14th International Conference on Structural Safety and Reliability - ICOSSAR'25
1-6 June, 2025 - Los Angeles California, USA
R. Ghanem, E. Johnson, S. Masri and A. Olivier (Eds)

OUTPERFORMING SUBSET SIMULATION BY SKIPPING: A CASE STUDY ON POWER GRID RELIABILITY ASSESSMENT

ZIRAN WANG[1], LEONARDO DUEÑAS-OSORIO[2], IASON PAPAIOANNOU[3] AND DANIEL STRAUB[4]

[1] Department of Civil and Environmental Engineering, Rice University
Houston, TX 77005, USA
E-mail: ziran.wang@rice.edu

[2] Department of Civil and Environmental Engineering, Rice University
Houston, TX 77005, USA
E-mail: leonardo.duenas-osorio@rice.edu

[3] TUM School of Engineering and Design, Technical University of Munich
Munich, BY 80333, Germany
E-mail: iason.papaioannou@tum.de

[4] TUM School of Engineering and Design, Technical University of Munich
Munich, BY 80333, Germany
E-mail: straub@tum.de

Key words: Subset simulation, Skipping subset simulation, Rare-event sampling, Power network reliability

Abstract. Estimating the failure probability of high-dimensional systems is crucial for infrastructure reliability assessment, especially when failures are rare but have significant consequences, as is the case with power systems. While subset simulation (SuS) is widely used for efficient rare-event probability estimation, its accuracy can suffer from large errors when the failure domain of the system is fragmented or its state space is rugged as in networked systems. This paper introduces skipping subset simulation (Skipping-SuS), an enhancement to the standard SuS that incorporates a skipping mechanism to improve global exploration of the state space. Skipping-SuS increases the transition probability between spatially separated failure regions, thereby increasing sample diversity and reducing both bias and variance in reliability estimates. More generally, Skipping-SuS features reduced error estimation losses to better support decision making. A case study on power grid reliability assessment using IEEE benchmark networks demonstrates that Skipping-SuS outperforms standard SuS by consistently identifying more failure regions and producing better estimates. Although the case study focuses on power systems, the proposed method is general and applicable to a broad range of rare-event problems in high-dimensional settings.

1 INTRODUCTION

Quantitative assessment of infrastructure network reliability is essential, especially as these systems continue to modernize and expand in scale. Such evaluations inform critical decisions in maintenance, design, resource allocation, and risk mitigation. Reliability is commonly assessed by estimating the failure probability p_f, defined as the probability that a performance metric of the network is negative. Mathematically, the failure probability can be expressed as:

$$p_f \triangleq \mathbb{E}_{\mathbf{X}}[\mathbb{I}\{g(\mathbf{X}) \leq 0\}] = \sum_{\mathbf{x} \in \Omega_{\mathbf{X}}} \mathbb{I}\{g(\mathbf{x}) \leq 0\} p_{\mathbf{X}}(\mathbf{x}) \tag{1}$$

where $g(\cdot)$ denotes the network performance function, and $\mathbf{X}$ is a random vector representing the state of network components. The sample space of $\mathbf{X}$ is $\Omega_{\mathbf{X}}$, and its probability mass function is $p_{\mathbf{X}}(\mathbf{x})$. The indicator function $\mathbb{I}\{\cdot\}$ equals one if the condition inside the braces holds and zero otherwise. Accordingly, failure is defined by the event $\mathcal{F} \triangleq \{g(\mathbf{X}) \leq 0\}$, and the set of $\mathbf{X}_{\mathcal{F}} \in \Omega_{\mathbf{X}}$ that satisfies this condition constitutes the failure region.

Sampling-based methods such as Monte Carlo simulation (MCS) are widely used due to their flexibility and general applicability. However, for engineered systems like power grids, where failures are rare events, particularly when aiming at an estimator with quality guarantees [1] or with reduced error loss via unbiasedness [2], crude MCS becomes computationally expensive and often impractical.

Subset simulation (SuS) has emerged as a widely adopted rare-event sampling technique due to its superior efficiency in high-dimensional problems [3]. SuS improves upon MCS by decomposing the rare failure event into a sequence of intermediate events with higher probabilities, enabling more tractable estimation of rare-event probabilities. Despite its advantages, SuS can be less effective when the failure domain is rugged or contains multiple disconnected regions with small probability mass. In such cases, the algorithm may miss relevant failure modes, leading to reliability estimates with large errors.

This paper introduces a enhancement to the standard subset simulation framework through a skipping mechanism [4, 5] designed to improve the algorithm's ability to discover disjoint and spatially separated failure regions. The proposed skipping subset simulation (Skipping-SuS) maintains the core structure of SuS but introduces controlled "skips" between candidate regions of the state space. This improves sample diversity and helps avoid local trapping, reducing both bias and variance in the estimated failure probability while maintaining computational efficiency.

To investigate the effectiveness of this approach, we conduct a case study in power grid reliability assessment—a challenging setting where failure regions are often sparse and non-local. Power grids are critical infrastructure systems vulnerable to a wide range of hazards, including equipment failures, natural disasters, and cyber-physical threats. Accurate reliability assessment is vital to support system planning and resilience analysis. However, the low probability of catastrophic failures renders conventional MCS infeasible for realistic systems. Our computational experiments on an IEEE benchmark test case show that Skipping-SuS consistently identifies a broader range of failure regions and yields more accurate reliability estimates than standard SuS. While the power grid serves as an illustrative example, the proposed method is broadly applicable to other domains involving high-dimensional rare-event modeling.

The remainder of this paper is structured as follows: Section 2 introduces the Direct Current Optimal Power Flow (DC-OPF) problem used as the testbed. Section 3 discusses the limitations of standard subset simulation and introduces the skipping subset simulation algorithm. Section 4 presents the experimental setup, analyzes the results, and highlights key insights. Section 5 concludes the paper and outlines directions for future work.

2 DIRECT CURRENT OPTIMAL POWER FLOW (DC-OPF) PROBLEM

We first briefly describe the testbed used for experiments, namely the Direct Current Optimal Power Flow (DC-OPF) problem. DC-OPF problem serves as a fundamental tool in power system operations for electricity market clearing, congestion management, generator dispatch scheduling, renewable integration studies, and contingency analysis [6, 7].

The power flow in transmission networks follows Kirchhoff's laws and operational constraints. While AC power flow models provide superior accuracy for transient stability analysis, their computational demands and data requirements render them impractical for reliability studies. The DC approximation offers a computationally efficient alternative by solving a linear system that neglects reactive power injections (Q_i) and assumes fixed voltage magnitudes ($V_i = 1$ per unit). This simplification proves adequate for reliability assessment, particularly when evaluating the most economical operating strategy that prevents component overloads.

We formulate the DC-OPF using optimization variables $\boldsymbol{\eta} \triangleq \{P_i^+, P_i^-, \theta_i\}_{i=1}^{n_b}$, where P_i^+ denotes power generation at each bus, P_i^- represents power demand at load buses, and θ_i indicates voltage angles at all buses. The standard DC-OPF problem is expressed as:

$$\min_{\boldsymbol{\eta}} \quad \mathcal{C}(\boldsymbol{\eta}) \tag{2}$$

$$\text{subject to} \quad \mathcal{G}(\boldsymbol{\eta}) = 0 \qquad \text{(Power balance)} \tag{3}$$

$$\mathcal{H}(\boldsymbol{\eta}) \leq \mathbf{0} \qquad \text{(Branch flow limits)} \tag{4}$$

$$\boldsymbol{\eta}^{\min} \leq \boldsymbol{\eta} \leq \boldsymbol{\eta}^{\max} \qquad \text{(Variable bounds)} \tag{5}$$

The power balance constraints $\mathcal{G}(\boldsymbol{\eta}) = 0$ ensure generation-load equilibrium, while $\mathcal{H}(\boldsymbol{\eta}) \leq \mathbf{0}$ enforces thermal limits on transmission branches. When exact branch ratings are unavailable, we assume each branch's flow limit equals twice its power flow in the intact network. The bounds $\boldsymbol{\eta}^{\min}$ and $\boldsymbol{\eta}^{\max}$ include a fixed reference bus angle (θ^{ref}) and operational limits for other variables. If a linear cost function is chosen, the optimization problem is linear and hence can be efficiently solved by various linear programming solvers. To this end, a positive constant cost c is associated with each unit of power loss, so that the cost function equals c multiplied by the total power loss as follows:

$$\mathcal{C}(\boldsymbol{\eta}) = c \cdot \left(P^{\text{dem}} - \sum_{i \in \Gamma_g} P_i^+ \right) \tag{6}$$

where the constant P^{dem} := total power demand in the intact network and Γ_g contains indices of all generators.

3 METHODOLOGY

3.1 Subset simulation and its limitations

Subset simulation (SuS) is a widely adopted rare-event sampling technique designed to efficiently estimate small failure probabilities in high-dimensional spaces [3]. Instead of directly sampling the rare failure event $\mathcal{F}$, SuS expresses the event as a sequence of more probable intermediate events:

$$\mathcal{F}_1 \supset \mathcal{F}_2 \supset \cdots \supset \mathcal{F}_m = \mathcal{F}, \tag{7}$$

where each $\mathcal{F}_k$ corresponds to a progressively stricter performance threshold. This decomposition enables efficient estimation by transforming the original rare-event probability into a product of conditional probabilities:

$$p_f = \mathbb{P}(\mathcal{F}_1) \prod_{k=2}^{m} \mathbb{P}(\mathcal{F}_k | \mathcal{F}_{k-1}). \tag{8}$$

Algorithm 1 Subset simulation (SuS) algorithm

Input: $p_0 \in (0, 1)$**, integer** N **multiple of** $\frac{1}{p_0}$

$l \leftarrow 0$, $b_l \leftarrow \infty$

while $b_l > 0$ **do**

  if $l = 0$ **then**

    Generate N samples $\{x_k\}_{k=1}^N$ from initial distribution $Q(\cdot|F_0)$

  else

    Generate N samples $\{x_k\}_{k=1}^N$ from $Q(\cdot|F_l)$ using MCMC with seeds $S^{(l)}$

  end if

  Sort $\{x_k\}_{k=1}^N$ by increasing $g(x)$ and denote as $\{\bar{x}_k\}_{k=1}^N$, $b_{l+1} \leftarrow g(\bar{x}_{p_0 \cdot N})$

  if $b_{l+1} \leq 0$ **then**

    $b_{l+1} \leftarrow 0$, $N_f \leftarrow \sum_{k=1}^{N} \mathbb{I}\{g(\bar{x}_k) \leq 0\}$

  end if

  $S^{(l+1)} \leftarrow \{\bar{x}_k\}_{k=1}^{p_0 \cdot N}$, $l \leftarrow l + 1$

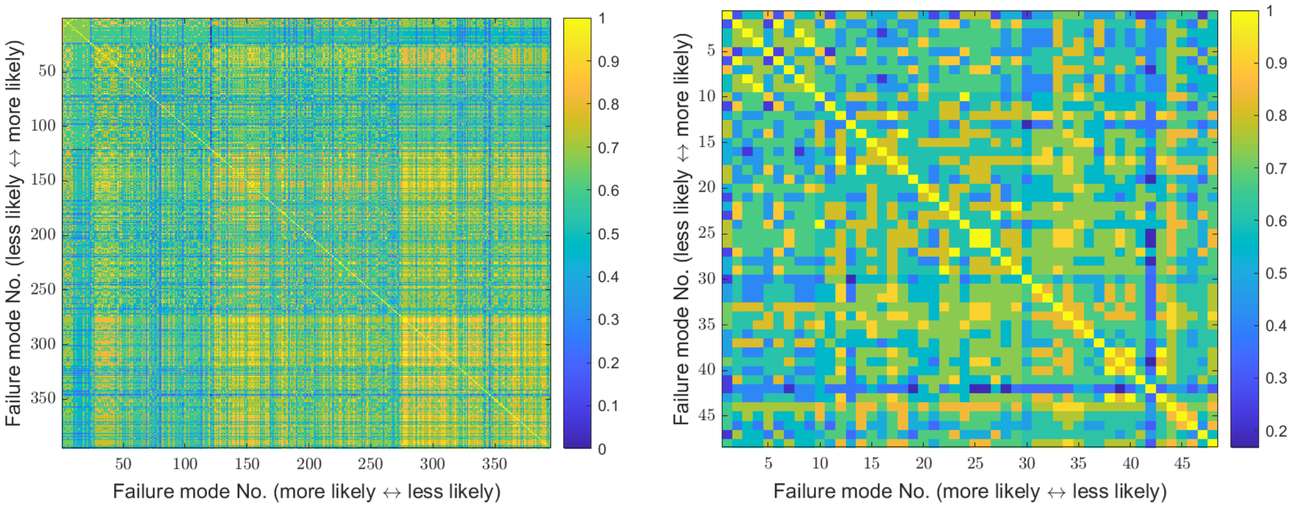
end while

Output: $\hat{p}_f = p_0^{l-1} \dfrac{N_f}{N}$

In SuS, an initial set of samples is drawn from the input distribution. A subset exceeding a prescribed performance threshold is selected, and these samples serve as seeds to generate subsequent samples via Markov Chain Monte Carlo (MCMC) methods, such as the Metropolis-Hastings (MH) algorithm. This sequential conditional sampling significantly reduces the number of samples required compared to crude Monte Carlo simulation. As summarized in Algorithm 1, SuS uses a conditional probability level $p_0 \in (0, 1)$, typically set to 0.1, which determines the fraction of samples selected as seeds for the next level. At level $l = 0$, SuS starts by sampling from the initial distribution $Q(\cdot \mid F_0)$, which is the unconditional distribution of the input random vector $\mathbf{X}$. In the context of the DC-OPF problem, $\mathbf{X}$ represents the states of network components, with F_0 being the entire sample space. At subsequent levels $l \geq 1$, SuS generates samples from the conditional distribution $Q(\cdot \mid F_l)$, where $F_l \triangleq \{\mathbf{X} : g(\mathbf{X}) < b_l\}$ is the intermediate failure event defined by the limit state function $g(\cdot)$ and threshold b_l.

Despite its success, SuS faces challenges when the failure domain is fragmented or contains multiple disconnected regions separated by low-probability barriers. In such cases, MCMC chains can become trapped in isolated regions, leading to poor exploration of the global failure domain and introducing errors into the failure probability estimator. Furthermore, poor chain mixing exacerbates estimation errors when failure regions are sparse or non-local.

As an illustrative example, a similarity analysis is conducted on the failure modes of power grid from the experiment Case 1-2 described later in Section 4.1. Using 10^8 Monte Carlo simulations, over 300 failure modes in Case 1 and over 40 in Case 2—each representing a unique combination of component states—are identified and sorted by likelihood of occurrence. Cosine similarity between all pairs of failure modes is calculated, and the resulting heatmaps are shown in Figure 1. The presence of "blue belts" in both cases reveals that some rare failures share little resemblance with any others. Such rare and isolated modes are difficult to discover through initial sampling of $Q(\cdot|F_0)$ due to their rarity, and even harder to reach during conditional samplings of $Q(\cdot|F_l)$, given the limited exploration capabilities of standard MCMC approaches.



(a) Case 1: $P_f = 1.1 \times 10^{-4}$

(b) Case 2: $P_f = 1.0 \times 10^{-5}$

Figure 1: Blue belts showing dissimilarity between failure modes in Case 1-2 (Section 4.1)

3.2 Skipping subset simulation

A Metropolis-class MCMC sampler, referred to as the skipping sampler, was proposed in [4, 5] to improve convergence rates by allowing a bounded number of consecutive rejections before terminating each proposal stage. According to [4], this approach preserves the desired stationary distribution while enhancing mixing and state space exploration and allows flexibility in the design of the proposal-rejection mechanism. By decoupling the acceptance process from the strict one-step update of classical Metropolis-Hastings, the skipping sampler allows chains to escape local modes more readily, particularly in high-dimensional or multi-modal distributions.

Inspired by this advancement, we incorporate the skipping mechanism into the subset simulation framework and propose skipping subset simulation (Skipping-SuS), which allows up to K consecutive rejections during the conditional sampling phase. The key idea is that by permitting a limited number of failed proposals, the MCMC chains gain increased freedom to explore disconnected or hard-to-reach failure regions, thus mitigating local trapping—a known limitation of standard SuS. Adding the skipping mechanism does not impair the core efficiency advantage of SuS. Instead, it improves accuracy by increasing the likelihood of identifying new and disconnected failure regions during the conditional sampling phase. The Skipping-SuS algorithm retains the main structure of subset simulation while enhancing its exploratory capability.

Algorithm 2 outlines the Skipping-SuS method, differing from standard SuS (Algorithm 1) solely in lines 6–18, where the skipping mechanism is implemented during MCMC sampling. The parameter

$K \geq 0$, a predefined integer, governs the maximum number of allowed skips: a larger K encourages broader exploration of the state space, while a smaller K results in behavior closer to standard SuS. Within the skipping loop, the variable $count$ tracks the number of consecutive rejections by incrementing each time a proposed candidate fails to satisfy the acceptance condition. If $count$ reaches K without an accepted candidate, the chain repeats the current state.

Algorithm 2 Skipping subset simulation (Skipping-SuS) algorithm

Input: $p_0 \in (0, 1)$**, integer** N **multiple of** $\frac{1}{p_0}$**, integer** K
$l \leftarrow 0$, $b_l \leftarrow \infty$
while $b_l > 0$ **do**
    if $l = 0$ **then**
        Generate N samples $\{x_k\}_{k=1}^{N}$ from initial distribution $Q(\cdot|F_0)$
    else
        for each seed s in $S^{(l)}$ **do**
            Generate $p_0 \cdot N$ samples $\{x_k\}_{k=1}^{p_0 \cdot N}$ from $Q(\cdot|F_l)$ using MCMC with seeds $S^{(l)}$
            for each sample **do**
                $count \leftarrow 0$, propose candidate
                while $count < K$ **do**
                    if accepted **then**
                        Move to candidate, **break**
                    else
                        $count \leftarrow count + 1$, propose candidate again from the current state
                    end if
                end while
            end for
        end for
    end if
    Sort $\{x_k\}$ by $g(x)$ and denote as $\{\bar{x}_k\}_{k=1}^{N}$, $b_{l+1} \leftarrow g(\bar{x}_{p_0 \cdot N})$
    if $b_{l+1} \leq 0$ **then**
        $b_{l+1} \leftarrow 0$, $N_f \leftarrow \sum_{k=1}^{N} \mathbb{I}\{g(\bar{x}_k) \leq 0\}$
    end if
    $S^{(l+1)} \leftarrow \{\bar{x}_k\}_{k=1}^{p_0 \cdot N}$, $l \leftarrow l + 1$
end while
Output: $\hat{p}_f = p_0^{l-1} \dfrac{N_f}{N}$

4 COMPUTATIONAL EXPERIMENT ON POWER GRID RELIABILITY ASSESSMENT

4.1 Experimental setup and software

The proposed Skipping-SuS is tested against standard SuS by conducting a computational experiment on DC-OPF-based reliability assessment with the IEEE-14 benchmark model. The objective is to estimate the probability that the percentage blackout size (PBS), defined as the percentage of load shed by DC-OPF, is above specified PBS thresholds in two experimental cases:

- Case 1: $P_f = 1.1 \times 10^{-4}$, $\text{PBS}_1 = 53.5\%$

- Case 2: $P_f = 1.0 \times 10^{-5}$, $\text{PBS}_2 = 57.8\%$

The optimization problem in Eqs. (2)–(5) with cost function in Eq. (6) equivalently quantifies the minimum power loss $l_p^{(\min)}(\mathbf{x})$ associated with the state of network components $\mathbf{x}$. The percentage blackout size is calculated as:

$$\text{PBS}(\mathbf{x}) \triangleq \frac{l_p^{(\min)}(\mathbf{x})}{P^{(\text{dem})}} \cdot 100\%. \tag{9}$$

The experiment uses the component failure probabilities specified in Table 1 for the IEEE-14 benchmark system. For each case, 200 independent runs are performed per method to ensure statistical significance. Both Skipping-SuS and standard SuS employ adaptive conditional sampling (aCS) [8] as the MCMC algorithm for conditional sampling, along with adaptive effort subset simulation (aE-SuS) [9–11] to avoid ambiguous definitions of intermediate failure events. The conditional probability level p_0 is set to 0.1 in both methods. In Skipping-SuS, the maximum number of allowed skips is fixed at $K = 1$ to evaluate the impact of the skipping mechanism.

Table 1: Failure probability of IEEE-14 model components

damage / type	**Complete (100%)**	**Major (60%)**	**Minor (20%)**	**Negligible (0%)**
Generator	0.01	0.19	0.30	0.50
Connecting bus	0.01	/	/	0.99
Transm. line	0.01	/	/	0.99

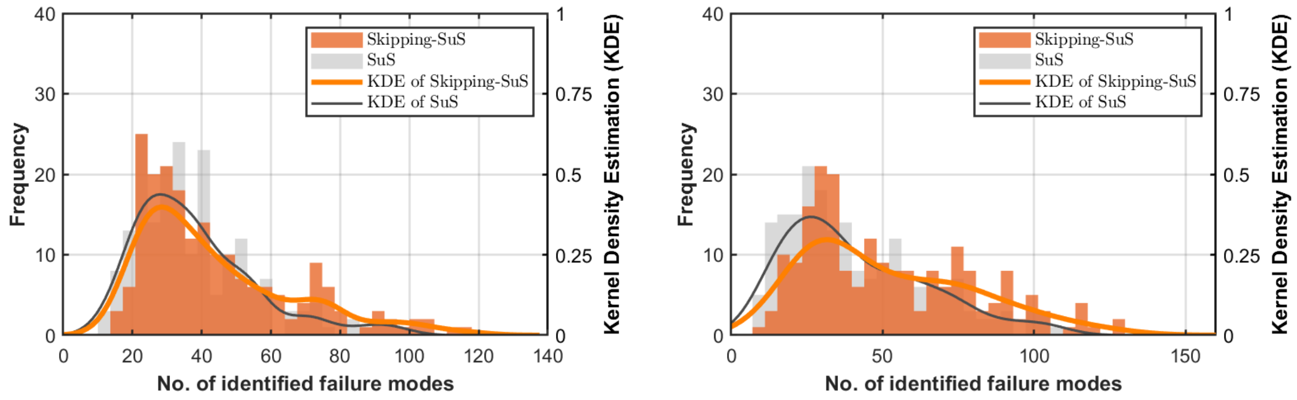
To ensure a fair comparison, both methods are evaluated under identical computational budgets, defined by the total number of limit-state function evaluations $g(\mathbf{X})$: 5000 evaluations for Case 1 and 7000 for Case 2. This total includes all function evaluations incurred during the skipping process in Skipping-SuS. Since the sample size per level (N, in both Algorithm 1 and Algorithm 2) is the only directly controllable parameter, we adjusted N iteratively until the average computational cost across 200 runs was equalized between the two methods.

All simulations were conducted in MATLAB R2024a using the DC-OPF solver provided in MATPOWER version 7.1 [12].

4.2 Results and Discussions

Figure 2 shows the distribution of numbers of unique failure modes identified with Skipping-SuS and SuS. In Case 1, Skipping-SuS identifies slightly more failure modes than SuS, with the difference becoming more pronounced in Case 2. This trend highlights the enhanced global exploration capability of Skipping-SuS, which is attributed to the introduction of the skipping mechanism during the conditional sampling phase, as anticipated.

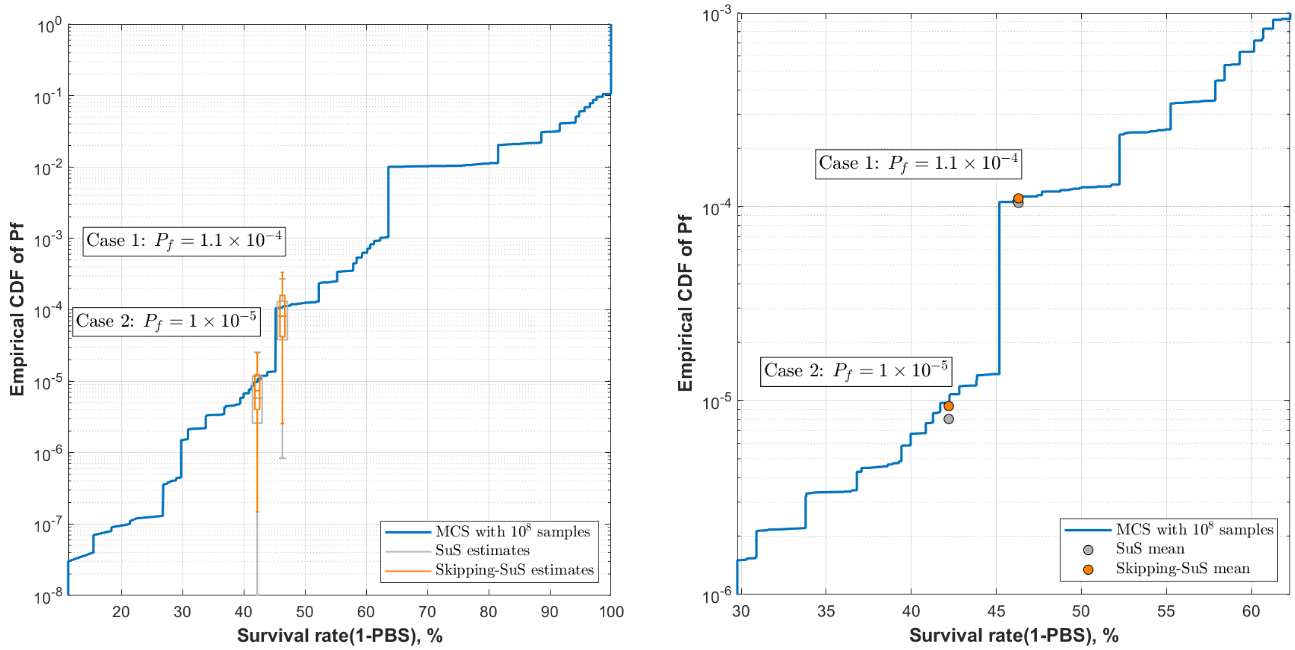
Figure 3 compares the failure probability estimates obtained by the two methods. Across all key point estimators—including variance, median, and mean—Skipping-SuS consistently outperforms SuS in both Case 1 and Case 2. In line with the observed improvements in the number of identified failure modes, the performance gap is notably larger in Case 2, demonstrating the advantage of Skipping-SuS in the more challenging reliability estimation problems.



(a) Case 1: $P_f = 1.1 \times 10^{-4}$

(b) Case 2: $P_f = 1.0 \times 10^{-5}$

Figure 2: Number of unique failure modes identified with Skipping-SuS and SuS.



(a) Boxplots of estimates

(b) Mean estimates

Figure 3: Failure probability (P_f) estimation with Skipping-SuS and SuS.

5 CONCLUSION AND FUTURE WORK

This paper introduced skipping subset simulation (Skipping-SuS), an enhanced variant of standard subset simulation, aimed at improving rare-event probability estimation in fragmented failure domains expected in power systems. By incorporating a skipping mechanism into the conditional sampling phase, Skipping-SuS enables more effective exploration of disjoint failure regions without sacrificing the computational efficiency of the original method. Computational experiments on an IEEE benchmark power system involving sparse and non-local failure regions demonstrate that Skipping-SuS consistently identifies more failure modes and achieves more accurate and robust reliability estimates compared to standard SuS.

This work represents an initial step in exploring the effect of integrating skipping mechanisms into subset simulation. Future research will focus on investigating alternative skipping strategies as well as the impact of increasing the maximum allowed number of skips, denoted as K, on sampling performance and estimator quality. In particular, dynamically adjusting K based on online performance indicators during simulation could further enhance the robustness and efficiency of the method. Additional extensions also involve applying Skipping-SuS to other classes of rare-event problems, such as those with continuous input spaces or highly nonlinear failure surfaces, to further validate its generality and effectiveness. Moreover, the improved accuracy of Skipping-SuS estimates makes them particularly well-suited for supporting decision analyses in power system reliability, where varying risk attitudes necessitate minimizing estimation errors to ensure informed and reliable decision-making.

6 ACKNOWLEDGMENTS

The first and the second authors are grateful for the U.S. National Science Foundation (NSF) Grant No. CMMI-2037545 and the National Institute for Testing and Technology (NIST) (Award No. 70NANB15H044).

References

[1] Dueñas-Osorio, L., Meel, K., Paredes, R. & Vardi, M. Counting-Based Reliability Estimation for Power-Transmission Grids. *Proc. AAAI Conf. Artif. Intell.* (2017) **31**(1). doi:10.1609/aaai.v31i1.11178.

[2] Paredes, R., Talebiyan, H. & Dueñas-Osorio, L. Path-dependent reliability and resiliency of critical infrastructure via particle integration methods. *Proc., 13th Int. Conf. on Structural Safety & Reliability*. Shanghai, China: Tongji Univ., 2022.

[3] Au, S.-K. and Beck, J.L. Estimation of small failure probabilities in high dimensions by subset simulation. *Probab. Eng. Mech.* (2001) **16**:263–277. doi:10.1016/S0266-8920(01)00019-4.

[4] Moriarty, J., Vogrinc, J. and Zocca, A. A Metropolis-class sampler for targets with non-convex support. *Stat. Comput.* (2021) **31**:72. doi:10.1007/s11222-021-10044-4.

[5] Goodridge, M.P., Lakshminarayana, S. and Zocca, A. Uncovering Load-Altering Attacks Against $N-1$ Secure Power Grids: A Rare-Event Sampling Approach. *IEEE Trans. Power Syst.* (2025) **40**:1269–1281. doi:10.1109/TPWRS.2024.3419725.

[6] Wood, A. J. & Wollenberg, B. F. *Power Generation, Operation, and Control*. Wiley, 1983.

[7] Billington, R. and Li, W. *Reliability assessment of electric power systems using Monte Carlo methods*. Springer Science & Business Media, 1994.

[8] Papaioannou, I., Betz, V., Zwirglmaier, K. and Straub, D. MCMC subset simulation. *Probab. Eng. Mech.* (2015) **41**:89–103. doi:10.1016/j.probengmech.2015.06.006.

[9] Chan, J., Papaioannou, I., Straub, D. An adaptive subset simulation algorithm for system reliability analysis with discontinuous limit states. *Reliab. Eng. Syst. Saf.* (2022) **225**:108607. doi:10.1016/j.ress.2022.108607.

[10] Chan, J., Paredes, R., Papaioannou, I., Dueñas-Osorio, L., Straub, D. A comparative study on adaptive Monte Carlo methods for network reliability assessment. *Proc., 14th Int. Conf. on Application of Statistics and Probability in Civil Engineering*. Belfast, Northern Ireland: Queen's Univ. Belfast, 2023.

[11] Chan, J., Paredes, R., Papaioannou, I., Dueñas-Osorio, L. and Straub, D. Adaptive Monte Carlo Methods for Estimating Rare Events in Power Grids. *ASCE-ASME J. Risk Uncertain. Eng. Syst., Part A: Civ. Eng.* (2025) **11**:04024082. doi:10.1061/AJRUA6.RUENG-1404.

[12] Zimmerman, R.D., Murillo-Sanchez, C.E. and Thomas, R.J. MATPOWER: steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Trans. Power Syst.* (2010) **26**:12–19.