Proceedings of the 2024 Winter Simulation Conference
H. Lam, E. Azar, D. Batur, S. Gao, W. Xie, S. R. Hunter, and M. D. Rossetti, eds.

IMPORTANCE SAMPLING OF RARE EVENTS FOR DISTRIBUTION NETWORKS WITH STOCHASTIC LOADS

Mark Christianen[1], Henry Lam[2], Maria Vlasiou[3], and Bert Zwart[1,4]

[1]Dept. of Math. and Comp. Sci., Eindhoven University of Technology, Eindhoven, THE NETHERLANDS
[2]Dept. of Industrial Eng. and Operations Research, Columbia University, New York, NY, USA
[3]Dept. of Electrical Eng., Math. and Comp. Sci., University of Twente, Enschede, THE NETHERLANDS
[4]Stochastics, CWI, Amsterdam, THE NETHERLANDS

ABSTRACT

Distribution networks are low-voltage electricity grids at the neighborhood level. Within these networks, failures can occur as rare events, triggered by stochastic loads that push voltage levels beyond safe limits. To assess the resilience and reliability of these networks, estimating voltage exceedance probabilities is therefore important. We develop importance-sampling strategies to estimate failure probabilities. We do so using two components. First, we propose a change of measure, using either the *Large Deviations Principle* and linear power flow equations or the *Cross-Entropy* method to improve sampling efficiency. Second, we determine feasibility of loads by using our previously developed duality method to overcome the computational complexity of directly solving nonlinear power flow equations using methods such as *Newton-Raphson* and *backward-forward sweep* algorithms. Experiments on a IEEE-15 bus network show that this methodology offers a fast and accurate estimation of failure probabilities in distribution networks.

1 INTRODUCTION

In estimating probabilities of rare events, the rare event itself happens infrequently in simulation experiments when using crude Monte Carlo (MC) simulation. This results in estimation inefficiency and motivates the use of variance-reduction techniques. Among them, importance sampling (IS) (Tokdar and Kass 2010) is one of the most popular methods. The idea behind IS involves introducing a change of measure to assign more weight to critical events. This alternative measure is then utilized to obtain simulation samples, and the outputs are re-weighted using the likelihood ratio, ensuring the unbiasedness of the estimator.

For distribution networks, the occurrence of rare events plays a significant role in the overall reliability and resilience of the network. These rare events correspond to infeasible loads (consisting of supply and demand for power of households) that fall outside the network's normal operating bounds, contrasting with feasible loads that comply with safety and operational standards, maintaining acceptable voltage levels and power flows. These events can occur both through excessive demands (e.g. through charging of electric vehicles) as well as excessive supply of electricity (e.g. through the production of excessive power through rooftop solar panels). If not properly managed, these events, not only disrupt the immediate power supply through a blackout but also include broader consequences including economic costs, compromised public services, and potential damage to critical infrastructure (Cadini et al. 2017).

In this paper, we develop IS strategies to estimate failure probabilities in distribution networks using two components: an efficient change of measure and an accurate feasibility check on whether the load violates the safety standards. We consider an indexed family of rare events $\{\mathscr{A}_\gamma\}_\gamma$ that represents a set of loads that lead to undesirable voltage levels (either too high or too low) within the network. The parameter γ serves as a "rarity parameter", and as $\gamma \to \infty$, the event $\mathscr{A}_\gamma$ becomes rarer so that the probability of the event $\mathscr{A}_\gamma$ tends to zero, i.e., $\mathbb{P}(\mathscr{A}_\gamma) \to 0$. Initially, the network's safe operating point is chosen to ensure

safe voltage levels. However, the presence of noise (modeled as Gaussian noise) can push the operating point into the direction of the unsafe loads in $\{\mathscr{A}_\gamma\}_\gamma$.

To apply a change of measure, we view the loads as a random vector with a mean equal to the safe operating point and shift the mean of this random vector, based on insights by the *Large Deviation Principle* (LDP) and linear approximations of the set of infeasible loads or the *Cross-Entropy* (CE) method. Other approaches to find efficient change of measures are closely related to instanton methods (Chertkov et al. 2011; Chertkov et al. 2011) and dominating points in large deviation techniques (Verseveldt 2023; Nesti et al. 2019) in power systems engineering. The main difference between these approaches and our work lies in their reliance on the simplified linear direct current (DC) model of the electricity network, instead of the more complex alternating current (AC) model (Molzahn and Hiskens 2019). The widely used DC model serves as a practical trade-off for computational tractability. However, the DC model assumes constant voltages magnitudes in the network, making it inadequate for our work as it fails to account for undesirable high or low voltages, i.e., voltage drop constraints. Instead, we employ the *Linearized Distflow* (LD) model and a polyhedral restriction of the feasible region based on the *Bus Injection model* (Christianen et al. 2023) to propose a shift in measure. Both models, the LD model and the polyhedral restriction, are able to deal with voltage drop constraints and represent ways to express the infeasible region through a union of events defined by linear constraints. The LD model offers an approximation, while the polyhedral restriction serves an an actual restriction of the feasible region. The approach using the LDP and linear approximations of the set of infeasible loads comes close to the work in Owen et al. (2019) and Vasmel (2019). However, there are two main differences. First, the authors consider a mixture IS strategy to estimate the probability of a union of events in electricity networks. Here, each component of the mixture corresponds to a shift in the Gaussian distribution's mean (used to model noise) to increase the likelihood of a failure. This strategy essentially combines multiple dominating points to increase estimation efficiency. While it is known that IS using all dominating points is asymptotically efficient, the results in Bai et al. (2023) highlight that using only the most significant dominating point – thus avoiding a mixture strategy – does not result in a poor experimental performance, an approach we adopt in our work. Second, Owen et al. (2019) and Vasmel (2019) use the linear DC model of the electricity network to propose a change of measure. The other approach, using the CE method (de Boer et al. 2005), has been applied to electricity networks. However, to the best of our knowledge, there are no studies on undesirable low or high voltages in the network, as considered in our work.

After a change of measure, we sample loads, which consist of consumer supply and demand of power. In AC models, power has both an active and reactive component; we consider both in our study (Molzahn and Hiskens 2019). Then, it becomes necessary to check whether a sample of loads from this new distribution is feasible. In other words, we must verify if a sample satisfies some power flow model equations. In our work, while the shift in measure is based on an approximation and restriction of the feasible region, the feasibility check of loads is based on the full feasible region defined by the Bus Injection Model, a popular nonlinear power flow model (Molzahn and Hiskens 2019). This is in contrast to the work done in Owen et al. (2019), Vasmel (2019), Chertkov et al. (2011), Verseveldt (2023), and Nesti et al. (2019) where the feasibility check is done using the linear DC model of the electricity network. When nonlinear power flow models are used, the Newton-Raphson (NR) method and traditional backward-forward sweep algorithms are typically used (Montoya et al. 2021). Our proposed duality method is specifically designed for distribution networks (Christianen et al. 2024). This method, a new backward-forward sweep-based algorithm, is different from existing methods since it does not require the computation of actual voltages in the distribution network, offering an improvement in speed over both the NR method and traditional backward-forward sweep-based algorithms, which are used for solving nonlinear power flow equations, and thus a valuable alternative for feasibility checks in distribution networks.

The insights of this paper can be summarized as follows. In terms of constructing a change of measure, our results reveal that using the LDP with our polyhedral restriction tend to produce overly conservative change of measures, leading to probability estimates of zero. Using either the LDP with the LD model or

the CE method yield refined estimates with correct magnitudes, with the former offering a faster approach to estimating rare-event probabilities compared to the latter. For the feasibility check, we show that an accurate estimation of rare-event probabilities via MC simulations necessitates the use of nonlinear power flow models in the feasibility check. In other words, using linear approximations or restrictions of the feasible region lead to rare-event probabilities that are orders of magnitude away from the true values. Additionally, to check whether the sampled loads satisfy some nonlinear power flow equations, the duality method is a faster and accurate alternative over the NR method for distribution networks that satisfy several assumptions, i.e., those with a high resistance-over-reactance ratio and predominance of active over reactive power in the network (Khatod et al. 2006; Tonso et al. 2005). Conversely, the NR method is also suitable for networks that do not follow these assumptions. The results of this paper are of potential relevance to the safe operation of distribution networks (Jang et al. 2024).

This paper is organized as follows. In Section 2, we describe the modeling of distribution networks, the power flow models and corresponding feasible regions that we use. Section 3 defines the problem. Sections 4 and 5 are the core of this paper: we design our IS strategies, and provide numerical experiments to test the validity of our approach on a IEEE 15-bus network. The conclusions can be found in Section 6.

2 DISTRIBUTION NETWORKS AND POWER FLOW MODELS

The distribution network is modeled as a rooted tree $\mathscr{G} = (\mathscr{V}, \mathscr{E})$, with nodes $\mathscr{V} = \{0\} \cup \{1, 2, \ldots, d\}$, where node 0 represents the generator and the rest represent load nodes. Edges $(i,k) \in \mathscr{E}$ are characterized by their impedance $z_{ik} = r_{ik} + \mathrm{i}x_{ik}$, incorporating the network's resistance (r_{ik}) and reactance (x_{ik}). Thus, for the network, we have in matrix form $\mathbf{Z} = \mathbf{R} + \mathrm{i}\mathbf{X}$ representing the *impedance bus matrix* (Peterson et al. 1989). We denote the generator's complex voltage, current, and power as V_0, I_0, and s_0, respectively. The complex voltages, currents, and powers (loads) at the load nodes are represented by the column vectors $\mathbf{V}$, $\mathbf{I}$, and $\mathbf{s}$, respectively. Each node's complex power $s_j = p_j + \mathrm{i}q_j$ includes active (p_j) and reactive (q_j) components; hence, $\mathbf{s} = \mathbf{p} + \mathrm{i}\mathbf{q}$ is the vector of powers. In general, for a complex number $\mathbf{w}$, we use $\mathbf{w}^*$ for its complex conjugate, and for a matrix $\mathbf{W}$, $\mathbf{W}_{i:j,k:l}$ is used to denote the submatrix of $\mathbf{W}$ consisting of rows i through j and columns k through l, and $\mathbf{w}^T$ and $\mathbf{W}^T$ denote the transpose of the vector $\mathbf{w}$ and $\mathbf{W}$, respectively. An all-ones column vector, denoted as $\mathbf{1}^k$, is a vector of length k where every element is 1.

In what follows, we introduce three different feasibility regions. First, we present the full nonlinear feasible region, used to determine the true feasibility of load samples. Second, we introduce a polyhedral restriction of this feasible region which, in combination with the LDP, is used to find a change of measure. Third, we define an approximation of the feasible region, derived using the LD model, which is also used to find a change of measure through the LDP.

2.1 Bus Injection Model and its Feasible Region

For the first model, the feasible region is defined by loads that satisfy the nonlinear Bus Injection Model and the voltage drop constraints. The Bus Injection Model is given by the implicit power flow equation

$$\mathbf{V} = V_0 \mathbf{1}^d - \mathbf{Z}\mathrm{diag}(\mathbf{V}^*)^{-1}\mathbf{s}^*. \tag{1}$$

The voltage drop constraints assure that the voltage magnitudes are within a certain range and are made precise by allowing a $\Delta \times 100\%$ voltage deviation from the generator's voltage V_0, denoted by pre-specified bounds $(1-\Delta)|V_0|$ and $(1+\Delta)|V_0|$. Therefore, for $\mathbf{s} = \mathbf{p} + \mathrm{i}\mathbf{q}$, the feasible region, denoted by $\mathscr{C}$, can be expressed as the collection of all active and reactive powers $\mathbf{s}$ such that:

$$\mathscr{C} := \left\{ \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix} : \mathbf{V} = V_0 \mathbf{1}^d - \mathbf{Z}\mathrm{diag}(\mathbf{V}^*)^{-1}\mathbf{s}^*, \quad (1-\Delta)|V_0| \le |V_j| \le (1+\Delta)|V_0|, \quad j \in \mathscr{V} \right\}. \tag{2}$$

These equations can result in safe regions that are non-convex, making optimization difficult and potentially intractable from a computational standpoint. Therefore, in the next two sections, we consider more tractable alternatives.

2.2 Polyhedral Restriction of the Feasible Region

In Christianen et al. (2023), we developed a procedure to restrict the feasible region of (2) to another set of equations and inequalities for which we can guarantee a solution to the Bus Injection Model and the satisfaction of the voltage drop constraints. In the restriction, we distinguish between power consumption s_j^c and generation s_j^g at each node j, where $s_j = s_j^c - s_j^g$, and s_j^c, s_j^g are complex, representing active and reactive components. To compactly represent these across a network, we use vectors $\mathbf{p}^c, \mathbf{q}^c$ for consumption, and $\mathbf{p}^g, \mathbf{q}^g$ for generation, incorporated in $\tilde{\mathbf{s}} = \begin{pmatrix} \mathbf{p}^{cT} & \mathbf{q}^{cT} & \mathbf{p}^{gT} & \mathbf{q}^{gT} \end{pmatrix}^T$. Then, in the procedure, it is necessary to have the knowledge of a pair $(\hat{\mathbf{V}}, \hat{\mathbf{s}})$ that satisfies the power flow equation in (1) (e.g., the trivial solution) and that the voltage angle differences between $\mathbf{V}$ and $\hat{\mathbf{V}}$ are small. This means that, for $\Delta \in [0,1)$, we define the set of Δ-stable voltage vectors as all vectors that satisfy the constraint $|V_k - \hat{V}_j| \leq \Delta |\hat{V}_j|$ for all $k, j \in \mathscr{V} \setminus \{0\}$. Defining $\hat{V}_{\min} := \min_j |\hat{V}_j|$, yields a polyhedral restriction of $\mathscr{C}$ in (2) as

$$\mathscr{C}_I := \left\{ \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix} : \begin{pmatrix} \mathbf{p} - \hat{\mathbf{p}} \\ \mathbf{q} - \hat{\mathbf{q}} \end{pmatrix}^T (\mathbf{A}^T + \Delta \mathbf{B}^T)_{1:2d,1:4d} \leq \left(\Delta(1-\Delta)^2 \hat{V}_{\min}^3 \mathbf{1}^{4d} - \Delta(2\mathbf{B} + (1-\Delta)\mathbf{C}) \hat{\tilde{\mathbf{s}}} \right)^T \right\} \subseteq \mathscr{C},$$

where the inequality is implied such that each element in the (1×4) block matrices on the left-hand side is compared to the corresponding element in the $4d$ vector on the right-hand side, and the block matrices $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{4d \times 4d}$ are given by

$$\mathbf{A} = \begin{pmatrix} -(\mathbf{R}+\mathbf{X}) & -(-\mathbf{R}+\mathbf{X}) & \mathbf{R}+\mathbf{X} & -\mathbf{R}+\mathbf{X} \\ -(-\mathbf{R}+\mathbf{X}) & -(-\mathbf{R}-\mathbf{X}) & -\mathbf{R}+\mathbf{X} & -\mathbf{R}-\mathbf{X} \\ -(\mathbf{R}-\mathbf{X}) & -(\mathbf{R}+\mathbf{X}) & \mathbf{R}-\mathbf{X} & \mathbf{R}+\mathbf{X} \\ -(-\mathbf{R}-\mathbf{X}) & -(\mathbf{R}-\mathbf{X}) & -\mathbf{R}-\mathbf{X} & \mathbf{R}-\mathbf{X} \end{pmatrix}, \ \mathbf{B} = \mathbf{J}_4 \otimes (\mathbf{R}+\mathbf{X}), \text{ and } \mathbf{C} = \mathbf{J}_4 \otimes |\mathbf{Z}|,$$

where $\mathbf{J}_4$ is a (4×4) all-ones matrix and $\otimes$ denotes the Kronecker-product; see also Christianen et al. (2023, Theorem 3.1). Then, the complementary set $\mathscr{C}_I^c$, containing all infeasible vectors of power according to the polyhedral restriction, is defined as

$$\mathscr{C}_I^c = \left\{ \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix} : \exists i \text{ s.t. } \left(\begin{pmatrix} \mathbf{p} - \hat{\mathbf{p}} \\ \mathbf{q} - \hat{\mathbf{q}} \end{pmatrix}^T (\mathbf{A}^T + \Delta \mathbf{B}^T)_{1:2d,1:4d} \right)_i > \left(\Delta(1-\Delta)^2 \hat{V}_{\min}^3 \mathbf{1}^{4d} - \Delta(2\mathbf{B} + (1-\Delta)\mathbf{C}) \hat{\tilde{\mathbf{s}}} \right)_i^T \right\}, \tag{3}$$

where the inequality is implied again elementwise.

2.3 Approximation of the Feasible Region

Instead of a restriction of the feasible region in (2), in this subsection, we provide an approximation. To do so, we approximate the voltages in the network by the LD model. When necessary, we denote quantities related to the LD model by a subscript L. In contrast to the Bus Injection Model, the LD model yields explicit expressions for the squared voltages on all nodes, represented by $(\mathbf{V}_L)^2$, and are calculated as: $(\mathbf{V}_L)^2 := V_0^2 \mathbf{1}^d - 2 \begin{pmatrix} \mathbf{R} & \mathbf{X} \end{pmatrix} \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix}$. Then, we can approximate the feasible region as

$$\mathscr{C}_L := \left\{ \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix} : (\mathbf{V}_L)^2 = V_0^2 \mathbf{1}^d - 2 \begin{pmatrix} \mathbf{R} & \mathbf{X} \end{pmatrix} \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix}, \quad (1-\Delta)|V_0|\mathbf{1}^d \leq |\mathbf{V}_L| \leq (1+\Delta)|V_0|\mathbf{1}^d \right\} \approx \mathscr{C},$$

and the complementary set $\mathscr{C}_L^c$, containing all infeasible loads, as:

$$\mathscr{C}_L^c := \left\{ \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix} : \exists i \text{ s.t. } \begin{pmatrix} \mathbf{R} & \mathbf{X} \end{pmatrix}_{i,1:2d} \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix} < -\frac{|V_0|^2 \Delta(2+\Delta)}{2} \text{ or } \begin{pmatrix} \mathbf{R} & \mathbf{X} \end{pmatrix}_{i,1:2d} \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix} > \frac{|V_0|^2 \Delta(2-\Delta)}{2} \right\}. \tag{4}$$

With a slight abuse of notation for brevity, we use $\mathbf{s}$ to denote that $\begin{pmatrix}\mathbf{p}^T & \mathbf{q}^T\end{pmatrix}^T$ belongs to a constraint set. In other words, when we say $\mathbf{s}$ is within a constraint set, we are specifically referring to the combined vector of active and reactive powers.

3 PROBLEM DESCRIPTION

For distribution network reliability, deterministic models often fail to capture uncertainties and fluctuations in real-world power systems. As a result, the use of stochastic models is necessary to address the stochastic nature of loads. However, this brings the risk of violating the network's voltage drop constraints. Therefore, we consider the events $\mathscr{A}_\gamma = \{\mathbf{s} \in \mathscr{C}^c\}$, where the set $\mathscr{C}^c$ represents all infeasible loads (see Section 2.1). Additionally, the loads $\mathbf{s}$ capture the network's stochasticity, characterized by a safe operating point and stochastic fluctuations around this point. The safe operating point is given by the vector $\boldsymbol{\mu}$, partitioned as $\boldsymbol{\mu} = \begin{pmatrix}\boldsymbol{\mu}_\mathbf{p}^T & \boldsymbol{\mu}_\mathbf{q}^T\end{pmatrix}^T$ where $\boldsymbol{\mu}_\mathbf{p}$ and $\boldsymbol{\mu}_\mathbf{q}$ denote the safe operating levels of active and reactive power, respectively. The stochastic fluctuations around the safe operating point are modeled as Gaussian noise with a mean of zero and a scaled covariance matrix $\boldsymbol{\Sigma}/\gamma$. It captures the variability and correlation between active and reactive power fluctuations. Here, the covariance matrix $\boldsymbol{\Sigma}$ is also partitioned according to $\mathbf{p}$ and $\mathbf{q}$, i.e.,

$$\boldsymbol{\Sigma} = \begin{pmatrix}\boldsymbol{\Sigma}_{\mathbf{pp}} & \boldsymbol{\Sigma}_{\mathbf{pq}} \\ \boldsymbol{\Sigma}_{\mathbf{qp}} & \boldsymbol{\Sigma}_{\mathbf{qq}}\end{pmatrix},$$

where the submatrices of $\boldsymbol{\Sigma}$ detail the variability in noise and dependencies between these components. By scaling the noise with $1/\gamma$, we adjust the impact of stochastic fluctuations on the network. Larger values of γ correspond to rarer events, as the variance of the distribution decreases with increasing γ: $\mathbf{s} \sim \mathscr{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/\gamma)$. The goal is to estimate

$$\mathfrak{p}_\gamma = \mathbb{P}(\mathscr{A}_\gamma) = \mathbb{P}(\mathbf{s} \in \mathscr{C}^c) \to 0 \text{ as } \gamma \to \infty. \tag{5}$$

Having discussed the power flow models and their feasible regions, we proceed with the IS strategies, which consists of two components: a change of measure and a feasibility check. In the following section, we focus on the change of measure, because the feasibility check uses previous developed methods. Specifically, the duality method has been covered in Christianen et al. (2024), and NR methods and backward-forward sweep-based algorithms are well-known; see e.g. Montoya et al. (2021).

4 CHANGE OF MEASURE

The goal of this section is to introduce the way we change measure. We first focus on changes of measure directly informed by the LDP which yields a change of measure through the solution of a constrained optimization problem. The optimization problem's constraints are defined by the feasible regions derived from power flow models. Here, we use linear models, which simplify the feasible regions to unions of half-spaces as described in the polyhedral restriction in Section 2.2 and the LD model in Section 2.3. Consequently, we have two options: the LDP with the polyhedral restriction and the LDP with the LD model. In the other direction, we introduce the CE method as an alternative way to change measure. The traditional CE method for distribution network reliability involves explicit computations of voltages in the network. Here, we use the duality method developed in previous work Christianen et al. (2024) which avoids this step. We adapt the CE method to be able to use duality. Alternatively, we use the CE method and compute explicitly the voltages in the network by the NR method. In this direction, we also have two options to change measure: the CE method using duality and the CE method using the NR method.

4.1 Change of Measure using the LDP and Linear Power Flow Equations

The LDP is a classical theory on the asymptotic behavior of sequences of probability distributions (Dembo and Zeitouni 1998). The LDP states that the most likely point in the rare event set $\mathscr{C}^c$ is important in

estimating the probability of rare events, $\mathfrak{p}_\gamma$. For loads following a multivariate Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, the most likely point, $\mathbf{s}^*$, is the solution of the constrained optimization problem:

$$\inf_{\mathbf{s}} \quad (\mathbf{s}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{s}-\boldsymbol{\mu}) \quad \text{s.t.} \qquad \mathbf{s} \in \mathscr{C}^c. \tag{6}$$

Thus, a reasonable IS strategy is to shift the mean of the loads to $\mathbf{s}^*$. To get an approximation of a solution of (6), which is constrained by the region $\mathscr{C}^c$ defined by the Bus Injection Model's nonlinear equations, we simplify by replacing $\mathscr{C}^c$ with a union of m half-spaces $\bigcup_{i=1}^m (\mathbf{s}-\boldsymbol{\mu})^T \boldsymbol{\xi}_i \geq b_i$ approximating the two model-specific sets: $\mathscr{C}_I^c$ in (3) from the polyhedral restriction and $\mathscr{C}_L^c$ in (4) from the LD model. Such a union involves vectors $\boldsymbol{\xi}_i$ and scalars b_i, $i = 1, \ldots, m$, leading to a more tractable optimization problem

$$\inf_{\mathbf{s}} \quad (\mathbf{s}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{s}-\boldsymbol{\mu}) \quad \text{s.t.} \qquad \bigcup_{i=1}^{m} (\mathbf{s}-\boldsymbol{\mu})^T \boldsymbol{\xi}_i \geq b_i, \tag{7}$$

where $\inf_{\mathbf{s}} (\mathbf{s}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{s}-\boldsymbol{\mu}) > 0$ since $\boldsymbol{\mu}$ corresponds to the safe operating point.

To solve the overall optimization problem in (7), we decouple the optimization problem into separate subproblems corresponding to each linear constraint, solve them individually using Lemma 1, and then select the minimum among all solutions according to Lemma 2.

Lemma 1 addresses the specific case of the constrained quadratic minimization problem where there is only one linear constraint. The lemma shows that a solution to the problem can be found in a closed form.

Lemma 1 Consider the constrained minimization problem in (7) where we include only one of the m linear constraints into the problem, indexed by i_0, so that $(\mathbf{s}-\boldsymbol{\mu})^T \boldsymbol{\xi}_{i_0} \geq b_{i_0}$. Then, the solution of this problem is given by $\mathbf{s} = \boldsymbol{\mu} + \frac{b_{i_0}}{\boldsymbol{\xi}_{i_0}^T \boldsymbol{\Sigma} \boldsymbol{\xi}_{i_0}} \boldsymbol{\Sigma} \boldsymbol{\xi}_{i_0}$.

Proof. The Lagrangian for this problem, including the inequality constraint, is $\mathscr{L}(\mathbf{s}, \lambda) = (\mathbf{s}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{s}-\boldsymbol{\mu}) + \lambda (b_{i_0} - (\mathbf{s}-\boldsymbol{\mu})^T \boldsymbol{\xi}_{i_0})$. The KKT conditions are as follows: by stationarity, $\nabla_{\mathbf{s}} \mathscr{L}(\mathbf{s}, \lambda) = 2\boldsymbol{\Sigma}^{-1} (\mathbf{s}-\boldsymbol{\mu}) - \lambda \boldsymbol{\xi}_{i_0} = \mathbf{0}$, by primal feasibility, $(\mathbf{s}-\boldsymbol{\mu})^T \boldsymbol{\xi}_{i_0} \geq b_{i_0}$, by dual feasibility, $\lambda \geq 0$, and complementary slackness, $\lambda (b_{i_0} - (\mathbf{s}-\boldsymbol{\mu})^T \boldsymbol{\xi}_{i_0}) = 0$. Now, solving the stationary condition $2\boldsymbol{\Sigma}^{-1} (\mathbf{s}-\boldsymbol{\mu}) - \lambda \boldsymbol{\xi}_{i_0} = \mathbf{0}$, yields $\boldsymbol{\Sigma}^{-1} (\mathbf{s}-\boldsymbol{\mu}) = \frac{\lambda}{2} \boldsymbol{\xi}_{i_0}$. Multiplying both sides by $\boldsymbol{\Sigma}$ and rearranging, we get: $\mathbf{s}-\boldsymbol{\mu} = \frac{\lambda}{2} \boldsymbol{\Sigma} \boldsymbol{\xi}_{i_0}$. Now, substituting this into the complementary slackness condition and rearranging: $\lambda (b_{i_0} - \frac{\lambda}{2} \boldsymbol{\xi}_{i_0}^T \boldsymbol{\Sigma} \boldsymbol{\xi}_{i_0}) = 0$. Since $\lambda \geq 0$, this implies that either $\lambda = 0$ or $b_{i_0} = \frac{\lambda}{2} \boldsymbol{\xi}_{i_0}^T \boldsymbol{\Sigma} \boldsymbol{\xi}_{i_0}$. If $\lambda = 0$, then from the stationary condition, $\mathbf{s} = \boldsymbol{\mu}$, which is not feasible due to the inequality constraint. Therefore, we must have $b_{i_0} = \frac{\lambda}{2} \boldsymbol{\xi}_{i_0}^T \boldsymbol{\Sigma} \boldsymbol{\xi}_{i_0}$. Substituting $\mathbf{s}-\boldsymbol{\mu} = \frac{\lambda}{2} \boldsymbol{\Sigma} \boldsymbol{\xi}_{i_0}$ back into the primal feasibility constraints: $(\mathbf{s}-\boldsymbol{\mu})^T \boldsymbol{\xi}_{i_0} = (\frac{\lambda}{2} \boldsymbol{\Sigma} \boldsymbol{\xi}_{i_0})^T \boldsymbol{\xi}_{i_0} = \frac{\lambda}{2} \boldsymbol{\xi}_{i_0}^T \boldsymbol{\Sigma} \boldsymbol{\xi}_{i_0} = b_{i_0}$, shows that the solution of this optimization problem is actually attained at equality. Finally, solving for $\frac{\lambda}{2}$ yields $\frac{\lambda}{2} = \frac{b_{i_0}}{\boldsymbol{\xi}_{i_0}^T \boldsymbol{\Sigma} \boldsymbol{\xi}_{i_0}}$, and substituting back into the stationary condition gives $\mathbf{s} = \boldsymbol{\mu} + \frac{b_{i_0}}{\boldsymbol{\xi}_{i_0}^T \boldsymbol{\Sigma} \boldsymbol{\xi}_{i_0}} \boldsymbol{\Sigma} \boldsymbol{\xi}_{i_0}$. □

Lemma 2 describes the setting of finding the infimum constrained by a union of half-spaces. The lemma asserts that a solution to the constrained quadratic minimization problem in (7) is given by considering the set of minimizers corresponding to each individual constraint.

Lemma 2 A solution of (7) is given by a minimizer of the set

$$\left\{ \boldsymbol{\mu} + \frac{b_i}{\boldsymbol{\xi}_i^T \boldsymbol{\Sigma} \boldsymbol{\xi}_i} \boldsymbol{\Sigma} \boldsymbol{\xi}_i, \quad i = 1, \ldots, m \right\}. \tag{8}$$

Proof. Suppose, for the sake of contradiction, that there exists a solution $\mathbf{s}^*$ that is not a minimizer of the set (8). Then, there are two options. The first option is that there is another member of the set (8) that minimizes the expression in (7), leading to a direct contradiction with the the minimality assumption. The second option is that there is a minimizer that lies in the interior of the infeasible region. In other words, there exists a minimizer $\mathbf{s}^*$ that is not a member of the set (8), such that at least for one constraint i, we have the strict inequality $(\mathbf{s}^* - \boldsymbol{\mu})^T \boldsymbol{\xi}_i > b_i$. However, by convexity of the objective function $(\mathbf{s}^* - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{s}^* - \boldsymbol{\mu})$, we can locally improve the objective by moving from $\mathbf{s}^*$ towards $\boldsymbol{\mu}$; hence, it contradictions with the minimality assumption. □

4.2 Change of Measure using the CE Method

Building upon the work of de Boer et al. (2005), our study applies the CE method to achieve an efficient change of measure for rare-event simulation. The CE method aims to minimize the Kullback-Leibler (KL) divergence between the optimal IS distribution, which ensures a zero-variance IS estimator, and a Gaussian distribution characterized by a mean v^*. This optimization is done by a multi-level algorithm. The algorithm constructs a sequence of reference parameters $\{\boldsymbol{\nu}_t, t \geq 0\}$ for the mean of the loads and a sequence of tolerance levels $\{\Delta_t, t \geq 1\}$ for the voltage drop constraint, and iterates in both Δ_t and $\boldsymbol{\nu}_t$, leading to a mean of the loads $\boldsymbol{\nu}^*$ in the final iteration.

On the one hand, for a fixed $\boldsymbol{\nu}_{t-1}$, let Δ_t be a $(1-\rho)$-quantile under $\boldsymbol{\nu}_{t-1}$. That is, Δ_t satisfies $\mathbb{P}_{\boldsymbol{\nu}_{t-1}}(\mathbf{s} \in \mathscr{C}^c) \geq \rho$, and $\mathbb{P}_{\boldsymbol{\nu}_{t-1}}(\mathbf{s} \in \mathscr{C}) \geq 1-\rho$, where the samples $\mathbf{s}$'s are drawn from a Gaussian distribution with mean $\boldsymbol{\nu}_{t-1}$. Notice that the set $\mathscr{C}^c$ depends on the voltage drop parameter Δ, yet the algorithm uses the parameter Δ_t at each iteration, slightly adjusting the criteria defining $\mathscr{C}^c$. To do the quantile estimation, we use an estimator $\hat{\Delta}_t$ of Δ_t that can be obtained by drawing N random samples $\mathbf{s}_1, \ldots, \mathbf{s}_N$ from a Gaussian distribution with mean $\boldsymbol{\nu}_{t-1}$, and a *binary search* strategy. This involves setting an interval $[a,b] = [0,\Delta]$, where the quantile estimate is presumed to reside, alongside establishing a tolerance level ε for the precision of the estimate and a target quantile level ρ. Then, the process iterates, narrowing the interval based on the midpoint $c = (a+b)/2$. At each iteration, the probability $\mathbb{P}_{\mathbf{v}_{t-1}}(\mathbf{s} \in \mathscr{C}^c)$ is approximated by $\hat{\mathfrak{p}}_{\Delta_t} = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}\{\mathbf{s} \in \mathscr{C}^c\}$, with the interval updated to $[c,b]$ if $\hat{\mathfrak{p}}_{\Delta_t} > \rho$ or to $[a,c]$ if $\hat{\mathfrak{p}}_{\Delta_t} \leq \rho$. This process repeats until the interval width $b-a$ is less than or equal to ε, at which point the midpoint of the interval is taken as the estimate for the $(1-\rho)$-quantile.

On the other hand, for fixed Δ_t and $\boldsymbol{\nu}_{t-1}$, update $\boldsymbol{\nu}_t$ according to (for details, see de Boer et al. (2005)):

$$\boldsymbol{\nu}_t = \frac{\sum_{i=1}^{N} \mathbb{1}\{\mathbf{s}_i \in \mathscr{C}^c\} W(\mathbf{s}_i; \boldsymbol{\mu}, \boldsymbol{\nu}_{t-1}) \mathbf{s}_i}{\sum_{i=1}^{N} \mathbb{1}\{\mathbf{s}_i \in \mathscr{C}^c\} W(\mathbf{s}_i; \boldsymbol{\mu}, \boldsymbol{\nu}_{t-1})}, \tag{9}$$

where $W(\mathbf{s}_i; \boldsymbol{\mu}, \boldsymbol{\nu}_{t-1})$ is the likelihood ratio at $\mathbf{s}_i$ between a Gaussian distribution with means $\boldsymbol{\mu}$ and $\boldsymbol{\nu}_{t-1}$.

In other words, we initialize the algorithm by choosing a not very small ρ, say $\rho = 10^{-2}$, a small level ε, such as $\varepsilon = 10^{-4}$, and by defining $\boldsymbol{\nu}_0 = \boldsymbol{\mu}$. Next, the level Δ_1 is updated to make the target event artificially less rare, i.e., we let $\Delta_1 (\Delta_1 < \Delta)$ be such that, under the original distribution $\mathscr{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/\gamma)$, for fixed γ, the probability $\mathfrak{p}_{\Delta_1} = \mathbb{P}_{\boldsymbol{\mu}}(\mathbf{s} \in \mathscr{C}^c)$ is at least ρ. We then update $\boldsymbol{\nu}_1$ according to (9). The algorithm iteratively repeats these two steps: updating Δ_t and $\boldsymbol{\nu}_t$. It terminates when at some iteration t a level is reached which is at least Δ and thus the original value of Δ can be used without getting too few rare samples.

5 SIMULATION RESULTS

In this section, we focus on IS strategies for an IEEE 15-node network to estimate failure probabilities. These strategies include a change of measure: the LDP with the LD model, the LDP with the polyhedral restriction and the CE method with either the duality or the NR method. Observe that in the CE method, see e.g. (9), in order to apply this method we continuously check whether a sample is in the feasible region or not, which we do by either the duality or the NR method. As a result, the feasibility check is integrated with the change of measure. In contrast, for the LDP with the LD model and for the LDP with

the polyhedral restriction, we first change the measure, and then verify whether a sample from this new distribution is in the feasible region or not, and we do so by either using the duality or the NR method. Thus, we consider different IS strategies: the first type is based on a change of measure based on the LDP with either the LD model or the polyhedral restriction, which is followed by a feasibility check using the duality or the NR method, and the second type is based on the CE method with an integrated feasibility check done by either the duality or the NR method.

5.1 Description of IEEE 15-node Network

We employ the IEEE 15-node tree network from the MATPOWER package Zimmerman et al. (2011), as depicted in Figure 1. We modify the network's parameters such that they closely follow the assumptions required for the duality method, which focuses on small x/r ratios and minimal reactive power magnitudes. To achieve this, we make specific adjustments to the original data from Zimmerman et al. (2011). We scale down both reactances and reactive power magnitudes to 10% of their initial values, and define a safe operating point ($\boldsymbol{\mu}_\mathbf{p}$ and $\boldsymbol{\mu}_\mathbf{q}$) which yields corresponding voltage levels ($\mathbf{V}$) within a 10% deviation from the generator node's voltage, standardized as $V_0 = 1$. These modifications are detailed in Table 1.

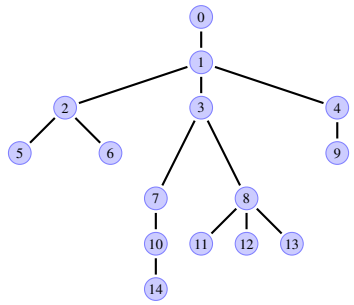


Figure 1: IEEE-15 node tree network showing the node numbers and connections. Node 0 is the generator node.

Table 1: IEEE-15 node network data.

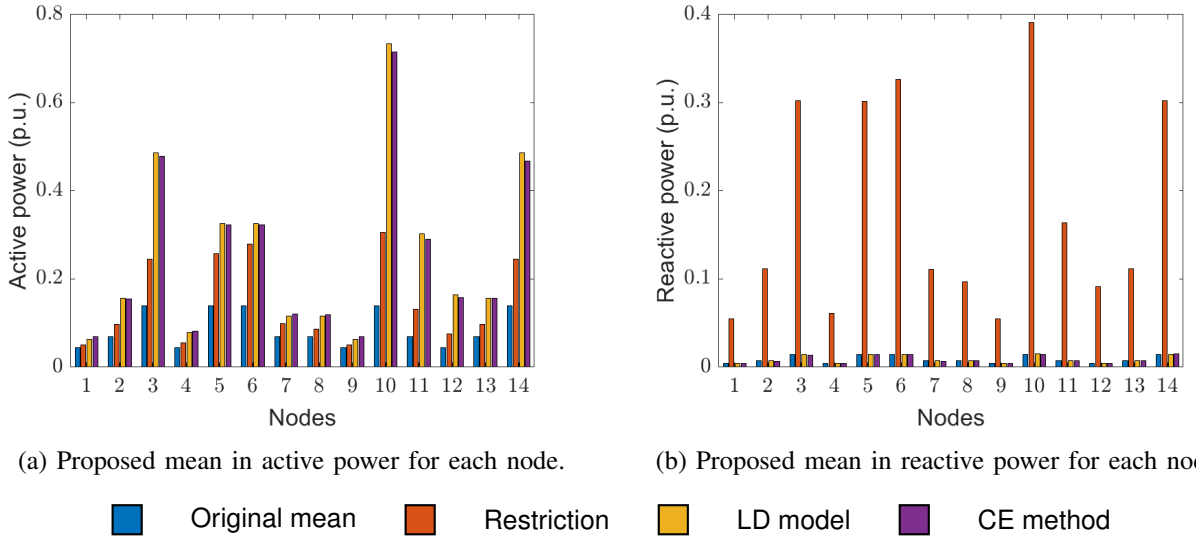
Line number	Bus From	Bus To	$\mathbf{r}$ (p.u.)	$\mathbf{x}$ (p.u.)	$\boldsymbol{\mu}_\mathbf{p}$	$\boldsymbol{\mu}_\mathbf{q}$	$\mathbf{V}$
1	0	1	1.1183	0.1094	44.1	4.50	0.9858
2	1	2	2.1134	0.1426	140.0	14.28	0.9786
3	1	3	0.9671	0.0946	70.0	7.14	0.9758
4	1	4	1.6638	0.1122	70.0	7.14	0.9652
5	2	5	1.0341	0.0698	70.0	7.14	0.9782
6	2	6	0.8993	0.0607	140.0	14.28	0.9769
7	3	7	1.4839	0.1001	140.0	14.28	0.9775
8	3	8	0.6951	0.0680	140.0	14.28	0.9839
9	4	9	1.3940	0.0940	44.1	4.50	0.9832
10	7	10	2.0235	0.1365	70.0	7.14	0.9747
11	8	11	0.9893	0.0667	140.0	14.28	0.9723
12	8	12	1.2591	0.0849	44.1	4.50	0.9716
13	8	13	1.8436	0.1244	70.0	7.14	0.9745
14	10	14	1.6638	0.1122	44.1	4.50	0.9744

We assume that the loads $\mathbf{s}$ are Gaussian distributed with mean $\boldsymbol{\mu} = \begin{pmatrix}\boldsymbol{\mu}_\mathbf{p}^T & \boldsymbol{\mu}_\mathbf{q}^T\end{pmatrix}^T$ as in Table 1. Each component has a variance given by the square of each element in $\boldsymbol{\mu}$, divided by the rarity parameter γ, denoted as $\boldsymbol{\mu}^2/\gamma$ and there is zero correlation between any two different components, i.e., $\text{corr}(\mathbf{s}_i;\mathbf{s}_j) = 0$ for any i, j with $i \neq j$. Consequently, the covariance matrix $\boldsymbol{\Sigma}$ is a diagonal matrix, with the variances $\boldsymbol{\mu}^2/\gamma$ on its diagonal and zeroes elsewhere. The rarity parameter γ is adjusted between 1 and 5 to control the frequency of the occurence of rare events. Our objective is to estimate the probability described in (5).

5.2 Change of Measure

This section compares different changes of measure. We identify those vectors of active and reactive power that are most likely to violate the voltage drop constraints, thus finding the most probable scenarios for constraint violation. The changes in mean for active and reactive power for all nodes, as suggested by each method, are depicted in Figure 2a for active power and Figure 2b for reactive power. In both figures, the powers at each node are according to the LDP with the LD model (yellow), the LDP with the polyhedral restriction (red), and the CE method based on the duality method (purple), against the original mean (blue). The powers according to the CE method based on the NR method are not shown in the figures, as they were identical to those of the CE method based on the duality method.

The proposed changes of measure by the LDP with the LD model and the CE method based on the duality method use a significant increase in active power, with relatively minor changes in reactive power compared to the original mean that corresponds to the safe operating point. This behavior deviates from the



(a) Proposed mean in active power for each node.

(b) Proposed mean in reactive power for each node.

Figure 2: Changes of measure proposed by the LDP with LD model (yellow), the LDP with the polyhedral restriction (orange), and the CE method using the duality method (purple), and the original mean (blue).

behavior of the polyhedral restriction, which demonstrates only a slight increase in active power while a significant increase in reactive powers. Examining the active power, there is a clear difference between the LDP with the LD model and the LDP with the polyhedral restriction. The LDP with the LD model proposes higher active powers than the more pessimistic approach by the LDP with the polyhedral restriction. This can be explained as follows: by using a restriction, we effectively narrow the feasible region. As a result, the estimated active powers under the polyhedral restriction are typically lower, reflecting a conservative estimation of what the system can reliably handle. Therefore, we exclude the approach using the LDP with the polyhedral restriction from our subsequent numerical experiments.

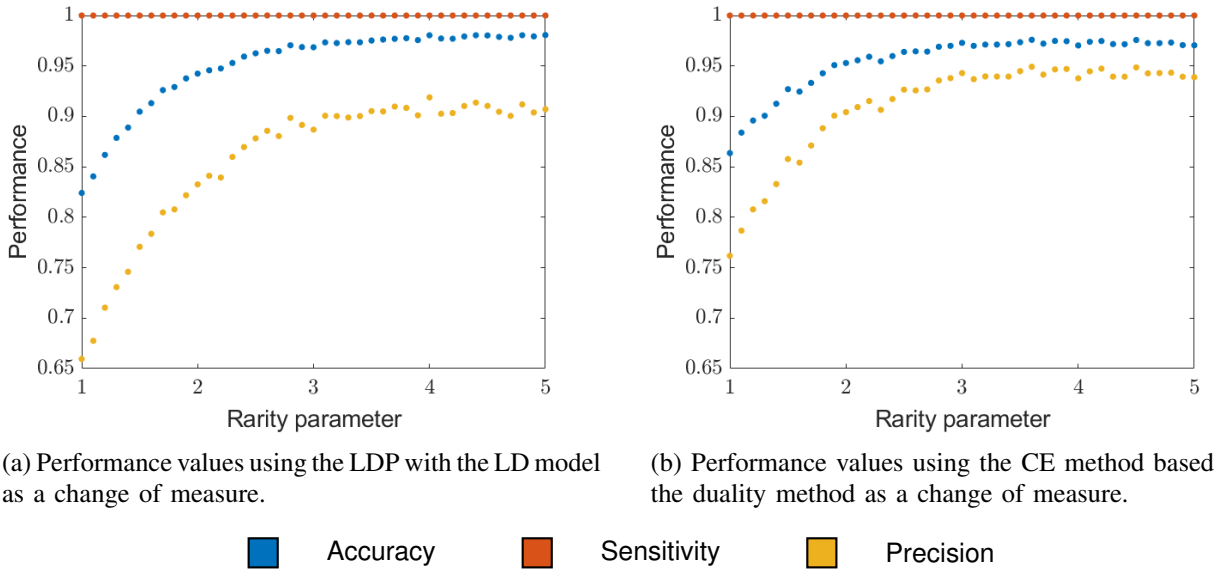
5.3 Feasibility Check

In this section, we focus on the other important component of our IS strategy: checking feasibility. In principle, in order to check feasibility, one needs to select a power flow model and verify whether a sample of loads satisfies the model's equations. In our numerical experiments, we have conducted extensive numerical experiments using the duality method, the NR method, backward-forward sweep-based algorithms, all of which rely on nonlinear power flow models. We have also tested the feasibility checks using the simpler, linear models and feasible regions provided by the LD model and the polyhedral restriction; see Sections 2.2 and 2.3. It was immediately clear that checking feasibility based on linearized models led to inaccurate results. Consequently, in this section, we focus on the feasibility checks using nonlinear models, which adhere more closely to the actual power flow equations. Specifically, we compare the efficacy of the duality-based approach in comparison with the NR method.

To check the limitations of the assumptions in checking feasibility using the duality-based approach, we compare the accuracy, sensitivity, and precision of the duality-based method with the NR method. Here, accuracy measures the correctness of the classification, considering both true positives (correctly identified feasible loads) and true negatives (correctly identified infeasible loads). Precision assesses the accuracy of the methods when they predict an infeasible load, measuring the proportion of true positives relative to all predicted positives, while sensitivity, in this context, focuses on the ability of the method to correctly identify feasible loads (true positives) relative to all actual feasible loads.

In our experiments, we increase the rarity parameter from 1 to 5, employing a change of measure determined by either the LDP with the LD model or the CE method based on the duality method. For

each rarity parameter, we generate 10^4 samples using these changes of measure and check their feasibility using both the duality method and NR methods to compare the performance of these approaches.



(a) Performance values using the LDP with the LD model as a change of measure.

(b) Performance values using the CE method based on the duality method as a change of measure.

Figure 3: Comparative metrics: accuracy (blue), sensitivity (orange), and precision (yellow) of feasibility checks based on the duality method against the NR method across rarity parameters.

Increasing the rarity parameter under these conditions leads to even better performance due to a reduction in variance, which diminishes the likelihood of encountering negative power consumption and increases the probability of meeting the duality method's assumptions. Notably, for the LDP with the LD model, the sensitivity remains consistently at 1, and the accuracy increases from 0.83 at rarity parameter equal to 1 to 0.97 at rarity parameter equal to 5. Similarly, for the CE method based on the duality method, the accuracy rises from 0.87 at rarity parameter equal to 1 to 0.98 at rarity parameter equal to 5. In terms of precision, the LDP with the LD model shows an increase from 0.66 at rarity parameter equal to 1 to 0.90 at rarity parameter equal to 5, while the CE method based on the duality method sees a rise from 0.76 at rarity parameter equal to 1 to 0.94 at rarity parameter equal to 5. For checking feasibility, the duality method is effective and performs best when combined with the CE method, as compared to its combination with the change of measure by the LDP with the LD model, under typical distribution network parameter settings.

5.4 Rare-event Probability Estimates

In this section, we estimate rare-event probabilities using IS strategies that employ a change of measure using either the LDP with the LD model or the CE method. We generate 10^4 samples using these measures, which are then checked by either the duality method or the NR method. The outputs are re-weighted using the likelihood ratio to ensure the unbiasedness of the estimators when using the NR method. To provide a detailed overview on the estimates derived from these four different combinations, we present the results in Table 2. The table shows point estimates along with their 95% confidence intervals (CI) for varying values of the rarity parameter, increasing from 1 to 5. In our experiments, we consider the NR method as the baseline for true probability estimates and compare it to the duality method. The estimates obtained by using either the LDP with the LD model or the CE method consistently result in estimates within the same order of magnitude as the true rare-event probabilities across this range.

Table 2: Point estimates and 95% CI for rare-event probabilities using IS for rarity parameters γ.

	LDP with the LD model		CE method	
γ	Duality	NR	Duality	NR
1	$2.1553(\pm 0.1536) \times 10^{-13}$	$4.1793(\pm 0.2441) \times 10^{-13}$	$0.6795(\pm 0.9652) \times 10^{-13}$	$3.9788(\pm 0.4528) \times 10^{-13}$
2	$1.2097(\pm 0.0918) \times 10^{-24}$	$2.0062(\pm 0.1459) \times 10^{-24}$	$0.7504(\pm 0.1936) \times 10^{-24}$	$1.6092(\pm 0.3185) \times 10^{-24}$
3	$0.5752(\pm 0.0494) \times 10^{-35}$	$1.1917(\pm 0.0994) \times 10^{-35}$	$1.2448(\pm 1.1489) \times 10^{-35}$	$1.1249(\pm 0.3288) \times 10^{-35}$
4	$3.1496(\pm 0.2969) \times 10^{-47}$	$6.9342(\pm 0.6523) \times 10^{-47}$	$1.7911(\pm 0.7657) \times 10^{-47}$	$0.5035(\pm 0.4204) \times 10^{-47}$
5	$1.4758(\pm 0.1589) \times 10^{-58}$	$4.7229(\pm 0.4859) \times 10^{-58}$	$1.4493(\pm 0.4397) \times 10^{-58}$	$4.0357(\pm 1.7105) \times 10^{-58}$

5.5 Computational Effort

To distinguish between these combinations, we measure the time required to obtain estimates for rare-event probabilities while increasing the rarity parameter from 1 to 5. The recorded times for all four combinations reveal two observations. First, using duality as a feasibility check instead of NR substantially reduces computational time by a factor of 3, regardless of the change of measure employed. Second, utilizing the LDP with the LD model as a change of measure instead of the CE method based on either the duality or the NR method results in a time reduction by a factor of 2. This difference arises from the iterative nature of the CE method, whereas the LDP with the LD model requires only a single step to compute the change of measure. Hence, adopting duality as the feasibility check instead of NR and utilizing the LDP with the LD model as the change of measure instead of the CE method allows for a significant reduction in total computational time, achieving a factor of 6 improvement compared to the base reference case using NR as a feasibility check and the CE method based on the NR method to find a change of measure.

6 CONCLUSION

We have proposed IS strategies to estimate failure probabilities. Within these IS strategies, we have used a change of measure informed by two types of approaches: using the LDP or the CE method. Empirical studies have shown the effectiveness of combining a change measure, as proposed by either the LDP with the LD model or the CE method based on the duality method, with an accurate feasibility check based on nonlinear models. Regarding the feasibility check, the use of linear approximations in the feasibility check has resulted in large under or overestimates of the true probabilities. To overcome this issue, we have proposed the duality method as a fast alternative for the classical NR method in checking feasibility. Consequently, combining either the LDP with the LD model with the duality method or the CE method based on the duality method has resulted in faster and accurate rare-event probability estimation strategies.

ACKNOWLEDGMENTS

This research is supported by the Dutch Research Council under contract number 613.001.801.

REFERENCES

Bai, Y., Z. Huang, H. Lam, and D. Zhao. 2023. "Over-Conservativeness of Variance-Based Efficiency Criteria and Probabilistic Efficiency in Rare-Event Simulation". *Management Science, Articles in Advance*:1–22.

Cadini, F., G. L. Agliardi, and E. Zio. 2017. "Estimation of Rare Event Probabilities in Power Transmission Networks Subject to Cascading Failures". *Reliability Engineering and System Safety* 158:9–20.

Chertkov, M., F. Pan, and M. G. Stepanov. 2011. "Predicting Failures in Power Grids: The Case of Static Overloads". *IEEE Transactions on Smart Grid* 2(1):162–172.

Chertkov, M., M. Stepanov, F. Pan, and R. Baldick. 2011. "Exact and Efficient Algorithm to Discover Extreme Stochastic Events in Wind Generation over Transmission Power Grids". In *2011 50th IEEE Conference on Decision and Control and European Control Conference*. December 12-15, Orlando, FL, USA, 2174-2180.

Christianen, M. H. M., S. van Kempen, M. Vlasiou, and B. Zwart. 2023. "Polyhedral Restrictions of Feasibility Regions in Optimal Power Flow for Distribution Networks". *arXiv preprint arXiv: 2312.04300*.

Christianen, M. H. M., M. Vlasiou, and B. Zwart. 2024. "Comparing Power Flow Models in Tree Networks with Stochastic Load Demands". In *Operations Research and Enterprise Systems*, edited by F. Liberatore, S. Wesolkowski, M. Demange, and G. H. Parlier, 138–167. Cham: Springer Nature Switzerland.

de Boer, P.-T., D. P. Kroese, S. Mannor, and R. Y. Rubinstein. 2005. "A Tutorial on the Cross-Entropy Method". *Annals of Operations Research* 134:19–67.

Dembo, A. and O. Zeitouni. 1998. *Large Deviations Techniques and Applications*. Berlin, Heidelberg: Springer.

Jang, S., N. Ozay, and J. L. Mathieu. 2024. "Probabilistic Constraint Construction for Network-Safe Load Coordination". *IEEE Transactions on Power Systems* 39(2):4473–4485.

Khatod, D. K., V. Pant, and J. Sharma. 2006. "A Novel Approach for Sensitivity Calculations in the Radial Distribution System". *IEEE Transactions on Power Delivery* 21:2048–2057.

Molzahn, D. K. and I. A. Hiskens. 2019. "A Survey of Relaxations and Approximations of the Power Flow Equations". *Foundations and Trends® in Electric Energy Systems* 4(1-2):1–221.

Montoya, O., A. Molina Cabrera, and J. . Hernández. 2021. "A Comparative Study on Power Flow Methods Applied to AC Distribution Networks with Single-Phase Representation". *Electronics* 10(21):2573.

Nesti, T., J. Nair, and B. Zwart. 2019. "Temperature Overloads in Power Grids under Uncertainty: A Large Deviations Approach". *IEEE Transactions on Control of Network Systems* 6:1161–1173.

Owen, A. B., Y. Maximov, and M. Chertkov. 2019. "Importance Sampling the Union of Rare Events with an Application to Power Systems Analysis". *Electronic Journal of Statistics* 13(1):231–254.

Peterson, W., E. Makram, and T. Baldwin. 1989. "A Generalized PC-based Bus Impedance Matrix Building Algorithm". In *Proceedings. IEEE Energy and Information Technologies in the Southeast'*. April 09-12, Columbia, SC, USA, 432-436.

Tokdar, S. T. and R. E. Kass. 2010. "Importance Sampling: A Review". *WIREs Computational Statistics* 2(1):54–60.

Tonso, M., J. Morren, S. W. De Haan, and J. A. Ferreira. 2005. "Variable Inductor for Voltage Control in Distribution Networks". In *2005 European Conference on Power Electronics and Applications*. September 11-14, Dresden, Germany.

Vasmel, N. 2019. "Electrical Grid Failures: An Analysis of Rare Events". Master's thesis, Leiden University.

Verseveldt, B. 2023. "Analysis of Rare Events for AC Power Flow Models". Master's thesis, Utrecht University.

Zimmerman, R. D., C. E. Murillo-Sánchez, and R. J. Thomas. 2011. "MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education". *IEEE Transactions on Power Systems* 26(1):12–19.

AUTHOR BIOGRAPHIES

MARK CHRISTIANEN has obtained his Ph.D. degree at the Department of Mathematics and Computer Science, Eindhoven University of Technology, the Netherlands. He received his B.Sc. and M.Sc. degree in Applied Mathematics from Delft University of Technology, Delft, the Netherlands. His research interests include stochastic networks and probability theory. His email address is m.h.m.christianen@tue.nl and his website is https://www.tue.nl/en/research/researchers/mark-christianen.

HENRY LAM is an Associate Professor in the Department of Industrial Engineering and Operations Research at Columbia University. His research interests include Monte Carlo methods, uncertainty quantification, risk analysis and data-driven optimization. He serves as the area editor for Stochastic Models and Data Science in Operations Research Letters, and on the editorial boards of Management Science, Operations Research, INFORMS Journal on Computing, Applied Probability Journals, Stochastic Models, Manufacturing and Service Operations Management, and Queueing Systems. His email address is henry.lam@columbia.edu and his website is http://www.columbia.edu/~khl2114/.

MARIA VLASIOU is a Professor of Stochastic Processing Interacting Networks (University of Twente). Her research focuses on the performance of stochastic processes, with applications on manufacturing, telecommunications, mathematical biology, and energy systems. Prof. Vlasiou serves as associate editor in four journals and is the co-chair of the TPCs of IFIP Performance 2023 and NMC 2023. She has received the best paper award in ICORES 2013, the Marcel Neuts student paper award in MAM8, a prize at the 8th conference in Actuarial Science, and the INFORMS UPS G. Smith 2022 award. Her email address is m.vlasiou@utwente.nl and her website is https://people.utwente.nl/m.vlasiou.

BERT ZWART is the Leader of the CWI Stochastics Group (Amsterdam) and Professor with the Eindhoven University of Technology, the Netherlands. His research interests include applied probability and stochastic networks. His work on power systems is focusing on the applications of probabilistic methods to scheduling and reliability issues and rare events such as cascading failures and blackouts. He has been Stochastic Models area editor for Operations Research, the flagship journal of his profession, from 2009 to 2017 and currently serves on the editorial boards of five journals. He was co-organizer of a special semester on the Mathematics of energy systems taking place in Cambridge, U.K., Spring 2019. His e-mail address is Bert.Zwart@cwi.nl and his website is https://www.cwi.nl/nl/people/bert-zwart/.