A Comparative Study on Adaptive Monte Carlo Methods for Network Reliability Assessment

Jianpeng Chan
Ph.D. student, TUM School of Engineering and Design, Technical University of Munich, Munich, Germany

Roger Paredes
Research associate, Department of Civil and Environmental Engineering, Rice University, Houston, USA

Iason Papaioannou
Senior lecturer, TUM School of Engineering and Design, Technical University of Munich, Munich, Germany

Leonardo Duenas-Osorio
Professor, Department of Civil and Environmental Engineering, Rice University, Houston, USA

Daniel Straub
Professor, TUM School of Engineering and Design, Technical University of Munich, Munich, Germany

ABSTRACT: Infrastructure networks, such as electrical power grids, transportation and water supply systems, support critical societal functions of society. Failures of such networks can have severe consequences, and quantification of the probability of failure of such systems is essential for understanding and managing their reliability. Analytical and simulation methods have been proposed to solve such kinds of problems, among which sampling methods feature prominently. Recently, the authors extended widely used structural reliability algorithms, subset simulation, cross-entropy-based importance sampling as well as uncertainty quantification methods built from particle integration methods and exact confidence, all for efficient reliability analysis in discrete spaces. This paper tests the performance of these algorithms for static network reliability assessment. In particular, we compare these methods for optimal power flow problems in various IEEE benchmark models. Overall, the cross-entropy-based method outperforms the other methods in all benchmark models except the largest IEEE 300, while the adaptive effort subset simulation and particle integration methods are more suitable for handling high-dimensional problems. By building up the benchmark models, we provide unified examples for comparing different emerging methods in static network reliability assessment and also to support improvement or combination of these methods.

1. INTRODUCTION

In network reliability assessment, one fundamental problem is to compute or estimate the failure probability p_f, the probability that the system performance is less or equal than a specified threshold. The system performance can be described by a limit state function (LSF) $g(\cdot)$, with failure defined as $F = \{g(\boldsymbol{X}) \leq 0\}$. In particular, let $\boldsymbol{X}$ describe the state of network components, whose probability mass function (PMF) is $p_{\boldsymbol{X}}(\boldsymbol{x}), \boldsymbol{x} \in \Omega_{\boldsymbol{X}}$, and let $\mathbb{I}\{\cdot\}$ denote the indicator function that takes value one when the statement inside the braces is true and zero otherwise. The failure probability can then be written as

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

When quantifying the reliability of a power system, $\boldsymbol{X}$ can represent the damage state of transmission lines and/or connecting buses, and the system performance can be either connectivity or power flow. In such settings, $\boldsymbol{X}$ usually contains discrete random variables, which can result in a LSF with discontinuous distribution as depicted in Figure 2.

A set of efficient non-sampling methodologies are applicable (Li and He, 2002; Hardy et al., 2007; Paredes et al., 2019), but many of them rely on specific assumptions such as binary state, independent components, perfect nodes, coherent systems, etc., that limit their generality. Other popular non-sampling-based methods such as the matrix-based system reliability assessment (Byun and Song, 2021) are restricted to small or moderate number of components.

Sampling-based methods, including crude Monte Carlo simulation (MCS) (Zio, 2013) and its different variants, trade efficiency, accuracy for their wide applicability. These methods are often non-intrusive and treat the network model as a black box, which facilitates using advanced network models in analysis. Nevertheless, the accuracy of these methods clearly depends on the number of samples and also the problems at hand, which highlights the importance of testing different emerging algorithms in a unified benchmark.

For rare event estimation, crude MCS is infeasible when the limit state function is expensive to compute, and hence, advanced variance reduction techniques or meta models have been proposed. These techniques include the standard subset simulation (SuS) (Zio and Pedroni, 2008; Zuev et al., 2015; Jensen and Jerez, 2018), various creation process embedded methods (Hui et al., 2005; Botev et al., 2013; Cancela et al., 2019), and actively trained meta models (Cadini et al., 2017; Dehghani et al., 2021).

Recently, Chan et al. (2022a,b,c) generalized two widely used structural reliability algorithms, the standard SuS method (Au and Beck, 2001) and improved cross entropy (iCE) method (Papaioannou et al., 2019), for sampling efficiently in discrete space. Concurrently, Paredes et al. (2019, 2022) built upon particle integration methods (PIMs) (Del Moral, 2013) and Gamma Bernoulli approximation scheme (GBAS) (Huber, 2017) for a priori accountable estimates.

The present paper focuses on testing the performance of these algorithms for rare event estimation in static networks. In particular, we consider the application to the optimal direct current (DC) power flow of IEEE benchmark models, which is used in reliability-based design, as well as in contingency analyses.

The remainder of the paper is organized as follows. A brief introduction of the optimal power flow problem is given in Section 2. In Section 3, we introduce the methodologies employed in this paper. A comparative study is then conducted in Section 4.

2. OPTIMAL POWER FLOW PROBLEM

The power flow in power transmission networks is driven by Kirchhoff's law and various operation strategies. While alternating current (AC) models can more accurately represent the power flow, their solution is computationally challenging due to the need for iterative solves and they require a number of inputs that are not generally available when doing the reliability analysis. For this reason, we use direct current (DC) power flow models, which approximate the flow through solving a linear equation set that is absolutely convergent, in which the net reactive power injection Q_i and voltage magnitude V_i at each bus are neglected. The results are less accurate than those of AC for transient analyses, but adequate for system reliability. In addition, instead of modelling cascading failures, we focus on the standard DC optimal power flow (DC-OPF) problems, where we compute the optimal power operation strategy that avoids network component failure and at the same time minimizes a specified cost function $\mathscr{C}(\cdot)$.

Specifically, let $\boldsymbol{\eta} \triangleq \{P_i, \theta_i\}_{i=1}^{n_b}$ collect the voltage angle θ_i and net active power injection P_i at each of the n_b buses, so as to formulate the standard DC-OPF as follows:

$$\begin{aligned} &\min_{\boldsymbol{\eta}} \mathscr{C}(\boldsymbol{\eta}) \\ &s.t. \mathscr{G}(\boldsymbol{\eta}) = \mathbf{0} \\ &\quad \mathscr{H}(\boldsymbol{\eta}) \leqslant \mathbf{0} \\ &\quad \boldsymbol{\eta}^{(min)} \leqslant \boldsymbol{\eta} \leqslant \boldsymbol{\eta}^{(max)} \end{aligned} \tag{2}$$

where equality constraints in $\mathscr{G}(\boldsymbol{\eta}) = \mathbf{0}$ represent the active power balance equations, and inequality constraints in $\mathscr{H}(\boldsymbol{\eta}) \leqslant \mathbf{0}$ result from the branch flow limits, i.e., the power flow over any branch is always below its capacity. The limits $\boldsymbol{\eta}^{(min)} \leqslant \boldsymbol{\eta} \leqslant \boldsymbol{\eta}^{(max)}$ include an equality constraint on the voltage angle of the reference bus $\theta^{(ref)}$ and lower and upper bound for other variables in $\boldsymbol{\eta}$. 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 to each unit of the power loss, and the cost function equals the constant c multiplied by the total power loss $l_p(\boldsymbol{\eta})$, i.e.,

$$\mathscr{C}(\boldsymbol{\eta}) \triangleq c \cdot l_p(\boldsymbol{\eta}) = c \cdot \left(P^{(dem)} - \sum_{i \in \Gamma_g} P_i \right), \tag{3}$$

where the constant $P^{(dem)}$ represents the total power demand, and Γ_g collects the index of all generator buses.

3. BRIEF SUMMARY OF THE ADAPTIVE MONTE CARLO METHODS

In this subsection, we give a brief introduction of the methods investigated in the comparative study.

3.1. Particle integration methods

Particle Integration Methods (PIMs) consist of sequential systems of samples, or particles, for approximating intractable integrals, such as the system failure probability in Eq. (1). Standard SuS can be regarded as one of these (interacting) particle systems (Del Moral, 2013), where the failure probability is represented as a sequence of products $p_f = p_0 p_1 \cdots p_k$, with the common choice of a fixed probability $p_0 = p_i$, for all $i < k$. The main ingredients of a PIM are: 1) an initial distribution, i.e., the probability distribution $p_{\mathbf{X}}$ of the input random vector $\mathbf{X}$; 2) a sequence of probability kernels, or MCMC samplers; and 3) a score function (or importance function) that is intimately related to the LSF. Typically, PIMs can be grouped as interacting (or adaptive-levels) methods and non-interacting (or fixed-levels) methods. Interacting methods can adaptively construct the sequence of products representation of the failure probability but they can be heavily biased, especially when the LSF presents discontinuities. In contrast, non-interacting particle methods are guaranteed to be consistent and unbiased estimators of the true failure probability; however, they assume the sequence of products representation to be known. Naturally, as suggested by Botev and Kroese (2012), one can combine the approaches as a two-step meta-algorithm where the interacting PIM (iPIM) is run first to learn the sequence of products, and then the non-interacting PIM is run second using the sequence of products and as an unbiased and consistent estimator of the true failure probability. We adopt the iPIM in Algorithm 2 of Paredes et al. (2022), which is a biased estimator similar to SuS. After this, we use the annealed PIM (aPIM), which is an unbiased estimator that takes the form of

$$\hat{p}_f^{(\text{aPIMs})} = \left(\frac{1}{p_0} \right)^{k-1} |\mathscr{X}_k|, \tag{4}$$

where $\mathscr{X}_k$ is the set of particles in the last level; see Algorithm 1 of (Paredes et al., 2022) for implementation details. Under certain conditions, the authors proposed an optimal tuning of such meta-algorithm by setting $p_0 = 0.2032$ and showing that popular MCMC samplers such as the preconditioned Crank-Nicolson and modified Metropolis-Hastings algorithms can scale well in high-dimensional problems with tens of thousands of random variables (Paredes et al., 2022).

3.2. Adaptive effort subset simulation method

Au and Wang (2014) identify issues associated with employing the standard SuS for LSFs with a discrete cumulative distribution function (CDF), where a fixed intermediate failure probability p_0 can lead to an ambiguous definition of the intermediate failure domains and hence inaccurate results. Chan et al. (2022a) address the problem in the context of network reliability assessment and propose the adaptive effort SuS (aE-SuS) algorithm that tackles this issue through adaptively choosing the intermediate failure probability p_0 and the number of samples per level N. aE-SuS can be combined with any Markov chain Monte Carlo (MCMC) algorithm, e.g., the adaptive sampling algorithm (essentially an adaptive variant of the preconditioned Crank-Nicolson algorithm (Papaioannou et al., 2015)), ideally suited for high-dimensional inputs, an independent Metropolis-Hasting algorithm that efficiently samples in low-dimensional discrete spaces (Chan et al., 2022a), or a Gibbs sampler for performing the reliability analysis conditional on data (Zwirglmaier et al., 2023). The algorithm starts with an initial choice of p_0 and N so that at least $N \cdot p_0 \cdot tol$ seeds (or failure samples) are obtained at each level, where tol is a prescribed hyperparameter. This leads to an adaptive estimate of the intermediate failure probabilities in terms of the failure samples and total number of samples at the respective level.

When the intermediate level and the MCMC chain length is predefined i.e., independent of the sampling process, SuS corresponds to the fixed effort generalized splitting method and is theoretically unbiased (Botev and Kroese, 2012). Although this does not hold for aE-SuS, we still find in practice that the bias is minor compared to the variance of the p_f estimator. However, the performance of the algorithm heavily depends on the choice of the MCMC algorithm. If an inappropriate MCMC is selected, the final estimator is highly skewed and estimating the mean and the variance of a highly skewed distribution is challenging.

3.3. Bayesian improved cross entropy method

The cross entropy method is an adaptive importance sampling (IS) method for rare event estimation. The method determines the IS distribution chosen adaptively through successively approximating a sequence of intermediate target distributions that gradually approaches the optimal IS distribution $p^*(\cdot) \propto p_{\boldsymbol{X}}(\cdot)\mathbb{I}\{g(\cdot) \leqslant 0\}$.

There are different ways of designing the intermediate target distributions. Our approach is provided by the iCE method (Papaioannou et al., 2019), an improved version of the standard cross entropy method (Rubinstein, 1997). iCE defines the sequence by smoothing the indicator function $\mathbb{I}\{\cdot\}$ in $p^*(\cdot)$ via the standard normal CDF $\Phi(\cdot)$, i.e.,

$$p_{\boldsymbol{X}}^{(t)}(\cdot) \propto p_{\boldsymbol{X}}(\cdot)\Phi\left(-\frac{g(\cdot)}{\sigma^{(t)}}\right), \, t = 1, ..., T \tag{5}$$

where $\sigma^{(t)}$ is the scaling parameter.

Given a parametric family $h(\cdot;\boldsymbol{v})$, the iCE method iteratively determines the distribution in $h(\cdot;\boldsymbol{v})$, or equivalently the parameter vector $\boldsymbol{v}$, through minimizing an estimate of its Kullback–Leibler divergence from $p^{(t)}(\cdot)$. This leads to the successive solution of the following optimization problem:

$$\widehat{\boldsymbol{v}}^{(t)} = \arg\max_{\boldsymbol{v}} \sum_{k=1}^{N} W_k^{(t)} \ln(h(\boldsymbol{x}_k;\boldsymbol{v})) \tag{6}$$

$$W_k^{(t)} \triangleq \frac{p_{\boldsymbol{X}}(\boldsymbol{x}_k)\Phi\left(-\frac{g(\boldsymbol{x}_k)}{\sigma^{(t)}}\right)}{h(\boldsymbol{x}_k;\widehat{\boldsymbol{v}}^{(t-1)})}, \;\; \boldsymbol{x}_k \sim h(\cdot;\widehat{\boldsymbol{v}}^{(t-1)})$$

with $h(\boldsymbol{x};\widehat{\boldsymbol{v}}^{(0)}) = p_{\boldsymbol{X}}(\boldsymbol{x})$.

$\sigma^{(t)}$ is chosen adaptively such that the effective sample size of the weighted data $\{\boldsymbol{x}_k, W_k^{(t)}\}_{k=1}^N$ is approximately equal at each level. One can prove that $\sigma^{(t)}$ decreases monotonically when the input distribution $p_{\boldsymbol{X}}(\cdot)$ is discrete and the intermediate target distributions $p_{\boldsymbol{X}}^{(t)}(\cdot), t = 1, ..., T$ can be perfectly restored from the parametric family, i.e., $h(\cdot;\widehat{\boldsymbol{v}}^{(t)}) = p_{\boldsymbol{X}}^{(t)}(\cdot), t = 1, ..., T$ (Chan et al., 2022c).

Eq. (6) indicates that $\widehat{\boldsymbol{v}}^{(t)}$ is the weighted maximum likelihood estimation of $\boldsymbol{v}$ given the data set $\{\boldsymbol{x}_k, W_k^{(t)}\}_{k=1}^N$. Hence, $\widehat{\boldsymbol{v}}^{(t)}$ may suffer from overfitting when the sample size N is small. Chan et al. (2022b) circumvent this issue through introducing Bayesian statistics in iCE and propose the

Bayesian iCE (BiCE) method. Specifically for network reliability assessment, they introduce a symmetric Dirichlet prior for the independent categorical distribution and substitute the weighted maximum likelihood estimation $\widehat{\boldsymbol{v}}^{(t)}$ with the weighted posterior predictive estimate, or weighted maximum a posteriori estimate. To further consider the dependence among network components, Chan et al. (2023) employ a more flexible categorical mixture, where the weighted maximum a posteriori can be efficiently approximated through a generalized expectation-maximization algorithm. The BiCE method is proved to be theoretical unbiased.

In addition to the parameters required by the standard iCE (δ_{tar} and N), BiCE introduces another parameter b that accounts for the 'strength' of the prior. In addition, the number of clusters in the mixture, K, should also be chosen in advance if the categorical mixture model is employed.

4. A COMPARATIVE STUDY

The methodologies introduced in Section 3 are tested for solving the DC-OPF problem in 5 IEEE benchmark models by using MATPOWER v7.1 (Zimmerman et al., 2010).

The objective is to estimate the probability that the percentage blackout size, that is the percentage of load shed by DC-OPF, is below a specified threshold thr. Note that the optimization problem in Eq.(2) with cost function in Eq.(3) equivalently quantifies the minimum power loss $l_p^{(min)}(\boldsymbol{x})$ associated with the state of network components $\boldsymbol{x}$, and the percentage blackout size of the network $PBS(\boldsymbol{x})$ can be calculated as:

$$PBS(\boldsymbol{x}) \triangleq \frac{l_p^{(min)}(\boldsymbol{x})}{P^{(dem)}} \cdot 100. \tag{7}$$

For each of the generator buses, we consider 4 damage states, namely negligible, minor, major and complete damage, which correspond to $0\%, 20\%, 60\%$ and 100% reduction of the power production, respectively. The remaining non-generator buses and all transmission lines have two damage states, either safe or failure.

We estimate the true failure probability p_f using a rigorous (ε, δ)-approximation method, i.e., a randomized approximation that has relative error at most ε with probability at least $1-\delta$, where both parameters $\varepsilon, \delta \in (0,1)$ are chosen by the user. For example, when $p_f = 10^{-4}$ and a user selects $\varepsilon = 5\%$ and $\delta = 1\%$, an (ε, δ)-approximation returns a sample value $\bar{p}_f$ in the range $[0.95 \times 10^{-4}, 1.05 \times 10^{-4}]$ with probability 99%. In this paper, we use the GBAS, an efficient (ε, δ)-approximation method introduced by Huber (2017), as the baseline for comparison and select $\varepsilon = 5\%$ and $\delta = 1\%$. In the GBAS, for the chosen values of ε and δ, the sample size is a random variable with a mean value of roughly $2662/p_f$. The results along with the detailed problem settings are reported for every network case in Table 1. In addition, the distribution of different network components is summarized in Table 2.

Table 1: Benchmark settings

	# nodes	# lines	LSF($g(\boldsymbol{x})$)	$\bar{p}_f (\times 10^{-4})$
IEEE 14	14	20	$40 - PBS(\boldsymbol{x})$	6.4
IEEE 30	30	41	$29 - PBS(\boldsymbol{x})$	6.9
IEEE 57	57	80	$40 - PBS(\boldsymbol{x})$	7.4
IEEE 118	118	186	$10 - PBS(\boldsymbol{x})$	7.0
IEEE 300	300	411	$12.5 - PBS(\boldsymbol{x})$	6.1

Table 2: Distribution of network components

prob. \ state / type	complete	major	minor	negligible
generator	0.01	0.19	0.3	0.5
non-generator	0.01	/	/	0.99
trans. line	0.01	/	/	0.99

We compare different methods by their relative efficiency with respect to the crude MCS. The concept is borrowed from statistics and the relative efficiency is defined as

$$\text{relEff}(\widehat{p}_f) \triangleq \frac{p_f \cdot (1-p_f)}{\text{MSE}(\widehat{p}_f) \times \text{Cost}(\widehat{p}_f)}, \tag{8}$$

where $\text{MSE}(\widehat{p}_f)$ represents the mean square error of the failure probability estimator $\widehat{p}_f$, and the cost of an algorithm is measured by the number of evaluations of the LSF. The relative efficiency of crude MCS is equal to one; the larger the relative efficiency of an estimator, the more efficient it is relative to MCS.

For aE-SuS, we employ the adaptive conditional sampler and set $N = 2,000, p_0 = 0.1, tol = 0.8$. For BiCE, we use $N = 2,000, \delta_{tar} = 1.5, b = 10$ and K equals either 1 or 10. For PIMs, we select $N = 2,000$, $p_0 = 0.25$, adopt the preconditioned Crank-Nicolson method as the MCMC sampler, and use the score function $thr/PBS(\mathbf{x})$. One single run of the iPIM is employed as the pilot run for fixing the intermediate levels in all 200 repetitions of aPIMs.

We report the relative efficiency computed from 200 independent runs of aE-SuS, BiCE, iPIM and aPIM in Table 3. Note that for unbiased estimators, such as aPIM and BiCE, the average values of the 200 repetitions can be used to improve the convergence of estimates to the true probability of failure. The boxplot of the failure probabilities by the four algorithms is also depicted in Figure 1.

Table 3: Relative efficiency of different sampling-based methods

	aE-SuS	BiCE(K=1)	BiCE(K=10)	iPIM	aPIM
IEEE 14	1.71	2.29	4.29	0.71	0.97
IEEE 30	1.82	3.26	4.05	0.79	0.94
IEEE 57	2.00	18.46	18.51	1.48	1.69
IEEE 118	1.99	5.52	2.41	1.87	1.93
IEEE 300	1.44	0.67	0.84	1.39	1.48

We can see that the BiCE method outperforms aE-SuS with a higher relative efficiency in all IEEE benchmarks except the IEEE 300, which is due to the correlation among samples generated in aE-SuS. For the IEEE 300 benchmark, there are in total 711 components, and performing BiCE in such high dimensions, just like other adaptive IS methods, often leads to the degeneration of the weights (Kroese et al., 2013) and hence to a poor result. The efficiency of the BiCE method is even lower than that of the crude MCS in IEEE 300.

In contrast, aE-SuS is more robust since the adaptive conditional sampler employed in aE-SuS is specially designed for tackling high-dimensional problems (Papaioannou et al., 2015). The mean value of the empirical CDF of $g(\boldsymbol{X})$ obtained with aE-SuS, together with the 10 and 90 percentile, is shown in Figure 2, which agrees well with that of crude MCS with 10^6 samples. In addition, the simulation results demonstrate that a large number of clusters K in the categorical mixture leads to increased performance in low-dimensional systems but results in a less efficient estimator than simply employing a single categorical distribution in BiCE as the dimension increases.

PIMs did not lead to a significant variance reduction; however, their performance tends to be better than MCS in larger benchmarks IEEE 57, IEEE 118 and IEEE 300. We attribute this to the smaller jump discontinuities in the larger networks (see Figure 2 introduced next), which allows for a proper representation of the failure probability as sequence of products $p_0 p_1 \cdots p_k$. Also, for the larger benchmarks, the unbiased estimator, aPIM, was more efficient than the biased estimator, iPIM, and is the most efficient in the largest benchmark, the IEEE 300.

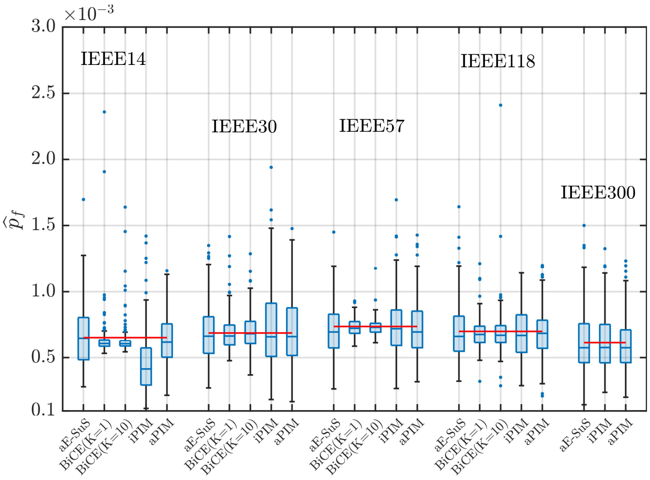

Figure 1: Performance of different sampling-based methods (the red solid line represents the GBAS estimate for each benchmark)

5. CONCLUSIONS

We test the performance of four recently developed sampling-based methods, namely the adaptive effort subset simulation (aE-SuS), Bayesian improved cross entropy method (BiCE) and particle integration methods (PIMs), either interacting or annealed, for rare event estimation in static networks. In particular, we compare these methods for optimal direct current optimal power flow problems in different IEEE benchmark models, whose dimension ranges from dozens to several hundreds. The aE-SuS method shows an efficiency

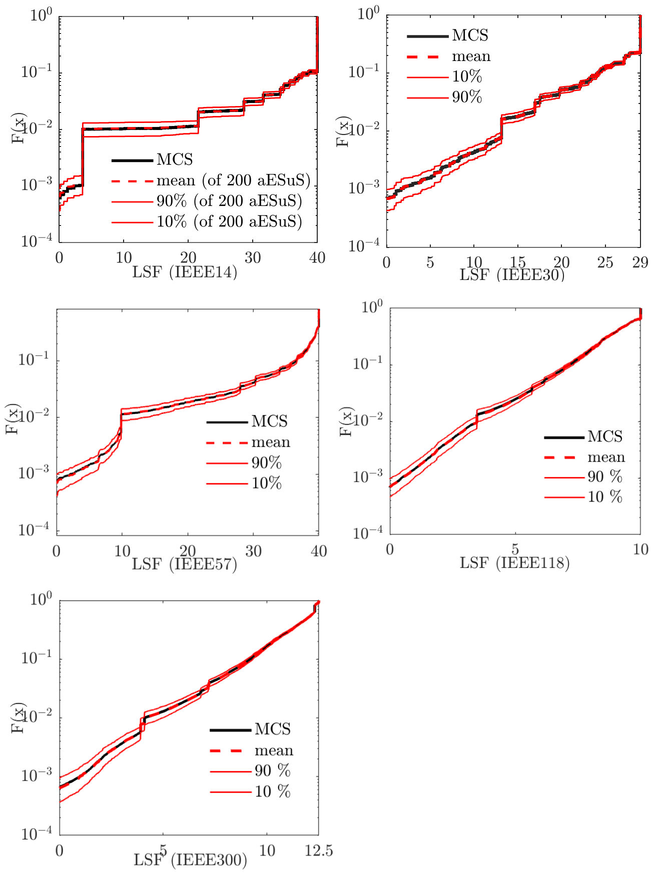

Figure 2: Empirical CDF from the aE-SuS method

that is higher than crude Monte Carlo simulation in all benchmarks. The BiCE method is the most efficient method in the benchmarks with low to moderate dimensions but performs poorly in the IEEE300 benchmark due to the degeneration of the IS weights. Here, its efficiency is less than that of crude Monte Carlo simulation. PIMs perform poorly in the low-dimensional benchmarks, but performs well in the large-scale applications where its performance is comparable to SuS. Overall, while BiCE appears to be the method of choice for smaller systems and provides a significant improvement over crude MCS, there is still room for more efficient methods to handle large-dimensional problem settings. SuS and PIM are both suitable candidates that should be further enhanced to provide improved efficiency in these cases.

For future work, note that the PIMs in this article used the same conditional failure probability value, p_0, for all levels. This is the optimal strategy for continuous score functions and failure probabilities that can be represented as $p_f = (p_0)^k$; however, for discontinuous score functions, such as the ones in this paper, a promising avenue would be to integrate the adaptive SuS ideas from Chan et al. (2022a) within PIMs. Additional future directions include the make-up of entrance states, the variance reduction possibilities of PIMs with cross entropy, etc.

6. ACKNOWLEDGMENTS

The first author gratefully acknowledges the financial support of the China Scholarship Council, and the second and the forth authors are grateful for the U.S. National Science Foundation (NSF) grant CMMI-2037545, and the National Institute for Testing and Technology (NIST) award F 70NANB15H044.

7. REFERENCES

Au, S.-K. and Beck, J. L. (2001). “Estimation of small failure probabilities in high dimensions by subset simulation.” *Probabilistic engineering mechanics*, 16(4), 263–277.

Au, S.-K. and Wang, Y. (2014). *Engineering risk assessment with subset simulation*. John Wiley & Sons.

Botev, Z. I. and Kroese, D. P. (2012). “Efficient Monte Carlo simulation via the generalized splitting method.” *Statistics and Computing*, 22(1), 1–16.

Botev, Z. I., L’Ecuyer, P., Rubino, G., Simard, R., and Tuffin, B. (2013). “Static network reliability estimation via generalized splitting.” *INFORMS Journal on Computing*, 25(1), 56–71.

Byun, J.-E. and Song, J. (2021). “Generalized matrix-based bayesian network for multi-state systems.” *Reliability Engineering & System Safety*, 211, 107468.

Cadini, F., Agliardi, G. L., and Zio, E. (2017). “Estimation of rare event probabilities in power transmission networks subject to cascading failures.” *Reliability Engineering & System Safety*, 158, 9–20.

Cancela, H., Murray, L., and Rubino, G. (2019). “Efficient estimation of stochastic flow network reliability.” *IEEE Transactions on Reliability*, 68(3), 954–970.

Chan, J., Papaioannou, I., and Straub, D. (2022a). "An adaptive subset simulation algorithm for system reliability analysis with discontinuous limit states." *Reliability Engineering & System Safety*, 108607.

Chan, J., Papaioannou, I., and Straub, D. (2022b). "Bayesian improved cross entropy method for network reliability assessment, <https://arxiv.org/abs/2211.09542.>.

Chan, J., Papaioannou, I., and Straub, D. (2022c). "Improved cross entropy-based importance sampling for network reliability assessment." *Proceedings of the 13th International Conference on Structural Safety & Reliability*, IASSAR.

Chan, J., Papaioannou, I., and Straub, D. (2023). "Bayesian improved cross entropy method with categorical mixture models.

Dehghani, N. L., Zamanian, S., and Shafieezadeh, A. (2021). "Adaptive network reliability analysis: Methodology and applications to power grid." *Reliability Engineering & System Safety*, 216, 107973.

Del Moral, P. (2013). *Mean Field Simulation for Monte Carlo Integration*. Chapman and Hall/CRC, <https://www.taylorfrancis.com/books/9781466504172> (may).

Hardy, G., Lucet, C., and Limnios, N. (2007). "K-terminal network reliability measures with binary decision diagrams." *IEEE Transactions on Reliability*, 56(3), 506–515.

Huber, M. (2017). "A Bernoulli mean estimate with known relative error distribution." *Random Structures and Algorithms*, 50(2), 173–182.

Hui, K.-P., Bean, N., Kraetzl, M., and Kroese, D. P. (2005). "The cross-entropy method for network reliability estimation." *Annals of Operations Research*, 134(1), 101.

Jensen, H. A. and Jerez, D. J. (2018). "A stochastic framework for reliability and sensitivity analysis of large scale water distribution networks." *Reliability Engineering & System Safety*, 176, 80–92.

Kroese, D. P., Taimre, T., and Botev, Z. I. (2013). *Handbook of Monte Carlo methods*, Vol. 706. John Wiley & Sons.

Li, J. and He, J. (2002). "A recursive decomposition algorithm for network seismic reliability evaluation." *Earthquake engineering & structural dynamics*, 31(8), 1525–1539.

Papaioannou, I., Betz, W., Zwirglmaier, K., and Straub, D. (2015). "Mcmc algorithms for subset simulation." *Probabilistic Engineering Mechanics*, 41, 89–103.

Papaioannou, I., Geyer, S., and Straub, D. (2019). "Improved cross entropy-based importance sampling with a flexible mixture model." *Reliability Engineering & System Safety*, 191, 106564.

Paredes, R., Dueñas-Osorio, L., Meel, K. S., and Vardi, M. Y. (2019). "Principled network reliability approximation: A counting-based approach." *Reliability Engineering & System Safety*, 191, 106472.

Paredes, R., Talebiyan, H., and Dueñas-Osorio, L. (2022). "Path-dependent reliability and resiliency of critical infrastructure via particle integration methods." *Proceedings of the 13th International Conference on Structural Safety & Reliability*, IASSAR.

Rubinstein, R. Y. (1997). "Optimization of computer simulation models with rare events." *European Journal of Operational Research*, 99(1), 89–112.

Zimmerman, R. D., Murillo-Sánchez, C. E., and Thomas, R. J. (2010). "MATPOWER: steady-state operations, planning, and analysis tools for power systems research and education." *IEEE Transactions on power systems*, 26(1), 12–19.

Zio, E. (2013). *Monte Carlo simulation: The method*. Springer.

Zio, E. and Pedroni, N. (2008). "Reliability analysis of discrete multi-state systems by means of subset simulation." *Proceedings of the 17th ESREL Conference*, 22–25.

Zuev, K. M., Wu, S., and Beck, J. L. (2015). "General network reliability problem and its efficient solution by subset simulation." *Probabilistic Engineering Mechanics*, 40, 25–35.

Zwirglmaier, K., Chan, J., Papaioannou, I., and Straub, D. (2023). "Hybrid bayesian networks for reliability assessment of infrastructure systems." *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*.