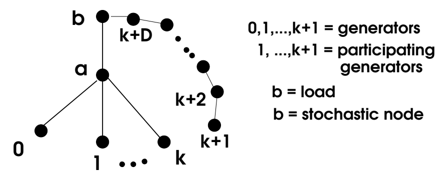

Fig. 1. High-variance example.

Here,

- k is large, as is D.
- The load at bus b equals L units.
- At bus b, stochastic output at bus is denoted by ω, with $\mathbf{E}(\omega) = \mu < L$ and $\mathbf{Var}(\omega) = \sigma^2$.
- At bus i ($0 \le i \le k+1$) there is a generator. Its cost function is linear, of the form $c_{i1}p_i$. Here $c_{01} < c_{11} = c_{21} \ldots = c_{k1} < c_{k+1,1}$.
- Generator at bus 0 has capacity larger than L and does not participate..
- At buses $1, 2, \ldots, k+1$ generators are all participating. Their lower limit is zero, their safety parameter is 3.
- There is no load, generation or stochastic injection at buses $a, k+2, \ldots, k+D$.
- Line limits are large.

Denote the participation factor for generator $i \ge 1$ as α_i.

It can be shown that the following solution,

$$\bar{p}_0 = L - \mu - 3\sigma \tag{16a}$$
$$\alpha_i = 1/k \text{ and } \bar{p}_i = 3\sigma/k, \text{ for } 1 \le i \le k. \tag{16b}$$
$$\alpha_{k+1} = \bar{p}_{k+1} = 0, \tag{16c}$$

is optimal (in fact: uniquely optimal) for the case of problem (15) in Figure 1. For a proof, see [18].

It is clear that the stochastic flow on line ab equals $L - \mu - \omega$, with variance σ^2. Thus f_{ab} displays the entirety of stochastic injection variance. Further, the sum of flow variances equals $\sigma^2(1 + 1/k) \approx \sigma^2$ since k is large.

Suppose we attempt to decrease $\mathbf{Var}(f_{ab})$ by a factor of $\Gamma < 1$. This goal will be achieved by setting $\sum_{i=1}^k \alpha_i = \sqrt{\Gamma}$ and thus $\alpha_{k+1} \approx 1 - \sqrt{\Gamma}$. In that case the sum of variances will be at least $\Gamma\sigma^2 + (D+1)(1-\sqrt{\Gamma})^2\sigma^2$. When $D = 10$ and $\Gamma = .5$ this quantity equals approximately $1.44\sigma^2$: in other words the sum of line flow variances has *increased* by more than 40%. As the reader no doubt has anticipated, putting emphasis on a particular variance metric brings forth a trade-off with other (perhaps equally compelling) metrics.

For a more extreme example, suppose now that the line limits are not very large. Consider an alternative variance metric that does take into account line limits:

$$\sum_{ij \in \mathcal{E}} \frac{\mathbf{Var}(f_{ij})}{(f_{ij}^{\max})^2} \tag{17}$$

We modify the data in the example above as follows:

- $\mu = L/4$ and $\sigma = \mu/2 = L/8$.
- Lines $0a$ and ab have limit $9L/8$.
- Lines ia $(1 \le i \le k)$ and lines on the path from $k+1$ to b all have limit 2σ.
- Line safety parameters are set to 3.

It can again be easily shown [18] that the solution given by (16) remains feasible, and is the sole optimal solution. The variance metric (17) attained by solution (16) equals $1/81 + 1/(4k) \approx .0123$ for k large.

Suppose that, as above, we wish to reduce the variance of f_{ab} by $0 \le \Gamma < 1$. The participation factors that attain the correction will now attain a variance metric of (at least) $\Gamma^2/81 + (1-\sqrt{\Gamma})^2(D+1)/4$. In the case $D = 10$, $\Gamma = 0.5$ this quantity is approximately 0.242. In other words, in our attempt to reduce variance in one line we have increased the variance metric (17) by approximately a factor of 20. And even if $D = 3$ the variance metric at the new solution equals 0.092: an increase by more than 8 times relative to the original value.

To conclude this section we point out that:

1) It is not true that *every* feasible solution to the security-constrained problem will accumulate variance on line ab. In fact our analyses show that we can transfer variance to the path $k+1, k+2, \ldots, b$. Doing so, however, will increase cost and total system variance.
2) This example is arguably unrealistic – different parameter choices will result in less radical behavior. Nevertheless, it is not difficult to see that the example can be modified to obtain a system setup that is not obviously extreme, while preserving or even amplifying the effect depicted above. The salient point is that the underlying economics and network structure is arguably efficient, because the large, most efficient generator can be used for deterministic demand coverage, while many participating generators are near the load.
3) The variance-concentration in the example, in summary, is due to interaction of network topology, load structure (i.e. location of loads, stochastic injection nodes and responding buses) and cost structure resulting in large variance caused by overlap of flow paths. To put it differently, a variance-aware reformulation of the security-constrained problem will tend to manage, or reduce, risky overlap of flow paths.

In the next section we will consider, in a systematic fashion, how to addresses the cost-variance trade-off.

V. Variance-aware problems

Here we describe a generic procedure that modifies problem (15), so as to better manage the trade-off between cost and a number of variance metrics. Our procedure seeks to find good solutions for a problem of the general type

$$\min \quad \sum_{i \in \mathcal{G}} \mathbf{E}(c_i(p_i)) \; + \; \Delta(\bar{f}, s^2) \tag{18a}$$

$$\text{s.t. } 15\text{b} - 15\text{i} \tag{18b}$$

where s^2 is a vector of variances. To fix ideas, an example is that where $s^2 = \{s^2_{ij}\}$ are the line flow variances of line flows, which depend on the participation matrix $\mathcal{A}$ as in (13), (14)). Finally,

$$\Delta(\bar{f}, s^2) \;=\; \sum_{ij \in \mathcal{E}} \Delta_{ij}(\bar{f}_{ij}, s^2_{ij}). \tag{19}$$

is the variance metric. Here each Δ_{ij} is a nonnegative function chosen to highlight a specific penalty as a function of expected flow and variance.

Formulation (18) includes all constraints of the safety-constrained problem (15) but the objective forces the desired trade-off. The formulation (and our templates below) are easily modified to account for many other variance metrics, for example, variability of output of participating generators.

As the reader can observe, one can scale the Δ term in (18a) by a positive constant, obtaining a problem with the same functional form that places a different emphasis on the cost-variance trade-off. Seen from this perspective, problem (18) bears a resemblance to traditional mean-variance optimization problems arising in financial portfolio analysis; see [23]. In Section VI we will return to this viewpoint.

A. Some variance metrics

Here we describe several relevant variance metrics. Also see Section VI.

(I) For all $ij \in \mathcal{E}$, Δ_{ij} is convex and nondecreasing in s^2_{ij}. The simplest case is that where $\Delta_{ij}(\bar{f}_{ij}, s_{ij}) = \psi_{ij} s^2_{ij}$ where $\psi_{ij} \ge 0$. When $\psi_{ij} = 1$ or $\psi_{ij} = (1/f^{max}_{ij})^2$ $\Delta(\bar{f}, s^2)$ we recover the metric in the examples in Section IV .

(II) Let $N > 0$ be a given integer, and consider a function Δ of the form

$$\Delta(\bar{f}, s^2) \;=\; \sum_{ij \in \mathcal{F}} \psi_{ij} s^2_{ij}. \tag{20}$$

Here the $\psi_{ij} \ge 0$ are given and $\mathcal{F} \subseteq \mathcal{E}$ is a *solution-dependent* set, for example:

(II.1) the set of N lines with largest average flow magnitude.

(II.2) the set of N lines with largest flow variance.

In either model, we do not know in advance the set $\mathcal{F}$ to be summed over. In case of Model (II.2), problem (18) can be written as

$$\min \quad \sum_{i \in \mathcal{G}} \mathbf{E}(c_i(p_i)) \; + \; \mathcal{V} \tag{21a}$$

$$\text{s.t.} \quad \mathcal{V} \ge \sum_{(i,j) \in \mathcal{H}} s^2_{ij}, \quad \forall\, \mathcal{H} \subseteq \mathcal{E} \text{ with } |\mathcal{H}| = N \tag{21b}$$

$$(15\text{b}) - (15\text{i}). \tag{21c}$$

Formulation (21a) is an SOCP of exponential size. However the formulation suggests a practical algorithm (i.e. a cutting-plane algorithm) as well as a theoretically sound polynomial-time method (relying on the ellipsoid method). We theorize that Model (II.1) can also be formulated as a convex program that is polynomially solvable.

(III) Let:

$$\begin{aligned}\Delta_{ij} &= -\rho_{ij}\log(s_{ij}^2 - b_{ij}^2\gamma_{ij}\Omega\gamma_{ij}^T), \quad \text{if } s_{ij}^2 > b_{ij}^2\gamma_{ij}\Omega\gamma_{ij}^T \\ &= +\infty, \quad \text{otherwise,}\end{aligned}$$

(where $\rho_{ij} > 0$ is given) i.e. the standard *logarithmic barrier* function [24]. In this setting, the conic constraints (15h) would not be used. When the ρ_{ij} are all equal to a common value ρ the solution to (18) (with (15h) not imposed) will converge to an optimum solution for (15) as $\rho \to 0^+$. Moreover, the structural form of this particular Δ function, for any choice of the ρ_{ij}, already imposes a trade-off of risk against operating cost – the tradeoff reflects the "nearness" of each line to its limit, in a natural sense, but without the explicit (and very computationally expensive) enforcement of the conics (15h). Of course, this risk perspective is, potentially, *not necessarily* the same as a variance-aware stance, because a line whose flow is not near the limit may still experience high variance of flow, and controlling such variance may be deemed important. Nevertheless, a line whose average flow is near the line limit is one where it is important to control variance, from a very direct perspective, as implied e.g. by constraints (15e), (15h) of (15).

We find this choice of metric appealing because of the success of very mature state-of-the-art logarithmic barrier solvers, which can furthermore be adapted to handle additional, non-convex features [25], [26].

B. Iterative solution

In this section we focus on problem (18). Rather than discussing direct solution to this problem, however, we will instead suggest an iterative corrective procedure that appropriately modifies an optimal solution to the standard safety-constrained problem (15) so as to make it into a "much more" variance-aware solution that is still near-optimal for (15).

There are two separate (but related) reasons for this goal. First of all, previous work in the literature on chance-constrained DC-OPF [3], [6] has shown that direct solution to (15), using a standalone solver, can be a challenging numerical task in the case of large transmission systems. However, this difficulty is easily resolved by employing a *cutting-plane* algorithm that iteratively outer-approximates the conic constraints (15h) using linear inequalities. In other words we rely on an algorithm that repeatedly solves problems of the form (15) with *all* constraints (15h) removed; the algorithm opportunistically detects which constraints (15h) are violated by a computed solution and approximates such violated constraints using tangent hyperplanes. It has been observed that such an approximation algorithm typically requires relatively few iterations to converge, and, most important, the number of lines ij that demand attention in the approximation is very small.

Beyond the purely computational insight (which of course is important) this observation brings forth an important risk-related perspective; namely, that only a few lines appear exposed to risk in the context of our generic problem (15). One may wonder why it is the case that realistic transmission systems have this behavior. To a certain extent this observation has been known in the industry in the context of non-risk aware DC-OPF. Recent work [27] performs systematic experiments that analyze the solution to the standard DC-OPF problem. They find that the number of tight constraints (1c) is limited to a few percent of the total and sometimes even less than that. Additionally, even large changes in the loads do not change the set of risky lines. We conclude that this is indeed a persistent feature in realistic cases.

In any case, these observations suggest an algorithm that shifts risk (in our context, variance) from what should be a small set of lines, to a broader set. The algorithm would start from an optimal solution to the standard safety-constrained problem (15), and correct that solution by taking simple, variance-metric reducing steps. In doing so we would be attaining three goals at once: (i) reduce the numerical complexity, which may be worse for the variance-aware problem (18) than it is for (15), (ii) leverage the underlying economic-risk dynamic we have just discussed, and (iii) explicitly reveal to a decision maker the set of lines that are exposed to high variance. A template for our scheme is given next:

Template V.1. *GENERIC CORRECTION TEMPLATE*

Input: an instance of the safety-constrained problem (15) and a variance metric.
Step I. Solve (15), with solution $(\bar{p}^*, \mathcal{A}^*)$.
Step II. Execute a small number of iterations that transform $(\bar{p}^*, \mathcal{A}^*)$ into a new feasible solution to (15) which attains a decreased value of the variance metric, while also moderately increasing generation cost.

An appropriate implementation of Step **II** is important. Below, we focus on a strategy that amounts to solving a convex optimization problem with relatively few constraints.

C. Implementing the template

To simplify the presentation we will assume that for each line ij, $\Delta_{ij}(\bar{f}_{ij}, s_{ij}^2) = \Delta_{ij}(s_{ij}^2)$, i.e. our metric only depends on variance, but we stress that the analysis of procedure V.3 given below extends to the more general case. We also assume that $\Delta_{ij}(s_{ij}^2)$ is convex nondecreasing.

Further,

Definition V.2. *Let $\bar{f}$ be a flow vector and $\mathcal{A} \in K$ a matrix of participation factors. We define the pair $(\bar{f}, \mathcal{A})$ as* **compatible** *(or say that $\mathcal{A}$ is compatible with $\bar{f}$) for (15), if there exist $\bar{p}, \bar{\theta}, D, \gamma$ and s so with $(\bar{p}, \bar{f}, \bar{\theta}, \mathcal{A}, D, \gamma, s)$ feasible for (15).*

Hence $(\bar{f}, \mathcal{A})$ is compatible if they give rise to a feasible solution to (15).

We describe an implementation of Step II of Template V.1 which repeatedly solves two convex optimization problems – a formal procedure is given in Procedure V.3 given below. We will next describe the two problems and comment on their use afterwards with Procedure V.3 presented after that. Let $0 < \tau < 1$ be fixed.

The first problem is denoted by $\mathbf{Reroute}(\hat{\mathcal{A}}, \tau)$. Its inputs are a compatible pair $(\hat{f}, \hat{\mathcal{A}})$ and the vector $\hat{s}^2 = \mathbf{V}(\hat{\mathcal{A}})$ of line flow variances (under participation factor matrix $\hat{\mathcal{A}}$).

$$\min_{\bar{p}, \bar{f}, \bar{\theta}} \quad \sum_{i \in \mathcal{G}} c_{i0}(\bar{p}_i^2 + \hat{\mathcal{A}}_i^T \Omega \hat{\mathcal{A}}_i) + c_{i1}\bar{p}_i + c_{i2} \tag{22a}$$

$$\text{s.t.} \quad B\bar{\theta} = \bar{p} + \mu - d \tag{22b}$$

$$b_{ij}|\bar{f}_{ij}| + \nu_{ij}\, \hat{s}_{ij} \ \le \ (1-\tau) f_{ij}^{\max} \quad \forall ij \in \mathcal{E}, \tag{22c}$$

$$\bar{f}_{ij} \ = \ b_{ij}(\bar{\theta}_i - \bar{\theta}_j) \quad \forall ij \in \mathcal{E}, \tag{22d}$$

$\forall i \in \mathcal{G}$:

$$p_i^{min} + \nu_i \sqrt{\hat{\mathcal{A}}_i^T \Omega \hat{\mathcal{A}}_i} \ \le \bar{p}_i \le p_i^{max} - \nu_i \sqrt{\hat{\mathcal{A}}_i^T \Omega \hat{\mathcal{A}}_i}. \tag{22e}$$

Our second problem takes as input a compatible pair $(\bar{f}', \mathcal{A}')$. In the description of the problem, we write $s'_{ij} = \sqrt{\mathbf{Var}(\mathcal{A}')_{ij}}$, and $\mathbf{T}(\bar{f}', \mathcal{A}', \tau)$ denotes the family of lines whose which the safety constraint is almost tight under the participation matrix $\mathcal{A}'$:

$$\mathbf{T}(\bar{f}', \mathcal{A}', \tau) = \{\, ij \ : \ |\bar{f}'_{ij}| + \nu_{ij} s'_{ij} \ge (1-\tau) f_{ij}^{max} \}. \tag{23}$$

The problem, denoted $\mathbf{VShift}(\bar{\mathbf{f}}', \mathcal{A}', \tau)$, is as follows:

$$\min_{s, \mathcal{A}} \quad \sum_{ij} \Delta_{ij}(s_{ij}^2) \tag{24a}$$

$$\text{s.t.} \quad \mathcal{A} \in \mathcal{K} \tag{24b}$$

$$s_{ij}^2 \ \ge \ b_{ij}^2 \pi_{ij}^T (I - \mathcal{A}) \Omega (I - \mathcal{A}^T) \pi_{ij} \quad \forall\, ij \in \mathcal{E} \tag{24c}$$

$$|\bar{f}'_{ij}| + \ \ \nu_{ij}\, s_{ij} \ \le \ f_{ij}^{\max} \quad \forall ij \in \mathbf{T}(f', \mathcal{A}', \tau). \tag{24d}$$

Comments: $\mathbf{Reroute}(\hat{\mathcal{A}}, \tau)$ minimizes expected generation cost using the constant participation matrix $\hat{\mathcal{A}}$. It imposes tighter line safety constraints (i.e. by a factor of $(1-\tau)$ in (22c)). Thus, this procedure reroutes flow so as to create slack capacity in all lines. Let an optimal solution be $(\bar{p}^*, \bar{f}^*, \bar{\theta}^*)$, assuming feasibility. Then clearly $(\bar{f}^*, \hat{\mathcal{A}})$ is compatible for (15), with "slack" at least $1-\tau$ on each line (multiplicatively). Choosing τ large can cause infeasibility of $\mathbf{Reroute}(\hat{\mathcal{A}}, \tau)$; but in our implementation of Procedure V.3, τ will be small, which as a byproduct will only slightly increase expected generation cost.

With regards to $\mathbf{VShift}(\bar{\mathbf{f}}', \mathcal{A}', \tau)$, its purpose is to improve on the variance metric while keeping flows and average generation fixed. Let us call the optimal solution $\hat{\mathcal{A}}$, and the corresponding line flow standard deviations, $\hat{s}$. The fact that the functions Δ_{ij} are nondecreasing implies, without loss of generality, that at optimality all constraints (24c) will be tight. Thus, by (24d), on lines $ij \in \mathbf{T}(f', \mathcal{A}', \tau)$ $|\bar{f}'_{ij}| + \nu_{ij}\hat{s}_{ij} \le f_{ij}^{\max}$. The definition of the set $\mathbf{T}$ makes it tempting to say that $(f', \hat{\mathcal{A}})$ are compatible: that $|\bar{f}'_{ij}| + \nu_{ij}\hat{s}_{ij} \le f_{ij}^{\max}$ holds for *all* lines ij and not just for $ij \in \mathbf{T}$, and thus we get a reduction of variance metric "for free" (without cost increase). As we will see below, the desired compatibility will in general not hold, and an appropriate algorithmic correction will be needed. The numerical difficulty entailed by problem (24) will depend on the number of constraints (24d) which, as we have discussed, is frequently quite small.

Our formal template is as follows:

Procedure V.3. *Variance-shifting*

Input: Feasible solution $(\bar{\mathbf{p}}^{\mathbf{0}}, \mathbf{f}^{\mathbf{0}}, \mathcal{A}^{\mathbf{0}})$ to safety-constrained problem (15), variance metric $\mathbf{\Delta}$, parameters $0 < \tau < 1$, $K > 0$. Let $\mathbf{s_0^2} = \mathbf{V}(\mathcal{A}^{\mathbf{0}})$.
For $k = 1, 2, \ldots, K$ **perform iteration k:**

1. Solve $\mathbf{Reroute}(\mathcal{A}^{\mathbf{k-1}}, \tau)$.
   If infeasible, **STOP.**
   Else, let $(\bar{\mathbf{p}}^{\mathbf{k}}, \bar{\mathbf{f}}^{\mathbf{k}}, \bar{\theta}^{\mathbf{k}})$ be the optimal solution.
2. Solve $\mathbf{VShift}(\bar{\mathbf{f}}^{\mathbf{k}}, \mathcal{A}^{\mathbf{k-1}}, \tau)$, with solution $(\hat{\mathbf{s}}_{\mathbf{k}}, \hat{\mathcal{A}}^{\mathbf{k}})$.
3. Choose $0 \le \lambda \le 1$ largest, so that
   $(\bar{\mathbf{f}}^{\mathbf{k}}, (\mathbf{1}-\lambda)\mathcal{A}^{\mathbf{k-1}} + \lambda\hat{\mathcal{A}}^{\mathbf{k}})$ is compatible.
4. Set $\mathcal{A}^{\mathbf{k}} \ \leftarrow \ (\mathbf{1}-\lambda)\mathcal{A}^{\mathbf{k-1}} + \lambda\hat{\mathcal{A}}^{\mathbf{k}}$, $\mathbf{s_k^2} = \mathbf{V}(\mathcal{A}^{\mathbf{k}})$.
5. If $\mathbf{\Delta}(\mathbf{s_k^2}) \ge \mathbf{\Delta}(\mathbf{s_{k-1}^2})$. **STOP**.

Next we comment on this algorithm[2] with a formal analysis provided below.

In Steps 3-4 we take a convex combination of the previous and the new participation matrices so as to obtain compatibility. For any line ij, $\mathbf{V}((1-t)\mathcal{A}^{k-1} + t\hat{\mathcal{A}}^k)_{ij}$ is a convex quadratic over $0 \le t \le 1$ and thus the determination of λ (Step 3) can be done exactly. Usually $\mathbf{\Delta}(\hat{\mathbf{s}}_{\mathbf{k}}^{\mathbf{2}}) < \mathbf{\Delta}(\hat{\mathbf{s}}_{\mathbf{k-1}}^{\mathbf{2}})$ (by choice of $(\hat{\mathbf{s}}_{\mathbf{k}}, \hat{\mathcal{A}}^{\mathbf{k}})$ in Step 2). Hence choosing λ as large as possible will help toward reducing variance metric. This is why we use inequality (24d) in $\mathbf{VShift}(\bar{\mathbf{f}}^{\mathbf{k}}, \mathcal{A}^{\mathbf{k-1}}, \tau)$: without it, each ij in $\mathbf{T}(\bar{f}^k, \mathcal{A}^{k-1}, \tau)$ with $\hat{s}_{k,ij}$ large would force a small value for λ.

Moreover, the procedure could terminate in Step 1 of some iteration. This feature is avoided by, e.g. scaling τ by a factor of $1/2$ and then repeating Step 1 (until feasibility is attained). We have not implemented this strategy as it did not prove necessary for small values of τ and K.

Next we provide an analysis of Procedure V.3. Lemmas 4 and 5 are the critical result; together they essentially show that, under fairly general conditions, any execution of Steps 2-3 is guaranteed to make progress in the sense that a positive length stepsize λ can be chosen. Theorem 6 is an important corollary; it shows that if the algorithm ever stops in Step 5, then under Model (I) variance metric has reached its minimum possible value.

Remark 3. *Let $\mathcal{A}, \mathcal{A}' \in \mathcal{K}$. Then for any $0 \le t \le 1$, $(1-t) + \mathcal{A} + t\mathcal{A}' \in \mathcal{K}$ and for any line ij, $\mathbf{V}((1-t)\mathcal{A} + t\mathcal{A}')_{ij}$ is a convex quadratic function of t.*

Proof. The first claim follows since $\mathcal{K}$ is convex and the second using expression (8). □

Lemma 4. *Suppose that in iteration k the algorithm reaches Step 2. Let $\mathcal{A} \in \mathcal{K}$ be an* **arbitrary** *participation matrix. Then there exists $0 < \gamma \le 1$ such that for all $0 \le t \le \gamma$,*

$$(\,\bar{f}^k, (1-t)\mathcal{A}^{k-1} + t\mathcal{A}\,)$$

is a compatible pair.

[2]motivated by [28]

Proof. Consider any line ij. For real $0 \le t \le 1$ let $s_{ij}^2(t) = \mathbf{V}((1-t)\mathcal{A}^{k-1} + t\mathcal{A})$. By construction in Step 1,

$$|\bar{f}_{ij}^k| + \nu_{ij} s_{ij}(0) \ \le\ (1-\tau) f_{ij}^{max}.$$

Hence we can find $\gamma_{ij} > 0$ with

$$|\bar{f}_{ij}^k| + \nu_{ij} s_{ij}(t) \ \le\ f_{ij}^{max}$$

for all $t \le \gamma_{ij}$. The result is obtained by setting $\gamma = \min_{ij} \gamma_{ij}$. □

Lemma 5. *Suppose that the algorithm reaches Step 2 of iteration k. Suppose $\bar{\mathcal{A}}$ is any matrix with $\Delta(V(\bar{\mathcal{A}})) < \Delta(s_{k-1}^2)$. Then for any $0 < t \le 1$,*

$$\Delta(\mathbf{V}(\,(1-t)\mathcal{A}^{k-1} + t\bar{\mathcal{A}}\,)) < \Delta(s_{k-1}^2). \tag{25}$$

Proof. Consider any line ij. Since

$$s_{ij}^2(t) \doteq \mathbf{V}((1-t)\mathcal{A}^{k-1} + t\bar{\mathcal{A}})$$

is a convex quadratic function of t,

$$s_{ij}^2(t) \le (1-t)s_{ij}^2(0) + t s_{ij}^2(1),$$

and therefore, since $\Delta_{ij}(s_{ij}^2)$ is convex and nondecreasing in s_{ij},

$$\Delta_{ij}(s_{ij}^2(t)) \le (1-t)\Delta_{ij}(s_{ij}^2(0)) + t\Delta_{ij}(s_{ij}^2(1)).$$

Summing this expression over all ij we obtain

$$\Delta(s^2(t)) \le (1-t)\Delta(s_{k-1}^2) + t\Delta(V(\bar{\mathcal{A}})) < \Delta(s_{k-1}^2) \tag{26}$$

since $t > 0$. □

Let $\mathcal{A}^*$ and Δ^* denote, respectively, the participation matrix with minimum variance metric, and the metric it attains, i.e.

$$\mathcal{A}^* \doteq \ \text{argmin}_{\mathcal{A} \in \mathcal{K}} \Delta(\mathbf{V}(\mathcal{A})) \text{ and } \Delta^* \doteq \ \Delta(\mathbf{V}(\mathcal{A}^*)). \tag{27}$$

Based on the above results we can state an important consequence under Model (I) of the variance metric (Section V-A).

Theorem 6. *Under Model (I) if Procedure (V.3) stops at Step 5 of iteration k then $\Delta(\mathbf{s_{k-1}^2}) = \Delta^*$.*

Proof. Assume, by contradiction, that $\Delta^* < \Delta(s_{k-1}^2)$. The contradiction we will provide will show that the procedure does not stop at Step 5. By Lemma 5, applied to matrix $\mathcal{A}^*$, we have that

$$\Delta(\mathbf{V}(\,(1-t)\mathcal{A}^{k-1} + t\mathcal{A}^*\,)) < \Delta(s_{k-1}^2) \ \forall\ 0 < t \le 1. \tag{28}$$

And if we apply Lemma 4, also to matrix $\mathcal{A}^*$, we have that there exists $0 < \gamma \le 1$ such that

$$(\,\bar{f}^k, (1-\gamma)\mathcal{A}^{k-1} + \gamma\mathcal{A}^*\,)$$

is a compatible pair. Hence $(1-\gamma)\mathcal{A}^{k-1} + \gamma\mathcal{A}^*$ is feasible for $\mathbf{VShift}(\mathbf{\bar{f}^k}, \mathcal{A}^{\mathbf{k-1}}, \tau)$. Since (Step 2) $\hat{\mathcal{A}}^k$ is the optimal solution to this problem, we therefore have by (28)

$$\Delta(\mathbf{V}(\hat{\mathcal{A}}^k)) \ < \ \Delta(\mathbf{V}(\mathcal{A}^{k-1})).$$

Now we apply Lemma 4 to $\hat{\mathcal{A}}^k$, and conclude that the stepsize λ computed in Step 5, is strictly positive. And if we also apply Lemma 5 to matrix $\hat{\mathcal{A}}^k$ we obtain that

$$\Delta(\mathbf{V}(\,(1-t)\mathcal{A}^{k-1} + t\hat{\mathcal{A}}^k\,)) < \Delta(s_{k-1}^2) \ \forall\ 0 < t \le 1. \tag{29}$$

Evaluating (29) at $t = \lambda$ yields the desired contradiction, i.e. $\mathbf{\Delta}(\mathbf{s_k^2}) < \mathbf{\Delta}(\mathbf{s_{k-1}^2})$. □

D. Numerical examples for the correction template

In this section we apply Procedure V.3 to the Polish grid example case2746wp [29]. This system has 2746 buses, 3514 branches, 520 generators total load 24873. We introduced 22 stochastic injection sites, with total average injection 4611.57 (i.e., roughly 18.5% penetration). The ratio of standard deviation to mean of stochastic injections was 0.3, and it is assumed that injections at different stochastic sites were independent. All safety parameters (for lines and generators) were set to 3. All generators were made available for balancing. These choices follow the setup in [3]. We also chose $\tau = 0.1$ and $K = 2$.

The variance metric we selected is nonconvex. We used

$$\sum_{ij \in \mathcal{F}} s_{ij}^2, \tag{30}$$

where the set $\mathcal{F} = \mathcal{F}(\bar{f}, s) \subset \mathcal{E}$ of lines is, as our notation indicates, flow and variance dependent and is designed to highlight lines that are "at risk"; it is made up by the combination of two sets:

(a) The 100 lines that attain largest average flow magnitude, as in Model (II.1).
(b) All lines ij for which $|\bar{f}_{ij}| + \nu_{ij} s_{ij} \ge (1-\tau) f_{ij}^{max}$, in other words, lines whose flow and variance characteristics place them in a near-risk condition.

The computation of the solution $(\mathbf{\bar{p}^0}, \mathbf{f^0}, \mathcal{A}^\mathbf{0})$ to safety-constrained problem (15) was performed using Cplex [34] and required approximately one CPU minute (on a standard workstation) and 29 cutting-plane steps. The computation selected 11 generators with positive participation parameer α and in the rest of the procedure we used this subset of buses as the set of participating buses.

We then applied the iterative component of Procedure V.3. We provide a brief outline of the run, next.

Iteration k = 1, Step 1. This computation took 1.12 seconds, and produced a solution whose expected generation cost $\approx 1.1 \times 10^{06}$, which is nearly identical to that in the initial chance-constrained OPF solution.

Iteration k = 1, Step 2. As anticipated by the above discussion, the set $\mathbf{T}$ (see (23)) of risky lines in the solution computed in Step 1 is quite small – just five lines are in the set. As a result, in (30) we have $|\mathcal{F}| = 105$, with variance metric taking value 6.3×10^{04}.

Problem $\mathbf{VShift}(\mathbf{\bar{f}^1}, \mathcal{A}^\mathbf{0}, \tau)$ had, roughly, 14000 variables and $1e06$ nonzeros. Nevertheless, it was solved using Gurobi 7.02 [35] in just 2.3 seconds. The updated variances, $\mathbf{\hat{s}_1^2}$ and $\mathbf{\bar{f}^1}$, attained metric (30) of $\approx\ 2.3 \times 10^{04}$– of course this computation is done using the set $\mathcal{F} = \mathcal{F}(\mathbf{\bar{f}^1}, \mathbf{s_0})$, rather than the correct set, $\mathcal{F}(\mathbf{\bar{f}^1}, \mathbf{\hat{s}_1})$. As such the value we have just provided is not a correct calculation of variance metric, however it indicates a large potential decrease of metric compared to that in the solution obtained in Step 1 (which was nearly three times larger). In fact, $\mathbf{\hat{A}^1}$ is not necessarily even compatible with $\mathbf{\bar{f}^1}$. Steps 3 and 4 will recover compatibility.

Iteration k = 1, Step 3. The stepsize computation yields $\lambda \approx 0.55$.

Iteration k = 1, Steps 4 and 5. Updating the participation matrix yields variances $\mathbf{s_1^2}$ whose metric is $\approx 4.65 \times 10^{04}$.
Summary. In just one iteration, Procedure V.3 keeps generation cost approximately constant and reduces our variance metric by, roughly, 35%.
Iteration k = 2, Step 1. Similar behavior as in iteration 1.
Iteration k = 2, Step 2. The number of nearly-tight lines grows to 24; and now $\mathcal{F} = 120$. Solution for $\mathbf{VShift}(\bar{\mathbf{f}}^{\mathbf{2}}, \mathcal{A}^{\mathbf{1}}, \tau)$ are similar to those for iteration 1. In this case, $\bar{f}^2$ together with $\hat{s}_2^2$ attain variance metric 2.89×10^{04}, with the same caveat as indicated for Iteration 1, Step 2.
Iteration k = 2, Step 3. Now $\lambda \approx 0.29$.
Iteration k = 2, Steps 4 and 5. After the update, $\bar{\mathbf{f}}^{\mathbf{2}}$ and $\mathbf{s_2^2}$ yield a metric of approximately 4.50×10^{04}.
Summary. Thus, in just two iterations of Procedure V.3 we once again maintain nearly constant expected generation cost, but variance metric is decreased by roughly 40% compared to the original value.

It is also useful to consider the structure of the power flows. At termination, the largest magnitude average line flow magnitude is approximately 817. The third largest flow has value 632 and the corresponding line attains the largest single standard deviation; roughly 91. In contrast, the 101^{st} largest flow magnitude is approximately 144 and lines below that ranking attain far smaller standard deviation of flow; approximately 15. In other words the procedure transfers variance away from high flow lines while avoiding the creation of lines with high variance but low expected flow.

VI. Numerical experiments with variance-aware problems

Our formal problem (18) enforces a tradeoff between "performance," i.e. expected cost given in the first term of (18a), and "risk" which is the second term in (18a). The more weight that is placed on the risk term, i.e. the more risk-aware the computation is, the less emphasis that the optimization will place on performance. This is the "cost of variance". For any given structural form of the variance metric one can scale (up or down) the contribution of variance to the objective of (18). For example, in Models (I) and (II) of Section V-A one can scale the parameters ψ_{ij} by a common positive constant so as to adjust the impact of variance. Different choices of this common scale factor imply a different risk posture and likely a different solution to problem (18). In essence, thus, problem (18) represents a number of choices of operating solutions, each driven by a different risk posture.

To better examine the performance-risk tradeoff we slightly alter formulation (18): we choose parameters $\Lambda = 0$ or 1 and $0 \leq \Pi$, and address the problem

$$\min \quad \Lambda \sum_{i \in \mathcal{G}} \mathbf{E}(c_i(p_i)) \ + \ \Pi \Delta(\bar{f}, s^2) \tag{31a}$$

$$\text{s.t. } 15b - 15i. \tag{31b}$$

When $\Lambda = 1$ and $\Pi = 0$ the variance term is ignored and we obtain the standard chance-constrained DCOPF problem. When $\Pi > 0$ and $\Lambda = 0$ the goal is to minimize variance metric and cost is ignored. For $\Lambda = 1$, larger choices of Π represent a more risk-aware stance; the larger Π is, the more emphasis is placed on variance reduction. Thus when $\Lambda = 1$, problem (31) amounts to a version of (18) with scaling of the variance metric. [Remark: (31) is more flexible than (18) because it allows us to ignore either cost or variance].

We first describe experiments using case2746wp as in Section V-D. Again we assume independence of the stochastic injections, and that the variance of the injection at bus k is given by σ_k^2. We simplify the controllable generator setup (3) so that for each controllable generator i there is a constant α_i with the stochastic response at i given by $p_i(\omega) = \bar{p}_i \ - \ \alpha_i \sum_{j \in \mathcal{B}} \omega_j$, rather than the more general (3). Then the variance of generation at i is given by $\mathbf{Var}(p_i) = \alpha_i^2 (\sum_j \sigma_j^2)$. We tackle problem (31) using a variant of the cutting-plane algorithm in [3], which computes a solution that is both optimal and feasible within numerical tolerance.

We first consider the variance metric given by $\sum_{i \in \mathcal{B}} \mathbf{Var}(p_i)$, i.e. the sum of variances of output of controllable generators. In Table I, "cost" is expected generation cost and "var" is the variance metric.

TABLE I
Results on problem (31) for variance metric $\sum_{i \in \mathcal{B}} \mathbf{Var}(p_i)$

$(\mathbf{\Lambda}, \mathbf{\Pi})$	$\mathbf{1.0, 0.0}$	$\mathbf{1, 1e1}$	$\mathbf{1, 1e2}$	$\mathbf{1, 1e3}$	$\mathbf{0.0, 1.0}$
cost	1.262e7	1.269e7	1.284e7	1.291e7	1.935e8
var	9.21e4	6.55e3	1.08e3	7.48e2	7.37e2

Comment. The problems solved when both $\Lambda = 1$ and $\Pi > 0$ show a large range of values for system variance, with small but significant changes in generation cost. Note that the cost in the $(1, 2e3)$ case is approximately 3% higher than in the standard chance-constrained solution (i.e. the $(1, 0.0)$ case), a nontrivial increase. The problems solved when $(\Lambda, \Pi) = (1, 0)$ or $(0, 1)$ disregard, respectively, variance and generation cost. The high variance metric attained in the $(1, 0)$ case highlights the potential danger in optimizing with respect to cost alone. In fact, a comparison of the $(1, 0)$ and $(1, 1e3)$ columns illustrates what might be termed the "cost of variance".

The second set of tests that we perform use, as variance metric, a sum of variances of flow on "important" lines. In particular we use the quantity $\sum_{ij \in \mathcal{E}(N)} s_{ij}^2$ where for an integer $N > 0$, $\mathcal{E}(N)$ is the set of lines with largest variance in the solution to the standard chance-constrained DCOPF problem (CCOPF problem in the sequel; this is related to Model (II) in Section V-A). To set up these tests, we first consider Table II, which shows various values of the variance metric that are attained by the solution of the standard problem. In our variance-aware problem, we again consider problem (31) with the metric $\sum_{ij \in \mathcal{E}(N)} s_{ij}^2$. We used $\Lambda = 1.0$. Parameter Π was chosen so as to attain a meaningful tradeoff between the two terms in (31a) while avoiding numerical

TABLE II
Variance metric $\sum_{ij \in \mathcal{E}(N)} s_{ij}^2$ in standard CCOPF solution

N	10	50	200	500
metric	1.2e4	4.03e4	7.35e4	1.13e5

difficulties in the underlying solver. Results are given in Table III. Note that the variance metric of interest (third row of Table III) is reduced as compared to Table II, as desired. In this case generation cost (fifth row) remains near-constant. Total line variance (fourth row) does not show large changes.

TABLE III
RESULTS ON PROBLEM (31) FOR VARIANCE METRIC $\sum_{ij\in\mathcal{E}(N)} s_{ij}^2$

N	10	50	200	500
Π	1e3	1e3	1e3	1e3
$\sum_{\mathbf{ij}\in\mathcal{E}(\mathbf{N})} \mathbf{s_{ij}^2}$	9.78e3	3.06e4	5.49e4	8.12e4
$\sum_{\mathbf{ij}\in\mathcal{E}} \mathbf{s_{ij}^2}$	2.07e5	1.90e5	1.86e5	1.97e5
cost	1.27e7	1.27e7	1.27e7	1.27e7

VII. SHARPE-RATIO PROBLEMS

In the previous section we performed experiments involving problem (31a) which optimizes a blend of performance and risk, with parameterization used to take a stance on the relative importance of variance reduction. This task may be nontrivial though an experienced operator may be able to tailor the trade-off based on an assessment of current operating conditions. In this section we describe an approach that bypasses the need for parameter choice and, informally, seeks to maximize "cost savings per unit of variance". We will make this statement clearer below.

For completeness, we first describe an approach found in the economics literature that provides an alternative to solving problem (18). Here, one chooses a desired level of maximum risk (variance metric) that can be tolerated and solve a CCOPF variant that caps variance metric at that maximum level. Or, conversely, one can place an upper bound on cost and minimize variance metric subject to that upper bound on cost. Formally, one would choose parameters ρ_s and ρ_c, and consider problems

$$c^*(\rho_s) \doteq \min \sum_{i\in\mathcal{G}} \mathbf{E}(c_i(p_i)) \tag{32a}$$

$$\text{s.t. } 15b - 15i \tag{32b}$$

$$\Delta(\bar{f}, s^2) \leq \rho_s \tag{32c}$$

and

$$\Delta^*(\rho_c) \doteq \min \Delta(\bar{f}, s^2) \tag{33a}$$

$$\text{s.t. } 15b - 15i \tag{33b}$$

$$\sum_{i\in\mathcal{G}} \mathbf{E}(c_i(p_i)) \leq \rho_c \tag{33c}$$

The first problem minimizes cost subject to a limit on variance, and viceversa. When $\Delta(\bar{f}, s^2)$ is convex, so are both problems. We note that (32) and (33) are analogues of well-known problems in the (financial) portfolio optimization literature, termed the *maximum return* and *minimum risk* problems (resp.), see [30]. We have the (straightforward) result:

Lemma 7. *Let $\hat{c}$ and $\hat{\Delta}$ denote the expected cost and variance metric (resp.) attained by an optimal solution to the variance-aware problem (18), or to problem (31) with $\Lambda = 1$ and $0 < \Pi$. Then: $c^*(\hat{\Delta}) = \hat{c}$ and $\Delta^*(\hat{c}) = \hat{\Delta}$.*

Clearly an approach relying on problem (32a) or (33a) is not parameter-free – we need to specify ρ_s or ρ_c. Either choice represents a risk-stance on the part of the planner. However, again, either problem may be of interest to an experienced operator.

In the remainder of this section we present a parameter-free approach, which applies when the generation cost functions are strictly convex and the when the variance metric is convex and of appropriate structure, and is motivated by the classical "Sharpe ratio" problem in economics [31]. Formally, we proceed as follows:

Step 1. Let us assume that the generation cost functions are of the form $c_i(p) = c_{i0}p^2 + c_{i1}p + c_{i2}$ (as discussed above) and thus, as per Lemma 2, $\mathbf{E}(c_i(p_i)) = c_{i0}(\bar{p}_i^2 + \mathcal{A}_i^T\Omega\mathcal{A}_i) + c_{i1}\bar{p}_i + c_{i2}$. **We will further assume that** $c_{i0} > 0$ for each generator $i \in \mathcal{G}$. Defining

$$v_i \doteq \bar{p}_i + c_{i1}/2c_{i0}$$

then $\mathbf{E}(c_i(p_i)) = c_{i0}v_i^2 + c_{i2} - \frac{c_{i1}^2}{4c_{i0}} + c_{i0}\mathcal{A}_i^T\Omega\mathcal{A}_i$. Hence, up to an additive constant, the expression $\sum_{i\in\mathcal{G}} \mathbf{E}(c_i(p_i))$ is a **convex quadratic** (i.e. no linear term!) of the v and $\mathcal{A}$.

Step 2. Formulation (31) includes linear equations and inequalities, as well as conic constraints. In order to simplify the discussion below, we will use streamlined notation, as follows.

- **g.1** We will use the generic term "x" so as to refer to the vector of variables appearing in problem (18), including the v just defined (see the paragraph preceding the statement of problem (15) for a complete list of all its variables). Some additional variables, such as the v_i above, will be included for problem reformulation.
- **g.2** Constraints (15b)-(15i) are represented using the generic form $\Phi(x) \leq b$. Here $\Phi(x)$ is a vector-valued function; each row of $\Phi(x) \leq b$ corresponds to a row of (15b)-(15i) with equations (such as (15c)) equivalently restated as two inequalities. We further assume that constraint (15b), i.e. $\mathcal{A} \in \mathcal{K}$, is polyhedral or second-order cone representable. Thus, in other words, the constraint system $\Phi(x) \leq b$ is made up of linear and second-order cone constraints.
- **g.3** The expected cost term, $\sum_{i\in\mathcal{G}} \mathbf{E}(c_i(p_i))$, is represented by a strictly convex quadratic $g(x)$ with **no linear term**.
- **g.4** The variance-metric term will be denoted as $S^2(x)$. We assume this function is convex and homogeneous of degree-2, i.e. $S^2(tx) = t^2S^2(x)$ for any scalar t. The variance metric given by (II.2) in Section V-A is of this form.

Under assumptions **g1.-g4** we can equivalently rewrite the generic variance-aware safety-constrained problem (31) (with $\Lambda = 1$) as

$$\min \ g(x) + \Pi S^2(x) \tag{34a}$$

$$\text{s.t.} \quad \Phi(x) \leq b \tag{34b}$$

Now we will replace problem (34) with one that does not require the choice of Π, and yet captures the interplay between performance and risk.

Step 3. We assume that there is a constant $\kappa^2 > 0$ such that a solution to the safety-constrained problem is of interest so long as its expected cost is less than κ^2. In other words this constant represents the maximum that we are willing to pay for a safe solution. We will further assume that such "cheap enough" safe solutions exist, and that they always attain positive (e.g. nonzero) variance metric.

With Steps 1-3 in place, we can now state a formal optimization problem, namely:

$$R_1 \doteq \max \frac{\kappa^2 - g(x)}{S^2(x)} \tag{35a}$$

$$\text{s.t.} \quad \Phi(x) \leq b \tag{35b}$$

We will show that this problem is SOCP-representable. Prior to proving this fact we can discuss the structure and rationale behind this problem. The proofs below, at a very high level, are directed by the traditional analysis of the Sharpe ratio.

- The numerator $\kappa^2 - g(x)$ in (35a) measures the savings in expected cost relative to the benchmark κ^2. Intuitively we are optimizing the "cost savings per unit of variance".
- By assumption (Step 3 above) the ratio $\frac{\kappa^2 - g(x)}{S^2(x)}$ is positive and well-defined (nonzero denominator) for solutions x of interest.
- Under the assumption that the generator cost functions are strictly convex quadratics, the units in $g(x)$ are "megawatts squared". In many cases of the variance metric described above, the units in $S^2(x)$ are of the same type.
- If the generation costs are all linear (which is the case in some of the larger examples available in e.g. MATPOWER) one can replace the ratio in (36) with one where the denominator is $\sqrt{S^2(x)}$ and the analysis below will carry through.
- In analysis of actual transmission systems, the solution of problem (35) may simply be used as a "benchmark" for other operating plans, rather than as a plan to be actually implemented. A network operator may choose to implement a particular plan if e.g. that plan looks "good enough" benchmarks well against the the solution to (35).

We will now prove the stated convexity result. Toward this goal, let

$$R_2 \doteq \max \frac{1}{S^2(x)} \tag{36a}$$

$$\text{s.t.} \ \Phi(x) - \Theta b \leq 0 \tag{36b}$$

$$\kappa^2\Theta^2 - g(x) \geq 1 \tag{36c}$$

$$\Theta \geq 0. \tag{36d}$$

Constraints (36b) are the *homogenization* of constraints (35b), and the new variable Θ is usually termed the *homogenization* constant. For example, constraint (15f) in (31), $\breve{B}\mathcal{A} = D$, becomes $\breve{B}\mathcal{A} - \Theta D = 0$ in (36), constraint (15e) $b_{ij}|\bar{f}_{ij}| + \nu_{ij}\, s_{ij} \leq f_{ij}^{\max}$ becomes $b_{ij}|\bar{f}_{ij}| + \nu_{ij}\, s_{ij} - f_{ij}^{\max}\Theta \leq 0$ and constraint (15h), $s_{ij} \geq b_{ij}\sqrt{\gamma_{ij}\Omega\gamma_{ij}^T}$, remains unchanged since it has no constant term.

We will next show that (a) problems (35) and (36) are equivalent (Lemmas (8) and (9)) and (b) problem (36) is convex. Prior to the proofs we note that:

- since $g(x)$ is strictly quadratic, g is homogeneous of degree-2, i.e. $g(tx) = t^2 g(x)$, for every scalar t,
- We also have $\Phi(tx) = t\Phi(x)$ for any scalar $t \geq 0$ (refer again to problem (15)).

Lemma 8. $R_1 \geq R_2$.

Proof. Let $(\bar{x}, \bar{\Theta})$ be optimal for (36). By (36c) $\bar{\Theta} > 0$ since $g(x) \geq 0$. Thus, define $z \doteq \bar{x}/\bar{\Theta}$. Then $f(z) = \bar{\Theta}^{-1} f(\bar{x}) \leq b$ (by (36b)) and so z is feasible for (35).

Moreover,

$$\kappa^2 - g(z) = \frac{1}{\bar{\Theta}^2}(\kappa^2\bar{\Theta}^2 - g(\bar{x})),$$

and $\bar{\Theta}^2 S^2(z) = S^2(\bar{x})$. Thus

$$\frac{1}{S^2(\bar{x})} = \frac{1}{\bar{\Theta}^2}\frac{1}{S^2(z)} \leq \frac{1}{\bar{\Theta}^2}\frac{\kappa^2\bar{\Theta}^2 - g(\bar{x})}{S^2(z)} = \frac{\kappa^2 - g(z)}{S^2(z)},$$

where the inequality follows from constraint (36c). The rightmost term in this expression is the objective attained by z for problem (35). □

Lemma 9. $R_2 \geq R_1$.

Proof. Let $\tilde{x}$ be optimal for (35). By assumption on the quantity κ^2, $\kappa^2 > g(\tilde{x})$. Define

$$\Theta \doteq (\kappa^2 - g(\tilde{x}))^{-1}, \text{ and } y \doteq \Theta\tilde{x}.$$

Then $\kappa^2\Theta^2 - g(y) = \Theta^2(\kappa^2 - g(\tilde{x})) = 1$, and $f(y) - \Theta b = \Theta(f(\tilde{x}) - b) \leq 0$, where the inequality follows from feasibility of $\tilde{x}$ for (35). In other words, (y, Θ) is feasible for problem (36). Moreover,

$$R_1 = \frac{\kappa^2 - g(\tilde{x})}{S^2(\tilde{x})} = \frac{1}{\Theta^2}\frac{\kappa^2\Theta^2 - g(y)}{S^2(\tilde{x})} = \frac{1}{S^2(y)},$$

from which the result follows. □

As a consequence of Lemmas 8 and 9 we obtain $R_1 = R_2$ and indeed the proofs show that problems (35) and (36) are equivalent.

There remains to show that problem (36) is convex. First, the objective can be rewritten as $\min S^2(x)$ which is convex and will be SOCP-representable in many of the cases described above. As discussed, the constraints (36b) are either linear or convex-conic. Finally we have $\kappa^2\Theta^2 - g(x) \geq 1$. This constraint, which implies $\Theta \geq 1/\kappa$, can be restated as

$$\kappa^2(\Theta + 1/\kappa)(\Theta - 1/\kappa) \geq g(x).$$

Since $g(x)$ is a convex quadratic this constraint is SOCP-representable. See, e.g. [32].

VIII. Conclusion

In this work we have considered the interplay between expected cost minimization in operation of power grids under stochastic injections, and variability. We presented numerically efficient procedures that post-process non-variance aware minimum-cost solution so as to reduce its variance while controlling costs as w ell as theoretically efficient cost-variance tradeoff procedures, in particular a generalized Sharpe-ratio maximization procedure. An interesting topic for future research concerns the analysis of concrete ("real-world") transmission system with particular emphasis on studying relevant variance metrics, i.e. variability of transformer operation. Such problems may lead to non-convexity, but as argued here may be amenable to efficient post-processing through appropriate algorithms. Capturing important engineering details may call for advanced numerical optimization tools, such as derivative-free optimization [33]. A further issue of interest is the exploration of the performance-risk frontier through techniques such as those described in Section VII, which may prove helpful in concrete pricing mechanisms that account for variability of renewable output.

IX. Acknowledgment

This work was supported by DOE award GMLC77; grants HDTRA1-13-1-0021, ONR N00014-16-1-2889 and a DARPA RADICS award.

References

[1] L. Roald, F. Oldewurtel, T. Krause, and G. Andersson, "Analytical Reformulation of Security Constrained Optimal Power Flow with Probabilistic Constraints," in *Proceedings of the Grenoble PowerTech, Grenoble, France*, 2013. 1, 2, 3

[2] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, "A probabilistic framework for security constrained reserve scheduling of networks with wind power generation," in *Energy Conference and Exhibition (ENERGYCON), 2012 IEEE International*, sept. 2012, pp. 452–457. 1, 2, 3

[3] D. Bienstock, M. Chertkov, and S. Harnett, "Chance-constrained DC-OPF," *SIAM Review*, vol. 56, pp. 461–495, 2014. 1, 2, 3, 5, 7, 8

[4] M. Lubin, Y. Dvorkin, and S. Backhaus, "A robust approach to chance constrained optimal power flow with renewable generation," vol. 31, no. 5, Sept 2016, pp. 3840–3849. 1, 2, 3

[5] M. Vrakopoulou and I. A. Hiskens, "Optimal policy-based control of generation and hvdc lines in power systems under uncertainty," in *2017 IEEE Manchester PowerTech*, June 2017, pp. 1–6. 1, 2

[6] L. Roald and G. Andersson, "Chance-Constrained AC Optimal Power Flow: Reformulations and Efficient Algorithms," 2017. 1, 5

[7] A. Morrison, "Optimized Risk Limits for Stochastic Optimal Power Flow," Master's thesis, ETH, 2016. 1, 2

[8] D. Phan and S. Ghosh, "Two-stage stochastic optimization for optimal power flow under renewable generation uncertainty," *ACM Trans. Model. Comput. Simul.*, vol. 24, no. 1, pp. 2:1–2:22, Jan. 2014. [Online]. Available: http://doi.acm.org/10.1145/2553084 1

[9] L. Roald, S. Misra, T. Krause, and G. Andersson, "Corrective Control to Handle Forecast Uncertainty: A Chance Constrained Optimal Power Flow," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1626–1637, March 2017. 1, 2, 3

[10] H. Zhang and P. Li, "Chance constrained programming for optimal power flow under uncertainty," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2417–2424, Nov 2011. 1

[11] K. Sundar, H. Nagarajan, M. Lubin, L. Roald, S. Misra, R. Bent, and D. Bienstock, "Unit commitment with n-1 security and wind uncertainty," in *Power Systems Computation Conference (PSCC), 2016*. IEEE, 2016, pp. 1–7. 1

[12] S. Baghsorkhi and I. Hiskens, "Impact of wind power variability on sub-transmission networks," in *Power and Energy Society General Meeting*. IEEE, 2012. 1

[13] M. Davari, F. Toorani, H. Nafisi, M. Abedi, and G. B. Gharehpetian, "Determination of mean and variance of LMP using probabilistic DCOPF and T-PEM," in *2008 IEEE 2nd International Power and Energy Conference*, Dec 2008, pp. 1280–1283. 1

[14] R. Jabr, "A Conic Optimal Power Flow with Uncertain Prices," in *Power Systems Computation Conference (PSCC), 2005*, 2005, pp. 1–7. 1

[15] H. Y. Yamin and S. M. Shahidehpour, "Risk and profit in self-scheduling for GenCos," *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 2104–2106, Nov 2004. 1

[16] R. Ferreira, L. Barroso, M. Carvalho, and M. Pereira, "A Mean-Variance Optimization Approach to the Development of Portfolios of Renewable Generation in Transmission-Constrained Systems," in *Proc. 2014 PSCC*, 2013, pp. 1–7. 1

[17] A. Schellenberg, W. Rosehart, and J. Aguado, "Cumulant based stochastic optimal power flow (s-opf) for variance optimization," in *IEEE Power Engineering Society General Meeting, 2005*, June 2005, pp. 473–478 Vol. 1. 1

[18] D. Bienstock and A. Shukla, "Variance-Aware Optimal Power Flow," in *Proc. 2018 PSCC*, 2018, pp. 1–8. 1, 2, 3, 4

[19] L. Roald, T. Krause, and G. Andersson, "Integrated balancing and congestion management under forecast uncertainty," in *2016 IEEE International Energy Conference (ENERGYCON)*, April 2016, pp. 1–6. 2, 3

[20] D. Bienstock, G. Munoz, C. Matke, and S. Yang, "Robust linear control of storage in transmission systems, and extensions to robust network control problems," 2017. 2

[21] G. C. Calafiore and L. El Ghaoui, "On distributionally robust chance-constrained linear programs," *Journal of Optimization Theory and Applications*, vol. 130, no. 1, pp. 1–22, 2006. 2

[22] T. Mülpfordt, T. Faulwasser, L. Roald, and V. Hagenmeyer, "Solving Optimal Power Flow with non-Gaussian Uncertainties via Polynomial Chaos Expansion," in *CDC, 2017*. IEEE, 2017. 3

[23] H. Markowitz, "Portfolio Selection," *The Journal of Finance*, vol. 7, pp. 77–91, 1952. 4

[24] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York: Springer, 2006. 5

[25] A. Wächter and L. T. Biegler, "On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming," *Mathematical Programming*, vol. 106, pp. 25 – 57, 2006. 5

[26] R. Byrd, J. Nocedal, and R. Waltz, "KNITRO: An integrated package for nonlinear optimization," in *Large-Scale Nonlinear Optimization*, G. di Pillo and M. Roma, Eds. Springer, 2006, pp. 35–59. 5

[27] L. Roald and D. Molzahn, "Implied constraint satisfaction in power systems optimization: The impacts of load variations," *arXiv:1903.11337v2*, 2019. 5

[28] Y. Nesterov, "Smooth Minimization of Non-smooth Functions," *Math. Program.*, vol. 103, no. 1, pp. 127–152, May 2005. [Online]. Available: http://dx.doi.org/10.1007/s10107-004-0552-5 6

[29] R. D. Zimmerman, C. E. Murillo-Sánchez, and D. Gan, "MATPOWER, A MATLAB Power System Simulation Package," *IEEE Trans. Power Sys.*, vol. 26, no. 1, pp. 12–19, 2011. 7

[30] D. Goldfarb and G. Iyengar, "Robust Portfolio Selection Problems," *Mathematics of Operations Research*, vol. 28, pp. 1–38, 2001. 9

[31] W. F. Sharpe, "The Sharpe Ratio," *The Journal of Portfolio Management*, vol. 21, p. 49–58, 1994. 9

[32] F. Alizadeh and D. Goldfarb, "Second-Order Cone Programming," *Mathematical Programming*, vol. 95, pp. 3–51, 2001. 10

[33] S. L. Digabel, "Algorithm 909: NOMAD: Nonlinear Optimization with the MADS algorithm," *ACM Transactions on Mathematical Software*, vol. 37, p. 44:144:15, 2011. 11

[34] IBM, "ILOG CPLEX Optimizer," http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/. 7

[35] GUROBI, "Gurobi Optimizer," http://www.gurobi.com/. 7