Importance Sampling Approach to Chance-Constrained DC Optimal Power Flow

Aleksander Lukashevich, Vyacheslav Gorchakov, Petr Vorobev Deepjyoti Deka, Yury Maximov

arXiv:2111.11729v1 [math.OC] 23 Nov 2021

***Abstract*—Despite significant economic and ecological effects, a higher level of renewable energy generation leads to increased uncertainty and variability in power injections, thus compromising grid reliability. In order to improve power grid security, we investigate a joint chance-constrained (CC) direct current (DC) optimal power flow (OPF) problem. The problem aims to find economically optimal power generation while guaranteeing that all power generation, line flows, and voltages simultaneously remain within their bounds with a pre-defined probability. Unfortunately, the problem is computationally intractable even if the distribution of renewables fluctuations is specified. Moreover, existing approximate solutions to the joint CC OPF problem are overly conservative, and therefore have less value for the operational practice. This paper proposes an importance sampling approach to the CC DC OPF problem, which yields better complexity and accuracy than current state-of-the-art methods. The algorithm efficiently reduces the number of scenarios by generating and using only the most important of them, thus enabling real-time solutions for test cases with up to several hundred buses.**

***Index Terms*—robust optimization, chance-constrained optimization, optimal power flow**

I. Introduction

In 2020 electricity produced approximately 25% of greenhouse gas emissions in the USA, and integration of a higher volume of renewable energy generation is seen as the primary tool to reduce the emission level. In turn, a higher amount of renewable generation increases the power grid uncertainty, compromises its security, and challenges classical power grid operation and planning policies.

The optimal power flow (OPF) problem, which determines the economically optimal operating level of power generation under given power balance equations and security constraints, is one of the most fundamental problems in grid operation and planning. Several extensions are proposed for the optimal power flow problem for addressing the uncertainty of power generation and consumption. Robust and chance-constrained power flow formulations are among the most popular ones. The robust OPF problem assumes bounded uncertainty and requires a solution to be feasible against any possible uncertainty realization within a given uncertainty set [1], [2], [3].

A more flexible chance-constrained approach requires security constraints to be satisfied a high probability while assuming the distribution of renewables is known in whole or in part [4], [5], [6], [7]. This paper considers a *joint chance-constrained* optimal power flow problem, where the joint probability of at least one failure of the security constraints (line load limits, voltage stability bounds) is bounded from above by a confidence threshold (JCC-OPF).

In contrast to a *single chance-constrained* formulation (SCC-OPF), which imposes individual failure probability thresholds for each of the constrains, the joint chance constraint is computationally hard even for (linear) direct current (DC) power balance equations, linear security limits and Gaussian uncertainty [8], [9]. Several tractable convex approximations have been proposed [10], [11], [12], [13] for joint chance-constrained optimization to overcome the computational hardness of the problem; however, they often lead to conservative solutions inapplicable for operational practice.

Other approaches such as scenario and sample average approximations [14], [15], [16], [11] consist of substituting the stochastic part with a set of deterministic inequalities based on the uncertainty realization. This approach can be distributionally robust and allow exploiting uncertainties beyond the Gaussian ones. A combination of an analytical approximation and sampling [17] can further improve the accuracy of the solution. However it may require a large number of samples. Scenario curation/modification heuristics have been designed to improve the sample complexity of JCC-OPF [18], although with formal analysis of the methodology. In work outside power grids, statistical learning has been used to approximate uncertain convex programs [19], [20], [21].

Nevertheless, the scenario approximation approach remains the most accurate among algorithms for solving the joint chance-constrained DC optimal power flow, although its complexity is often unacceptable for large-scale power grids [22]. To this end, the paper suggests using importance sampling to reduce the complexity and improve the accuracy of the scenario approximation to chance-constrained optimal power flow. Importance sampling allows generating more informative samples and results in an optimization problem with fewer constraints.

The contribution of the paper is as follows.

1) we propose a novel computationally efficient approach to the joint chance-constrained DC OPF problem. The

algorithm exploits physics-informed importance sampling to refine the classical scenario approximation [14];
2) we prove the algorithm to converge to a sub-optimal solution with a guaranteed accuracy under mild technical assumptions;
3) we demonstrate the proposed algorithm's high time efficiency and accuracy over multiple real and synthetic test cases.

The paper is organized as follows. In Section II we present the joint chance-constrained optimal power flow problem and introduce notation used in the paper. We outline the algorithm and provide its theoretical analysis in Section III. Empirical study and comparison to state-of-the-art methods are given in Section IV. In Section V we conclude with a brief summary and discussion on possible applications of our results.

II. Background and Problem Setup

A. Notation

The direct current (DC) power flow model remains extremely popular yet simple for the analysis because of a linear relation between powers and phase angles that is typical high-voltage power grids.

In the following, we consider a power grid given by a graph $\mathcal{G} = (V, E)$ with a set of nodes/buses V and edges/lines E. Assume n is a number of buses, and m is a number of lines. Let p be a vector of power injections $p = (p_F, p_R, p_S)^\top$, where p_F corresponds to buses with deterministic/fixed (F) power injections, p_R to buses with random (R) injections, and p_S is the injection at the slack bus (S). The power system is balanced, i.e. the sum of all power injections equals to zero, $p_S + \sum_{r\in R} p_r + \sum_{f\in F} p_f = 0$. Let θ be a vector of phase angles. Without a loss of generalization, We assume that the phase angle on the reference slack bus $\theta_S = 0$. Let $B \in \mathbb{R}^{N\times N}$ be an admittance matrix of the system, $p = B\theta$. The components B_{ij} are such that $B_{ij} \neq 0$ if there is a line between nodes i and j, the diagonal elements $B_{ii} = -\sum_{j\neq i} B_{ij}$, i.e. B is a Laplacian matrix. Let $B^\dagger$ be the pseudo-inverse of B, i.e. $\theta = B^\dagger p$.

The DC power flow equations, generation and stability constraints are

$$p = B\theta \tag{1}$$
$$p_i^{\min} \leq p_i \leq p_i^{\max}, i \in V \tag{2}$$
$$|\theta_i - \theta_j| \leq \bar{\theta}_{ij},\ (i,j) \in E \tag{3}$$

Let $A \in \{-1, 0, 1\}^{m\times n}$ be the incidence matrix of grid G, i.e. if nodes i and j are connected by edge k then $A_{ki} = +1$, $A_{jk} = -1$ and all other elements are equal to zero. Then the phase angle constraints in (3) can be represented as $AB^\dagger p \leq \bar{\theta}$, $-AB^\dagger p \leq \bar{\theta}$.

Let $C \in \mathbb{R}^{n\times n}$ be a symmetric matrix such that for all random nodes r, fixed nodes f and the slack node S, $C_{rr} = 1$, $C_{ff} = 1$, $C_{Sr} = -1$, $C_{Sf} = -1$, while all other elements are equal to zero. The Dc OPF constraints are given by the following system of inequalities

$$\underbrace{(AB^\dagger C, -AB^\dagger C, C, -C)^\top}_{W} p \leq \underbrace{(\bar{\theta}, \bar{\theta}, p_{\max}, p_{\min})^\top}_{b}.$$

Let J be a number of constraints, $J = 2m + 2n$, then $W \in \mathbb{R}^{J\times n}$ and $b \in \mathbb{R}^{n\times 1}$. We refer a feasibility polytope as a set $\mathcal{P}$, $\mathcal{P} = \{p : Wp \leq b\} = \{p : \bigcap_{i=1}^{J} \omega_i^\top p \leq b_i\}$.

Finally, we assume that fluctuations of power injections p are Gaussian, $p = x + \xi$, where ξ is a Gaussian uncertainty, $\xi \sim \mathcal{N}(0, \Sigma)$ and x is the deterministic part of power injections.

The paper notation is summarized in Table I. We use lower indices for coordinates of vectors and matrices, lower-case letters for probability density functions (PDFs), and upper-case letters for cumulative distribution functions (CDFs). When it does not lead to confusion, we use $\mathbb{P}$ and $\mathbb{E}$ to denote probability and expectation without explicitly mentioning a distribution.

E	set of lines	$\nu(p)$	nominal distribution PDF
V	set of buses	$\mathcal{V}(p)$	nominal distribution CDF
B	admittance matrix	$\nu(p, x)$	parametric distribution CDF
m	number of lines	$\mathcal{V}(p, x)$	parametric distribution CDF
n	number of buses	$p^{\max}$	generation upper limits
p	power injections	$p^{\min}$	generation lower limits
θ	voltage phases	$\mathbb{P}(\cdot)$, $\mathbb{E}(\cdot)$	probability, expectation
I_n	$n \times n$ identity matrix	ξ	vector of stochastic power fluctuations
J	number of constraints		
$\bar{\theta}_{ij}$	voltage angle limits	$\mathcal{N}(\mu, \Sigma)$	Gaussian distribution
ξ	power injection uncertainty, $p \sim \mathcal{N}(0, \Sigma)$	$\mathcal{P}$	feasibility polytope $\mathcal{P} = \{p : \bigcap_{i\leq J} \omega_i^\top p \leq b_i\}$
p	power injections $p = x + \xi$	x	deterministic part of power injections, $x = \mathbb{E}_\xi p$

TABLE I
Paper notation.

B. Problem Setup

The joint chance-constrained optimal power flow problem is:

$$\min_x \mathbb{E}_{\xi\sim\mathcal{N}(0,\Sigma)} \operatorname{cost}(x, \xi) \tag{4}$$
$$\text{s.t.} : \mathbb{P}_{\xi\sim\mathcal{N}(0,\Sigma)}(x + \xi \in \mathcal{P}) \geq 1 - \eta, 0 < \eta \leq 1/2,$$

where η is a preset confidence parameter, and cost is a cost function convex in x for any realization of ξ. In other words, we assume that power flow balance equations (Eqs. (1)) are satisfied almost surely, and the probability that at least one of the security constraints (Eqs. (2) and (3)) fails is at most η.

Notice, that despite the convexity of the cost function Problem (4) is non-convex as its feasibility set is non-convex for a sufficiently high level of uncertainty.

C. Scenario Approach

Over the last two decades, the scenario approach [11], [14] remains the state-of-the-art method for solving joint chance-constrained optimization. The scenario approach consists in substituting the probabilistic constraints with a larger number

of deterministic ones with each constraint standing for some uncertainty realization:

$$\min_x \frac{1}{N}\sum_{t=1}^{N}\text{cost}(x,\xi_t) \tag{5}$$
$$\text{s.t.}: p_t^{\min} \le x+\xi^t \le p_t^{\max},\ 1\le t\le N$$
$$|\theta_i(\xi^t)-\theta_j(\xi^t)| \le \bar\theta_{ij}, (i,j)\in\mathcal{E},\ 1\le t\le N$$
$$x+\xi^t = B\theta(\xi^t),\ 1\le t\le N$$

where N is a number of scenarios, and $\{\xi^t\}_{i=1}^N$ is a series of uncertainty realization. We assume below that the generation cost, $\text{cost}(x,\xi)$, does not depend on the randomness in power fluctuation, and thus $\text{cost}(x,\xi)=\text{cost}(x)$ for any uncertainty realization ξ. Note that dependence of expected cost on known asymptotic parameters of the distribution (variance, mean) are allowed within this framework. We next discuss our method to solve Problem (5).

III. Algorithm

A. Idea and Sketch

The algorithm consists of several steps: constructing an inner approximation to the feasibility set, generating samples outside of the approximation, and, finally, solving the scenario approximation problem (5) with a new collection of samples.

First, using the fact that *the probability of a union of events is bounded from below by the maximum of individual event probabilities*, we construct a lower bound on the probability of constraint feasibility $\mathbb{P}(x+\xi\in\mathcal{P})$. The latter allows to add a set of constraints, $x\in\mathcal{P}_m$, so that if $x\notin\mathcal{P}_m$ then $\mathbb{P}(x+\xi\in\mathcal{P})>\eta$. In other words, if the solution $\bar x$ of the scenario approximation (5) with samples from the nominal distribution $\mathcal{N}(0,\Sigma)$ satisfies $\mathbb{P}(x+\xi\in\mathcal{P})\ge 1-\eta$, then adding additional inequalities $x\in\mathcal{P}_m$ does not change $\bar x$.

Second, using the aforementioned bound, we design a polytope $\mathcal{P}_{in}=\{p: W_{in}\xi\le b_{in}, p=x+\xi\}$ around x with derived W_{in} and b_{in} independent of x. Figure 1 illustrates the idea. Then, we show that for any sample ξ^t and feasible x, if $p=x+\xi^t\in\mathcal{P}_{in}$, then p also necessarily belongs to the constraint feasibility set $\mathcal{P}$. To this end scenarios $\xi^t: x+\xi^t\in\mathcal{P}_{in}$ can be eliminated from the optimization problem (see Eq. (5)) without impacting the approximation accuracy.

Finally, we sample scenarios outside of the polytope $\mathcal{P}_{in}$ using the state-of-the-art importance sampling methods [23], [24] and solve the Scenario Approximation Problem (5) with the collection of samples generated from importance distribution. Later in this section, we provide rigorous proof of the algorithm efficiency and justify its empirical performance in Section IV.

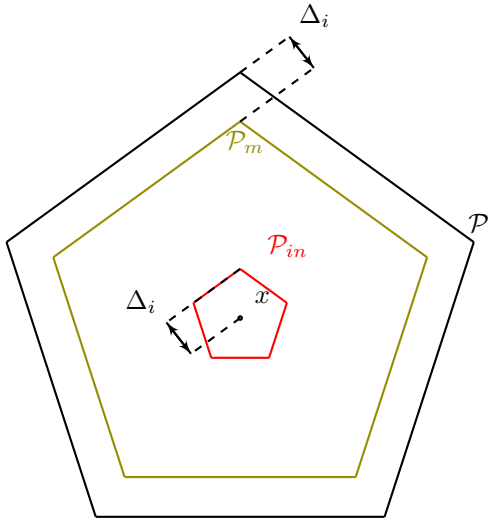


Fig. 1. We consider operating point x and derive a polytope $\mathcal{P}_m$ so that no optimal solution of Prob. (4) belongs to $\mathcal{P}\setminus\mathcal{P}_m$. As it is shown on the figure, $\mathcal{P}_m$ induces a polytope $\mathcal{P}_{in}$ so that none of the power generations $p\in\mathcal{P}_{in}$ affect the scenario approximation solution.

B. Inner Approximation

Consider a probability for the power generation p of being inside the feasibility polytope, $p\in\mathcal{P}$:

$$\mathbb{P}(p\in\mathcal{P})=\mathbb{P}(p: Wp\le b)=$$
$$\mathbb{P}\left(p:\bigcap_{i=1}^{J}\omega_i^\top p\le b_i\right)=1-\mathbb{P}\left(p:\bigcup_{i=1}^{J}\omega_i^\top p> b_i\right)\le$$
$$1-\max_{1\le i\le J}\mathbb{P}\left(p:\omega_i^\top p>b_i\right).$$

So, if there exists x such that for some i, $\mathbb{P}(\omega_i^\top p>b_i)>\eta$, then $\mathbb{P}(p\in\mathcal{P})<1-\eta$. Thus, to satisfy the joint chance constraint for $p=x+\xi$, $\xi\sim\mathcal{N}(0,\Sigma)$, we need

$$\eta\ \ge\mathbb{P}(\omega_i^\top p>b_i)=\mathbb{P}(\omega_i^\top\xi+\omega_i^\top x>b_i)=$$
$$\mathbb{P}\left(\frac{\omega_i^\top\xi}{\|\Sigma^{1/2}\omega\|_2}>\frac{b_i-\omega_i^\top x}{\|\Sigma^{1/2}\omega\|_2}\right)=$$
$$\mathbb{P}\left(\zeta>\frac{b_i-\omega_i^\top x}{\|\Sigma^{1/2}\omega\|_2}\right)=\Phi\left(\frac{\omega_i^\top x-b_i}{\|\Sigma^{1/2}\omega\|_2}\right), \tag{6}$$

where $\zeta\sim\mathcal{N}(0,1)$, Φ is a CDF of the standard normal distribution. Notice, that the function $\Phi(\cdot)$ is convex as soon as its argument is negative, i.e., Ineq. (6) is convex as soon as $x\in\mathcal{P}$.

A set of inequalities $\eta\ge\Phi\left((\omega_i^\top x-b_i)/(\|\Sigma^{1/2}\omega_i\|_2)\right)$, $1\le i\le J$, defines a polytope $\mathcal{P}_m$ as follows:

$$\mathcal{P}_m=\left\{x:\omega_i^\top x\le b_i-\Delta_i\right\}, \tag{7}$$

where $\Delta_i=\|\Sigma^{1/2}\omega_i\|_2\Phi^{-1}(1-\eta)$, $\eta\le 1/2$. Figure 1 illustrates mutual arrangement of the polytopes $\mathcal{P}_m\subset\mathcal{P}$. Note that $\mathcal{P}_m$ defines an outer approximation of the non-convex chance-constrained feasibility set, which is itself an inner approximation of the constraint feasiblity set without any uncertainty. Eq. (7) implies Theorem III.1.

Theorem III.1. *The joint chance-constrained optimal power flow problem* (4) *has the same set of optimal solutions* X *as*

$$\begin{aligned} &\min_x \mathbb{E}_{\xi\sim\mathcal{N}(0,\Sigma)} \operatorname{cost}(x,\xi) \\ &\text{s.t.}: \mathbb{P}_{\xi\sim\mathcal{N}(0,\Sigma)}(x+\xi\in\mathcal{P}) \geq 1-\eta, 0<\eta\leq 1/2, \\ &\quad x\in\mathcal{P}_m \end{aligned} \tag{8}$$

Proof. Equations $x \in \mathcal{P}_m$ does not affect the solution of Problem (4) as the feasiblity set of the chance-constrained optimization problem exists inside $\mathcal{P}_m$. □

C. Redundant Scenarios

Another useful consequence of the fact that the optimal solution of the chance-constrained optimal power flow problem is well separated from the boundary of the polytope $\mathcal{P}$ is that certain scenarios may be removed as they do not improve the accuracy of scenario approximation (5).

Let the optimal solution $\bar{x}$ of the problem (5) be feasible for the chance-constrained OPF problem. Then by Theorem III.1, it necessarily also belongs to $\mathcal{P}_m$. Mathematically,

Theorem III.2. *Any solution of the Scenario approximation problem* (5) *that is feasible for the Chance-constrained OPF problem* (4) *is also a solution of*

$$\min_x \ \operatorname{cost}(x)$$

$$\text{s.t.}: \ p_t^{\min} \leq x+\xi^t \leq p_t^{\max}, \ 1\leq t\leq N \tag{9a}$$

$$|\theta_i(\xi^t)-\theta_j(\xi^t)| \leq \bar{\theta}_{ij}, (i,j)\in\mathcal{E}, \ 1\leq t\leq N \tag{9b}$$

$$x+\xi^t = B\theta(\xi^t), \ 1\leq t\leq N \tag{9c}$$

$$x\in\mathcal{P}_m \tag{9d}$$

Let $\mathcal{P}_{in} = \{\xi : \omega_i^\top \xi \leq \Delta_i\}$ with $\Delta_i = \|\Sigma^{1/2}\omega_i\|\Phi^{-1}(1-\eta)$. Theorem III.2 implies that for any $x \in \mathcal{P}_m$, if $\xi^t \in \mathcal{P}_{in}$, then automatically $p = x + \xi^t \in \mathcal{P}$. In other words, one can exclude scenario $\xi \in \mathcal{P}_{in}$ from Problem (5) as soon as $p = x + \xi \in \mathcal{P}$. Figure 1 illustrates the idea and the geometry of $\mathcal{P}, \mathcal{P}_m$, and $\mathcal{P}_{in}$.

Theorem III.3 follows from the result of Calafiore and Campi [14] and establishes approximation properties of a solution of the Problem (9). Assumption 1 is the main technical assumption used in the proof of Theorem III.3.

Assumption 1. *Assume that for all possible uncertainty realizations* $\xi^1, \ldots, \xi^N$*, the optimization problem* (9) *is either infeasible or, if feasible, it attains a unique optimal solution.*

Theorem III.3. *Let* $\bar{x}_N$ *be a unique solution of the Scenario optimization Problem* (9) *with* N *i.i.d. samples, so that none of the samples belong to* $\mathcal{P}_{in}$*. Moreover, assume that for any* N *the assumption 1 is fulfilled. Then for any* $\delta \in (0,1)$ *and any* $\eta \in (0, 1/2]$*,* $\bar{x}_N$ *is also a solution for the Chance-constrained optimal power flow Problem* (4) *with probability at least* $1-\delta$ *if*

$$N \geq \left\lceil 2\frac{(1-\pi)\ln\frac{1}{\delta}}{\eta} + 2d + 2d(1-\pi)\frac{\ln\frac{2(1-\pi)}{\eta}}{\eta} \right\rceil,$$

where d *is a dimension of the space of controllable generators, and* π *is a probability of a random scenario* ξ *to belong to* $\mathcal{P}_{in}$*, and* $\pi < 1$*.*

Proof. First, notice that discarding scenarios $\xi^t \sim \mathcal{N}(0,\Sigma)$ is equivalent in solving the scenario approximation problem with $\xi \sim D \Leftrightarrow \xi \sim \mathcal{N}(0,\Sigma)\,\text{s.t.}: \xi \notin \mathcal{P}_{in}$.

According to the result of Calafiori and Campi [14], for any probability $\delta \in (0,1)$ and any confidence threshold probability ε, and dimension of the space of parameters d one has, for N_1

$$N_1 \geq \left\lceil \frac{2}{\varepsilon}\ln\frac{1}{\delta} + 2d + \frac{2d}{\varepsilon}\ln\frac{2}{\varepsilon} \right\rceil \tag{10}$$

scenarios from D and the optimal solution $\bar{x}$ of the Problem (9), the probability of failure is bounded as

$$\mathbb{P}_D(p \notin \mathcal{P}) \leq \varepsilon$$

with probability at least $1-\delta$.

Notice, that the bounds on the number of samples (see Eq. (10)) is strictly decreasing in ε for $\varepsilon \in (0,1)$. As scenarios in $\mathcal{P}_{in}$ do not cause failure, to get a probability of failure η according to measure $\mathcal{N}(0,\Sigma)$, we need the failure probability according to D to be at least $\varepsilon = \eta/(1-\pi)$. Thus, taking $\varepsilon = \eta/(1-\pi)$ and using monotonicity of Ineq. (10) one gets the statement of the theorem. □

Theorem III.3 establishes the number of scenarios sufficient for the scenario approximation solution being feasible for the chance-constrained optimal power flow problem. This number significantly decreases if one can come up with a sufficiently accurate inner approximation of the feasibility set. Notice, that without an inner approximation, i.e. for $\pi = 0$, one gets the result of [14, Theorem 1].

D. Importance Sampling

Although scenario optimization with scenarios that do not belong to the polytope $\mathcal{P}_{in}$ obey a nice complexity bound, it requires on average $1/(1-\pi)$ or more samples to generate at least one point outside of $\mathcal{P}_{in}$ and decreases the overall efficiency of the approach. The problem is especially challenging when dealing with rare events, i.e., the confidence level $\eta \to 0$.

Importance sampling is a general technique that helps to improve the efficiency of scenario generation [25]. It consists of a change of the probability distribution to sample rare events with a higher probability. Figure 2 illustrates the concept.

Unfortunately, there is no exact and time-efficient algorithm for sampling outside of a convex polytope from Gaussian distribution [9]; however, the ALOE algorithm [23], [24] proposes an elegant way for approximating the distribution of interest by sampling from a mixture of distributions.

We consider Gaussian fluctuations of power injections, $\xi \sim \mathcal{N}(0,\Sigma)$, with known covariance $\Sigma \in \mathbb{R}^{n\times n}$ and aim to sample scenarios outside of $\mathcal{P}_{in}$ so that the probability distribution to sample from is as close as possible to the conditional Gaussian distribution $\xi \sim \mathcal{N}(0,\Sigma)\,\text{s.t.}: \xi \notin \mathcal{P}_{in}$.

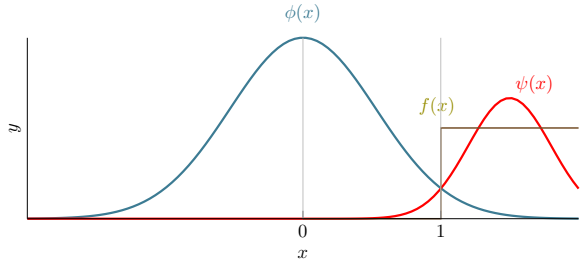


Fig. 2. Our goal is to sample points in the area $\{x : f(x) = 1\}$. Sampling from the nominal distribution $\phi(x)$ is inefficient as one needs to sample from the distribution tail. The probability distribution $\psi(x)$ is a better choice as the probability of getting $f(x) = 1$ is substantially higher when sampling from it.

The method essentially samples from a weighted mixture of conditional Gaussian distributions D_i

$$\xi \sim D_i \iff \xi \sim \mathcal{N}(0, \Sigma)\,\mathrm{s.t.}: \omega_i^\top \xi > \Delta_i.$$

Consider the set of inequalities $\{\omega_i^\top \xi > \Delta_i\}_{i \le J}$ in more detail. First, let $\zeta \sim \mathcal{N}(0, I_n)$ then the system is equivalent to $\{(\Sigma^{1/2}\omega_i)^\top \zeta > \Delta_i\}_{i \le J}$.

Distribution D_i can be simulated exactly using the inverse transform method [26], [27] that admits conditional sampling $\xi \sim \mathcal{N}(0, \Sigma)$ s.t. $\omega_i^\top \xi \ge \Delta_i$.

1) Sample $z \sim \mathcal{N}(0, I)$ and sample $u \sim U(0, 1)$
2) Compute $y = \Phi^{-1}(\Phi(-\Delta_i) + u(1 - \Phi(-\Delta_i)))$
3) Set $\phi = \phi y + (I - \phi\phi^\top)z$, $\phi = \Sigma^{1/2}\omega_i / \|\Sigma^{1/2}\omega_i\|_2$
4) Set $\xi = \Sigma^{1/2}\phi$.

In [23], [24], the authors proposed a slightly refined version of the algorithm above that exhibits better numerical stability. We refer to the same papers for the corresponding proofs and analysis. Figure 3 illustrates the idea.

Finally, ALOE proposes to sample scenarios from a weighted mixture

$$D = \sum_{i=1}^{J} \alpha_i D_i,\ \alpha_i \ge 0, \sum_{i=1}^{J} \alpha_i = 1$$
$$\alpha_i \propto \Phi(-\Delta_i / \|\Sigma^{1/2}\omega_i\|_2), \qquad (11)$$

where Φ is a CDF of the standard normal distribution. Let $q_D(\xi)$ be a PDF of distribution D, then Theorem III.4 established a maximal ratio of the conditional Gaussian density $\xi \sim \mathcal{N}(0, \Sigma)\,\mathrm{s.t.}: \xi \notin \mathcal{P}_{in}$ and $q_D(\xi)$.

Theorem III.4. *Let $p(\xi)$, $q_D(\xi)$ be PDFs of $\xi \sim \mathcal{N}(0, \Sigma)$ $\mathrm{s.t.}: \xi \notin \mathcal{P}_{in}$, and a mixture density (see Eq. (11)) resp. Then for any $\xi \notin \mathcal{P}_{in}$, we have*

$$p(\xi) \le M q_D(\xi), M = \frac{\sum_{i \le J} \Phi(-\Delta_i / \|\Sigma^{1/2}\omega_i\|_2)}{\max_{i \le J} \Phi(-\Delta_i / \|\Sigma^{1/2}\omega_i\|_2)}, \qquad (12)$$

where D and α_i are given in Eq. (11).

Proof. Let $\phi(\xi)$ be PDF of $\xi \sim \mathcal{N}(0, \Sigma)$. Notice, that the conditional densities D_i have probability density functions

$$q_{D_i}(\xi) = \begin{cases} \phi(\xi)/\Phi(-\Delta_i / \|\Sigma^{1/2}\omega_i\|_2), & \omega_i^\top \xi > \Delta_i \\ 0, & \text{otherwise} \end{cases}$$

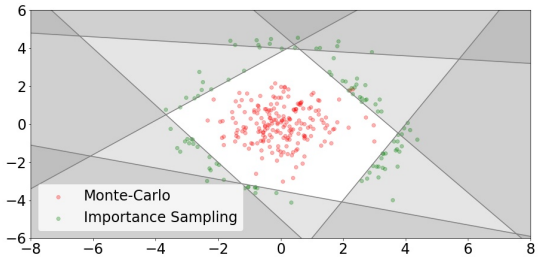


Fig. 3. The white area stands for generations that do not exceed operating limits. A set of generations leading to at least one constraint violation is in grey. Two or more reliability constraints are not satisfied in the dark grey area, also samples from a nominal distribution and the constructed mixture marked in red and green respectively

Thus the density of distribution D is $\sum_{i \le J} \alpha_i q_{D_i}(\xi)$. Similarly, density $p(\xi)$ outside of the polytope $\mathcal{P}_{in}$ is $\phi(\xi)/\mathbb{P}_{\xi \sim \mathcal{N}(0,\Sigma)}(\xi \notin \mathcal{P}_{in})$ which is less or equal then $\phi(\xi)/\max_{i \le J} \Phi(-\Delta_i / \|\Sigma^{1/2}\omega_i\|_2)$ for any $\xi \notin \mathcal{P}_{in}$.

Finally, taking the ratio of $q_D(\xi)$ and $p(\xi)$ and using the value of α_is, we get the lemma statement. □

Theorem III.4 implies that if a probability of a set with respect to measure q_D is less or equal then ε, then the probability of the same set with respect to measure p does not exceed $M\varepsilon$.

E. Scenario Approximation with Importance Sampling

In this section, we present a scenario approximation for the chance-constrained optimal power flow with a set of scenarios generated by the ALOE algorithm [23]. A particular advantage of this approach is that every scenario is generated outside of $\mathcal{P}_{in}$. The latter substantially improves the accuracy and efficiency of the scenario approximation. In particular, we solve the following optimization problem instead of Problem (5):

$$\min_x \ \mathrm{cost}(x)$$
$$\mathrm{s.t.}: \ p_t^{\min} \le x + \xi^t \le p_t^{\max},\ 1 \le t \le N \qquad (13a)$$
$$|\theta_i(\xi^t) - \theta_j(\xi^t)| \le \bar{\theta}_{ij}, (i,j) \in \mathcal{E},\ 1 \le t \le N \qquad (13b)$$
$$x + \xi^t = B\theta(\xi^t),\ 1 \le t \le N \qquad (13c)$$
$$x \in \mathcal{P}_m \qquad (13d)$$
$$\xi^1, \xi^2, \ldots, \xi^N \sim D, \qquad (13e)$$

where D is the probability distribution defined by Eq. (11).

Notice, that sampling from distribution D allows to efficiently generate scenarios outside of the polytope $\mathcal{P}_{in}$.

However, they follow distribution D instead of $\xi \sim \mathcal{N}(0, \Sigma)\,\mathrm{s.t.}: \xi \notin \mathcal{P}_{in}$. As these distributions are sufficiently close to each other, Theorem III.5 establishes efficient complexity bounds for the scenario approximation with importance sampling.

Theorem III.5. *Let $\bar{x}_N$ be a unique solution of the Scenario optimization Problem (13) with N i.i.d. samples follow distribution D. Moreover, assume that for any N the assumption 1*

is fulfilled. Then for any $\delta \in (0,1)$ and any $\eta \in (0, 1/2]$, $\bar{x}_N$ is also a solution for the Chance-constrained optimal power flow Problem (4) *with probability at least $1-\delta$ if*

$$N \geq \left\lceil 2M\frac{(1-\pi)\ln\frac{1}{\delta}}{\eta} + 2d + 2dM(1-\pi)\frac{\ln\frac{2M(1-\pi)}{\eta}}{\eta} \right\rceil,$$

where d is a dimension of the problem and π is a probability of a random scenario ξ to belong to $\mathcal{P}_{in}$, $\pi < 1$, and constant M is defined by Theorem III.4.

Proof. The proof is similar to the one of Theorem III.3. Application of theorem III.4 allows to upper-bound the probability of an event in measure D with respect to its probability in measure $\mathcal{N}(0,\Sigma)\,\mathrm{s.t.}: \xi \notin \mathcal{P}_{in}$. □

IV. Empirical Study

We compare the performance of the scenario approximation approaches with and without importance sampling over real and simulated test cases with dimensions varying from several dozens to hundreds of variables. We limit the empirical setting to considering Gaussian distributions and linear feasibility constraints only.

A. Compared Algorithms

In this study, we have compared the classical scenario approximation [14] with the importance sampling-based scenario approximation, where the samples are generated with ALOE algorithm [23] from the outside of the polytope $\mathcal{P}_{in}$ that contains obsolete samples only (see Section III-C for details). We omit detailed comparison with other importance sampling strategies [28], [24], [29] when generating scenarios because of the paper space limitation and for the sake of empirical study clarity.

B. Implementation details

We have used Julia 1.5.3. and PowerModels.jl [30] on MacBook Pro (2.4GHz, 8-Core Intel Core i9, 64 GB RAM). In the experiments, the computational time for each case has not exceeded 10 minutes, which makes the solution applicable for the operational practice. We will make our code publicly available for research purposes.

When solving scenario approximation optimization problems, we use CVX [31] and GLPK optimization solver [32].

C. Test Cases and Numerical Results

In our experiments we use both synthetic and realistic power flow IEEE test cases [33].

a) Synthetic Example: We first studied efficiency of our algorithm on one dimensional test case:

$$\begin{aligned} &\max\ x \\ &\mathrm{s.t.}: \mathbb{P}_\xi(x+\xi \leq a) \geq 1-\eta, \xi \sim \mathcal{N}(0,1) \end{aligned}$$

for $0 < \eta < 1/2$ and a positive parameter a. In this case, the chance-constrained optimization problem admits an exact solution, $x^* = a - \Phi^{-1}(1-\eta)$, the polytopes $\mathcal{P}_m$ and $\mathcal{P}_{in}$ are $\{x : x \leq x^*\}$ and $\xi : \xi \leq \Phi^{-1}(1-\eta)$ respectively. To illustrate the role of an inner approximations $\mathcal{P}_m$ and $\mathcal{P}_{in}$ we consider different polytopes $\mathcal{P}_m = \{x : x \leq b\}$. The latter affects the efficiency of sampling. Figure 4 illustrates the evolution of scenario approximation without importance with the number of samples and the size of the polytope $\mathcal{P}_m$.

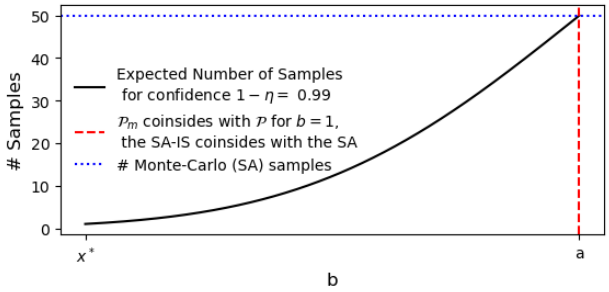


Fig. 4. The figure demonstrates feasibility of the scenario approximation with importance sampling depending on the size of $\mathcal{P}_m = \{x : x \leq b\}$, $b \leq a - \Phi^{-1}(1-\eta)$. The more accurate approximation $\mathcal{P}_m$ (and $\mathcal{P}_{in}$ resp.) is the less number of samples required by the SA-IS algorithm.

Although, one can consider an arbitrary polytope $\bar{\mathcal{P}} \subset \mathcal{P}_{in}$ which then will satisfy the conditions of Theorem III.5. Figure 4 illustrates the efficiency of the importance sampling approximation for more and less conservative inner polyhedral approximations. Notice that a less conservative approximation leads to a less accurate solution we get. To this end, deriving a non-conservative inner approximation is crucial for the importance sampling approach success.

b) Power grid test cases: We address the chance-constrained DC optimal power flow problem under Gaussian fluctuations of renewables by comparing the algorithms' efficiency in various test cases with up to several hundred buses. We have used DC power flow cases accessible in Matpower [34] and PowerModels.jl [30].

Table II summarized the empirical efficiency of the classical scenario approximation and the importance of sampling-based scenario approximation algorithms. For all consider cases (IEEE 30, IEEE 57, and IEEE 118), we assumed the power generation and consumption level to fluctuate with the standard deviation 0.07 of its nominal value. We refer to SA and SA-IS as scenario approximations with and without importance sampling, respectively.

Our experiments show that with a moderate number of samples the SA approach comes up with a lower cost solution; however, it does not meet security guarantees in out-of sample testing. So the SA method requires a much higher number of samples to deliver constraint satisfaction with the required level of confidence. In contrast, the SA-IS requires much less samples to meet the security constraints with a required probability (validated in out of sample guarantees), with the improvement being significant for higher confidence levels $1-\eta$. Furthermore this incures a very minimal increase in cost, making our approach useful for solving problems with rare events and stricter guarantees.

V. Conclusion

In this paper, we investigated the scenario approximation for the chance-constrained optimal power flow. We showed

Case	Conf. $1-\eta$	DC-OPF cost	SA cost	SA-IS cost	SA conf.	SA-IS conf.
IEEE 30	0.95	5669	5712	5735	0.86	**0.96**
IEEE 30	0.99	5669	5712	5760	0.86	**0.99**
IEEE 30	0.995	5669	5712	5780	0.86	**1.00**
IEEE 57	0.95	25016	25044	25095	0.84	**0.97**
IEEE 57	0.99	25016	25044	25110	0.84	**1.00**
IEEE 57	0.995	25016	25044	25165	0.84	**1.00**
IEEE 118	0.95	84840	85607	86018	0.47	**0.96**
IEEE 118	0.99	84840	85607	86230	0.47	**0.99**
IEEE 118	0.995	84840	85607	86321	0.47	**1.00**

TABLE II

EMPIRICAL COMPARISON OF THE SCENARIO APPROXIMATION APPROACHES WITH AND WITHOUT IMPORTANCE SAMPLING. IN ALL CASES WE USED **600** SCENARIOS AND SOLVED THE PROBLEM 50 TIMES OVER INDEPENDENT SAMPLES. SA AND SA-AS CONF. STAND FOR THE AVERAGE CONFIDENCE WITH OUT OF SAMPLE TEST (1000 SAMPLES FROM THE NOMINAL UNCERTAINTY DISTRIBUTION).

that the importance sampling technique used for scenario generation leads to better accuracy and time complexity in theory and practice. Finally, the approach can be extended to automated real-time control of bulk power systems.

REFERENCES

[1] A. Ben-Tal and A. Nemirovski, "Robust optimization–methodology and applications," *Mathematical programming*, vol. 92, no. 3, pp. 453–480, 2002.

[2] T. Ding, Z. Bie, L. Bai, and F. Li, "Adjustable robust optimal power flow with the price of robustness for large-scale power systems," *IET Generation, Transmission & Distribution*, vol. 10, no. 1, pp. 164–174, 2016.

[3] A. A. Sousa, G. L. Torres, and C. A. Canizares, "Robust optimal power flow solution using trust region and interior-point methods," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 487–499, 2010.

[4] M. Lubin, Y. Dvorkin, and S. Backhaus, "A robust approach to chance constrained optimal power flow with renewable generation," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3840–3849, 2015.

[5] L. Roald and G. Andersson, "Chance-constrained ac optimal power flow: Reformulations and efficient algorithms," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2906–2918, 2017.

[6] D. Bienstock, M. Chertkov, and S. Harnett, "Chance-constrained optimal power flow: Risk-aware network control under uncertainty," *Siam Review*, vol. 56, no. 3, pp. 461–495, 2014.

[7] A. Pena-Ordieres, D. K. Molzahn, L. A. Roald, and A. Wächter, "Dc optimal power flow with joint chance constraints," *IEEE Transactions on Power Systems*, vol. 36, no. 1, pp. 147–158, 2020.

[8] B. Cousins and S. Vempala, "A cubic algorithm for computing gaussian volume," in *Proceedings of the twenty-fifth annual ACM-SIAM symposium on discrete algorithms*. SIAM, 2014, pp. 1215–1228.

[9] L. Khachiyan, "Complexity of polytope volume computation," in *New trends in discrete and computational geometry*. Springer, 1993, pp. 91–101.

[10] A. Nemirovski and A. Shapiro, "Convex approximations of chance constrained programs," *SIAM Journal on Optimization*, vol. 17, no. 4, pp. 969–996, 2007.

[11] ——, "Scenario approximations of chance constraints," in *Probabilistic and randomized methods for design under uncertainty*. Springer, 2006, pp. 3–47.

[12] A. Nemirovski, "On tractable approximations of randomly perturbed convex constraints," in *42nd IEEE International Conference on Decision and Control (IEEE Cat. No. 03CH37475)*, vol. 3. IEEE, 2003, pp. 2419–2422.

[13] A. Trofino and C. E. De Souza, "Bi-quadratic stability of uncertain linear systems," in *Proceedings of the 38th IEEE Conference on Decision and Control (Cat. No. 99CH36304)*, vol. 5. IEEE, 1999, pp. 5016–5021.

[14] G. C. Calafiore and M. C. Campi, "The scenario approach to robust control design," *IEEE Transactions on automatic control*, vol. 51, no. 5, pp. 742–753, 2006.

[15] S. Garatti and M. Campi, "Risk and complexity in scenario optimization," *Mathematical Programming*, pp. 1–37, 2019.

[16] 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*, vol. 28, no. 4, pp. 3885–3896, 2013.

[17] A. M. Hou and L. A. Roald, "Chance constraint tuning for optimal power flow," in *2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*. IEEE, 2020, pp. 1–6.

[18] I. Mezghani, S. Misra, and D. Deka, "Stochastic ac optimal power flow: A data-driven approach," *Electric Power Systems Research*, vol. 189, p. 106567, 2020.

[19] V. N. Vapnik, "An overview of statistical learning theory," *IEEE transactions on neural networks*, vol. 10, no. 5, pp. 988–999, 1999.

[20] Y. Maximov and D. Reshetova, "Tight risk bounds for multi-class margin classifiers," *Pattern Recognition and Image Analysis*, vol. 26, no. 4, pp. 673–680, 2016.

[21] M. C. Campi and S. Garatti, "Scenario optimization with relaxation: a new tool for design and application to machine learning problems," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 2463–2468.

[22] N. N. Sakhavand, "New algorithms for stochastic power systems planning and operations problems," Ph.D. dissertation, University of Texas Arlington, 2020.

[23] A. B. Owen, Y. Maximov, M. Chertkov *et al.*, "Importance sampling the union of rare events with an application to power systems analysis," *Electronic Journal of Statistics*, vol. 13, no. 1, pp. 231–254, 2019.

[24] A. Lukashevich and Y. Maximov, "Power grid reliability estimation via adaptive importance sampling," *IEEE Control Systems Letters*, 2021.

[25] S. T. Tokdar and R. E. Kass, "Importance sampling: a review," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 1, pp. 54–60, 2010.

[26] P. L'Ecuyer and A. B. Owen, *Monte Carlo and Quasi-Monte Carlo Methods 2008*. Springer, 2009.

[27] J. Morlet, "Sampling theory and wave propagation," in *Issues in acoustic Signal processing and recognition*. Springer, 1983, pp. 233–261.

[28] A. Genz, F. Bretz, T. Miwa, X. Mi, F. Leisch, F. Scheipl, B. Bornkamp, M. Maechler, T. Hothorn, and M. T. Hothorn, "Package 'mvtnorm'," *J. of Computational and Graphical Statistics*, vol. 11, pp. 950–971, 2020.

[29] M. F. Bugallo, V. Elvira, L. Martino, D. Luengo, J. Miguez, and P. M. Djuric, "Adaptive importance sampling: The past, the present, and the future," *IEEE Signal Processing Magazine*, vol. 34, no. 4, pp. 60–79, 2017.

[30] C. Coffrin, R. Bent, K. Sundar, Y. Ng, and M. Lubin, "Powermodels. jl: An open-source framework for exploring power flow formulations," in *2018 Power Systems Computation Conference*. IEEE, 2018, pp. 1–8.

[31] S. Diamond and S. Boyd, "CVXPY: A Python-embedded modeling language for convex optimization," *Journal of Machine Learning Research*, vol. 17, no. 83, pp. 1–5, 2016.

[32] A. Makhorin, "Glpk (gnu linear programming kit)," Available at `http://www.gnu.org/software/glpk/glpk.html`.

[33] L. Thurner, A. Scheidler, F. Schäfer, J.-H. Menke, J. Dollichon, F. Meier, S. Meinecke, and M. Braun, "pandapower—an open-source python tool for convenient modeling, analysis, and optimization of electric power systems," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6510–6521, 2018.

[34] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on power systems*, vol. 26, no. 1, pp. 12–19, 2010.