IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 24, 2019, accepted January 6, 2020, date of publication January 17, 2020, date of current version January 31, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2967599

Low-Overhead Evaluation of Multiuser Detection Performance for Physical-Layer Multiple Access Systems

FENGXIA HAN[1], SHENGJIE ZHAO[1], (Senior Member, IEEE), HAO JIANG[2], (Member, IEEE), HONG CHEN[3], (Senior Member, IEEE), AND CHENXI ZHANG[4]

[1]School of Software Engineering, Key Laboratory of Embedded System and Service Computing, School of Software Engineering, Tongji University, Shanghai 201804, China
[2]College of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing 210044, China
[3]School of Automotive Studies, Tongji University, Shanghai 201804, China
[4]School of Software Engineering, Tongji University, Shanghai 201804, China

Corresponding author: Shengjie Zhao (shengjiezhao@tongji.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61936014, and in part by the National Key Research and Development Project under Grant 2019YFB2102300 and Grant 2019YFB2102301.

ABSTRACT In order to facilitate massive connectivity in fifth-generation (5G) systems and make full use of advanced coding schemes, designing effective multiuser detection algorithms is necessary for mitigating the interference in multiple access systems. To evaluate the performance of different detection algorithms, the conventional method is implementing the Monte Carlo (MC) simulation to estimate the detection error rate. In this paper, we present a novel simulation scheme based on adaptive importance-sampling (AIS) theory, which accelerates the simulation speed for estimating the extremely low detection error rate in multiple access systems. Specifically, by restricting the generation of random codewords to the joint Gaussian distribution biased with scaling parameters, two algorithms are proposed to determine the optimal biased parameters such that the estimated variance or the cross-entropy resulted in the AIS simulation is minimized respectively. Our proposed simulation scheme is compared with the standard MC simulation in the performance evaluation of message passing algorithm (MPA) for uplink sparse code multiple access (SCMA) system. Numerical results show that our proposed scheme provides a feasible estimation of extremely low detection error rate and achieves significant performance gain with reduced simulation overhead.

INDEX TERMS Multiple access, multiuser detection, low-overhead simulation, adaptive importance-sampling.

I. INTRODUCTION

The Internet of Things (IoT) has emerged as an intelligent network for the fifth-generation (5G) wireless communications, which supports connections among a large number of users and/or smart devices [1]–[3]. To meet the increasing demand for low-latency and high-reliability transmission in multiple access systems, enhanced physical-layer coding and modulation schemes have been highly expected to accommodate massive connectivity and satisfy diverse service requirements [4], [5].

Recently, several non-orthogonal multiple access (NOMA) schemes have attracted a lot of attention, which can provide a good tradeoff between employing the limited spectral resources and mitigating the co-channel interference [6]–[9]. Among them, the sparse code multiple access (SCMA) induced in [10] has demonstrated its advantages in terms of multidimensional constellation gain and overloaded reception as a novel code-domain non-orthogonal multiple access scheme. Moreover, to achieve effective multiuser detection, message passing algorithm (MPA) has been proposed as a dominant detection scheme with near-optimal performance [11] in SCMA systems. However, the high complexity of existing multiuser detection algorithms poses an obstacle to the hardware implementation especially with massive users accessing. In this regard, seeking the low-complexity detection algorithms with superior performance is necessary for deploying multiple access systems [12], [13].

The associate editor coordinating the review of this manuscript and approving it for publication was Kezhi Wang.

The performance evaluation of different multiuser detection algorithms is usually carried out by observing the resulted bit error rate (BER) or symbol error rate (SER) [14]. Due to the inherent complexity in multiple access systems, the analytical solutions of the detection error rate is mathematically intractable. To facilitate efficient evaluation among different multiuser detection algorithms, Monte Carlo (MC) simulations have been widely utilized. However, the conventional MC scheme often requires generating a large number of random simulation samples to reach specific accuracy, and will lead to prohibitive computational complexity when the estimated detection error rate is extremely low [15]. Toward this end, a more efficient simulation scheme (with low complexity and high accuracy) to evaluate the performance of different multiuser detection algorithms is expected.

It is well known that the complexity of MC scheme is mainly related to the estimated variance caused by random events [16]. Such variance can be large or even infinite when the target event is rare, e.g., the extremely rare detection error., for which the required number of MC trials may be prohibitively high to maintain specific accuracy for estimating the error rate [17], [18]. To deal with such issue, importance-sampling (IS) scheme has been proposed to reduce the estimated variance to an acceptable level [19]. The basic idea of IS scheme is to design a biased distribution for event generation such that the occurrence frequency of rare event is enhanced. Although the zero-variance IS biased distribution has theoretically existed with simple representation, it involves a priori knowledge of the value to be estimated [20]. Therefore, a large body of related work [21]–[23] have been devoted to searching for the implementable and suboptimal biased distribution, that facilitates both fast simulation and acceptable accuracy compared with the conventional MC scheme.

One of the commonly used approaches for designing the biased distributions is to restrict them to a selected family of density functions indexed by one or more parameters. Such selection can be carried out by identifying a density distribution that closely resembles the zero-variance IS biased distribution [24]. In addition, it could also refer to the direct transformation, e.g., scaling, translation, or exponential twisting, of the original density distribution used for generating the random event [25], [26]. Once the selection is made, the rest procedure is mainly about determining the optimal biased parameters with minimized estimated variance or the other metrics, such as the cross-entropy [27]. Work [28] proposed to restrict the biased distribution to Bernoulli distribution for generating random linear block codewords over binary symmetric channels, based on which the optimal IS biased distribution was analytically proved. To estimate the low error rate of low-density parity-check (LDPC) codes over fading channels, [29] regarded the IS biased distribution as a mean-shifted version of the original noise density, where the involved parameters were approximately derived in closed-form solutions.

However, the aforementioned optimization procedure cannot always proceed directly and sometimes may lead to numerically tedious exercise. According to Bayesian theory, the unknown posterior can be approximated by employing the generated samples from the prior. In this way, the adaptive importance-sampling (AIS) scheme has been emerged recently [30]–[32], which suggests to iteratively improve the selected biased distribution as the simulations proceed simultaneously. More explicitly, by generating random events according to an initial biased distribution, the corresponding estimation results can be collected to produce a better biased distribution. Thus, the learning process takes place from events obtained in previous iterations and the selection of biased distributions keeps adapting, which provides methods to effectively optimize the biased parameters with lower complexity. Works [33], [34] have investigated the superiority of AIS for dealing with the infeasible optimization of biased parameters. By delicately designing the biased distributions, the simulation overhead for evaluating the performance of different coding or decoding algorithms can be significantly saved.

Unfortunately, most of the existing work have assumed to apply the IS-based simulations into the point-to-point communication system, which is not extended to more general cases of 5G wireless networks, i.e., the physical-layer transmission is realized among a large number of users or devices [35]. Therefore, how to exploit the potential benefits of the IS scheme in more complex systems is challenging and necessary.

In this paper, we will discuss the AIS-aided fast simulation scheme for performance evaluation of multiuser detection algorithms in the physical-layer multiple access system. Specifically, we focus on the uplink SCMA system detected through the low-complexity MPA algorithm. Since the detection error rate is mainly caused by the background noise and since the co-channel interference can be effectively eliminated through multiuser detection, we regard the random noise as the random event and consider to restrict its biased distribution to the Gaussian distribution with alterable noise variance. Then, two algorithms are proposed to determine the suboptimal biased parameters, which are aimed at minimizing the estimated variance and the cross-entropy, respectively. Simulation results show that our proposed scheme will significantly reduce the simulation overhead compared with the conventional MC scheme, especially for estimating the extremely low detection error rate.

II. SYSTEM MODEL

We consider an uplink sparse code multiple access (SCMA) system, where J single-antenna users transmit signal simultaneously to the base station in K orthogonal resources ($J > K$). For each user j, every L incoming bits are mapped into a K-dimensional complex codeword $\mathbf{x}_j = (x_{1,j}, \cdots, x_{K,j})^T$ selected from a K-dimensional complex codebook $\mathcal{X}_j$ of size $Q = 2^L$. In order to accommodate the demand of massive connectivity in 5G system, the overloading factor is defined as $\lambda = J/K$ and $\lambda > 1$ in general.

The K-dimensional received signal $\mathbf{y} = (y_1, \cdots, y_K)^T$ at the base station can be expressed as

$$\mathbf{y} = \sum_{j=1}^{J} \text{diag}(\mathbf{h}_j)\mathbf{x}_j + \mathbf{n}, \tag{1}$$

where $\mathbf{h}_j = (h_{1,j}, \cdots, h_{K,j})^T$ denotes the channel vector between the base station and user j, whose elements obey the complex Gaussian distributions $\mathcal{CN}(0, 1)$. $\mathbf{n}$ denotes the Gaussian noise vector, and follows the complex Gaussian distribution $\mathcal{CN}(0, \sigma^2\mathbf{I})$.

Given the observed $\mathbf{y}$, the base station needs to choose a specific multiuser detection algorithm (e.g., message passing algorithm (MPA)) to recover the transmitted codewords $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_J)$. The performance of different detection algorithms can be measured by calculating the probability of detection error, i.e., the probability that the detected codewords is different from the transmitted codewords, which is mathematically given by

$$P(e) = \mathbb{E}\left[\frac{c(\mathbf{y})}{J}\right] = \mathbb{E}[g(\mathbf{y})], \tag{2}$$

where $c(\mathbf{y})$ denotes the sum of codewords detected erroneously given each observation $\mathbf{y}$; $g(\mathbf{y})$ denotes the average number of codewords detected erroneously for each individual user; and $\mathbb{E}[\cdot]$ represents the expectation operation among all potential observations. For ease of notation, we refer $P(e)$ as the symbol error rate (SER) hereinafter.

Since the detection error induced by most detection algorithms is mainly related to the background noise, we can regard that $\mathbb{E}[g(\mathbf{y})]$ equals to $\mathbb{E}[g(\mathbf{n})]$ given the same transmitted codewords $\mathbf{X}$, where $g(\mathbf{n})$ denotes the number of codewords detected erroneously under noise $\mathbf{n}$. Then, we have the following transformation for $P(e)$ as,

$$P(e) = \mathbb{E}[g(\mathbf{y})] = \mathbb{E}[g(\mathbf{n})] = \int_{\mathbf{n}} g(\mathbf{n})p(\mathbf{n})\mathrm{d}\mathbf{n}, \tag{3}$$

where $p(\mathbf{n}) = \prod_{k=1}^{K} p(n_k)$ denotes the joint PMF of the occurrence of noise vector $\mathbf{n}$.

Due to the unknown background noise and the inherent complexity of detection algorithms in SCMA systems, $g(\mathbf{n})$ has no explicit form and thus the SER defined in (3) can not be solved analytically. Therefore, to evaluate and compare the performance of different detection algorithms, employing the simulation scheme to generate random samples and estimate the SERs has been widely utilized.

A. MONTE–CARLO (MC) SIMULATION

MC simulation is one of the commonly used approaches for estimating SERs in SCMA systems, which is executed upon averaging the detection error over the input space via naive event generation. More specifically, the MC estimator $\hat{P}_{MC}(e)$ for SER is given by

$$\hat{P}_{MC}(e) = \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} g(\mathbf{n}_i), \tag{4}$$

where $\mathbf{n}_i$ denotes the i-th random noise event and N_{MC} represents the number of simulation trails for obtaining the MC estimator.

As discussed in [16], [17], the MC estimator defined in (4) is an unbiased estimator for SER whose variance is given by

$$\text{Var}\left[\hat{P}_{MC}(e)\right] = \frac{1}{N_{MC}}\left(\mathbb{E}[g^2(\mathbf{n})] - P^2(e)\right), \tag{5}$$

based on which the relative error of the MC estimator can be further defined as

$$\kappa_{MC} \triangleq \frac{\sqrt{\text{Var}\left[\hat{P}_{MC}(e)\right]}}{P(e)} = \frac{1}{P(e)}\sqrt{\frac{\mathbb{E}[g^2(\mathbf{n})] - P^2(e)}{N_{MC}}}, \tag{6}$$

from which we observe that the relative error of the MC estimator is inversely proportional to N_{MC} and can be made arbitrarily small by increasing the number of simulation trials. Therefore, for the conventional MC scheme, the number of simulation trials N_{MC} may turn to be prohibitively large to maintain a fixed level of relative accuracy, especially when the estimated SER is sufficiently small.

Take the uplink SCMA systems considered in this paper as an example. If we adopt MPA as the multiuser detection algorithm and assume the system operates at high SNR regimes with low detection error rate, e.g., $P(e) \leq 1 \times 10^{-6}$, then the MC estimator requires more than $10^8/J$ independent trials to achieve a relative accuracy of $\kappa_{MC} = 0.1$.

B. IMPORTANCE-SAMPLING (IS) SIMULATION

IS simulation is in essence a variance-reduction scheme employed to increase the occurrence frequency of the rare event such that the low SER performance can be efficiently estimated. Before proceeding to introduce the IS scheme, we rewrite the mathematical expression of $P(e)$ in (3) as

$$P(e) = \int_{\mathbf{n}} g(\mathbf{n})p(\mathbf{n})\mathrm{d}\mathbf{n} = \int_{\mathbf{n}} g(\mathbf{n})\frac{p(\mathbf{n})}{p^*(\mathbf{n})}p^*(\mathbf{n})\mathrm{d}\mathbf{n}, \tag{7}$$

where $p^*(\mathbf{n})$ denotes the biased distribution and is exactly the main part to be designed for employing the IS scheme. Then, similar as the MC estimator defined in (4), the IS estimator $\hat{P}_{IS}(e)$ for SER is given by

$$\hat{P}_{IS}(e) = \frac{1}{N_{IS}} \sum_{i=1}^{N_{IS}} g(\mathbf{n}_i)\frac{p(\mathbf{n}_i)}{p^*(\mathbf{n}_i)}$$

$$= \frac{1}{N_{IS}} \sum_{i=1}^{N_{IS}} g(\mathbf{n}_i)\omega(\mathbf{n}_i), \tag{8}$$

where $\omega(\cdot)$ is called the weight function given by

$$\omega(\mathbf{n}) = \frac{p(\mathbf{n})}{p^*(\mathbf{n})}. \tag{9}$$

Note that for obtaining the IS estimator in (8), we restrict the generation of random noise to a newly designed biased distribution, i.e., $\mathbf{n}_i \sim p^*(\mathbf{n})$, instead of following the original noise density distribution as in the MC scheme, i.e., $\mathbf{n}_i \sim p(\mathbf{n})$. And to ensure the equivalence between the two estimators, we introduce the weight function $\omega(\cdot)$ for coordination.

Then, according to the relations (5) and (6), the variance and the relative error for the IS estimator can be respectively expressed as

$$\mathrm{Var}\Big[\hat{P}_{IS}(e)\Big] = \frac{1}{N_{IS}}\Big(\mathbb{E}_*[g^2(\mathbf{n})\omega^2(\mathbf{n})] - P^2(e)\Big), \tag{10}$$

and

$$\kappa_{IS} \triangleq \frac{\sqrt{\mathrm{Var}\Big[\hat{P}_{IS}(e)\Big]}}{P(e)} = \frac{1}{P(e)}\sqrt{\frac{\mathbb{E}_*[g^2(\mathbf{n})\omega^2(\mathbf{n})] - P^2(e)}{N_{IS}}}, \tag{11}$$

where $\mathbb{E}_*[\cdot]$ represents the expectation with respect to the biased distribution $p^*(\cdot)$.

The main objective in designing the IS scheme is to find a biased distribution that makes $\mathrm{Var}[\hat{P}_{IS}(e)]$ as small as possible. Obviously, the optimal biased distribution is theoretically existed that contributes to zero variance as

$$p^*_{\mathrm{opt}}(\mathbf{n}) = \frac{g(\mathbf{n})p(\mathbf{n})}{P(e)}. \tag{12}$$

However, such zero-variance biased distribution is not feasible in practical simulations, since it requires the prior knowledge of $P(e)$, which is exactly the value to be estimated.

Therefore, we need to search for the biased distribution that can be implemented easily, and at the same time, requires as fewer simulation trials as possible. In the following sections, we will provide two algorithms to determine the suboptimal biased distributions, which aim at minimizing the estimated variance and the cross-entropy for the IS estimator, respectively.

III. STOCHASTIC APPROXIMATION BASED ALGORITHM

In this section, we will propose an adaptive importance-sampling algorithm based on stochastic approximation (SA) theory, which can solve a wide range of suboptimal biased distributions for the IS scheme to estimate the SER in uplink SCMA systems.

A. PROBLEM FORMULATION

As illustrated in Section II.B, the superiority of the IS estimator reflects in the lower variance and thus we hope to find the biased distribution with minimized estimated variance. Motivated by [28], [29], such problem can be simplified if the biased distribution is restricted to a selected family of density functions and is parameterized by several scalars. To this end, we assume that the biased distribution $p^*(\mathbf{n})$ is parameterized by a single scalar θ and denoted by $p^*(\mathbf{n};\theta)$. Then, the optimization problem can be formulated as

$$\theta^{\mathrm{opt}} = \arg\min_{\theta} \mathrm{Var}\Big[\hat{P}_{IS}(e)\Big]. \tag{13}$$

From (8), the biased distribution is used to change the original distribution of random events (i.e., random noise). Recall the assumption of complex Gaussian noise for the considered uplink SCMA systems as indicated in Section II. If the biased distribution is restricted to the Gaussian distribution with alterable noise variance, the new distribution for noise generation will be easily implemented through simulation trials and the parameters to be designed are only related to the noise variance. In this way, we declare that $p^*(\mathbf{n};\theta)$ has the following form

$$p^*(\mathbf{n};\theta) = p(\mathbf{n};\sigma_*) = \prod_{k=1}^{K} \frac{1}{\sqrt{2\pi}\sigma_*} \exp\Big(-\frac{|n_k|^2}{2\sigma_*^2}\Big), \tag{14}$$

where σ_* denotes the altered variance and corresponds to the variable to be solved for determining the biased distribution.

Then, by denoting $p(\mathbf{n};\sigma)$ and $p(\mathbf{n};\sigma_*)$ as the original and biased distributions for random noise, respectively. The IS estimator defined in (8) can be transformed as

$$\begin{aligned}\hat{P}_{IS}(e) &= \frac{1}{N_{IS}} \sum_{i=1}^{N_{IS}} g(\mathbf{n}_i)\frac{p(\mathbf{n}_i;\sigma)}{p(\mathbf{n}_i;\sigma_*)} \\ &= \frac{1}{N_{IS}} \sum_{i=1}^{N_{IS}} g(\mathbf{n}_i)\frac{\frac{1}{(\sqrt{2\pi}\sigma)^K}\exp(-\frac{\|\mathbf{n}_i\|^2}{2\sigma^2})}{\frac{1}{(\sqrt{2\pi}\sigma_*)^K}\exp(-\frac{\|\mathbf{n}_i\|^2}{2\sigma_*^2})} \\ &= \frac{1}{N_{IS}} \sum_{i=1}^{N_{IS}} g(\mathbf{n}_i)\Big(\frac{\sigma_*}{\sigma}\Big)^K \exp\Big(\frac{\|\mathbf{n}_i\|^2}{2\sigma_*^2} - \frac{\|\mathbf{n}_i\|^2}{2\sigma^2}\Big) \\ &= \frac{1}{N_{IS}} \sum_{i=1}^{N_{IS}} g(\mathbf{n}_i)\omega(\mathbf{n}_i;\sigma,\sigma_*),\end{aligned} \tag{15}$$

where the weight function is given by

$$\omega(\mathbf{n}_i;\sigma,\sigma_*) = \Big(\frac{\sigma_*}{\sigma}\Big)^K \exp\Big(\frac{\|\mathbf{n}_i\|^2}{2\sigma_*^2} - \frac{\|\mathbf{n}_i\|^2}{2\sigma^2}\Big). \tag{16}$$

Based on the above discussions, the optimization problem in (13) can be further reformulated as

$$\sigma_*^{\mathrm{opt}} = \arg\min_{\sigma_* \in \mathcal{R}^+} \mathrm{Var}\Big[\hat{P}_{IS}(e)\Big]$$

$$= \arg\min_{\sigma_* \in \mathcal{R}^+} \; I(\sigma_*), \tag{17}$$

where

$$I(\sigma_*) = \mathbb{E}_*[g^2(\mathbf{n})\omega^2(\mathbf{n}; \sigma, \sigma_*)]. \tag{18}$$

In the following, we will discuss how to deal with the problem (17) and obtain the suboptimal biased parameters σ_*.

B. STOCHASTIC APPROXIMATION (SA) BASED ALGORITHM

Before proceeding to deal with the problem (17), we assume the function $I(\sigma_*)$ defined in (18) can be uniquely minimized. Then, the optimal solution to (17) is determined by $I'(\sigma_*^{\text{opt}}) = 0$. Inspired by [20], [23], a stochastic approximation procedure can be used to gradually approximate such optimal solution through limited recursions, i.e.,

$$\sigma_*^{(t+1)} = \sigma_*^{(t)} - \delta \frac{I'(\sigma_*^{(t)})}{I''(\sigma_*^{(t)})}, \quad t = 1, 2, \cdots, T, \tag{19}$$

where t denotes the t-th recursion step; δ denotes the rate factor used to control the convergence speed; and the first and second derivatives of $I(\sigma_*)$ are respectively given by

$$I'(\sigma_*) = \mathbb{E}_*[g^2(\mathbf{n})\omega(\mathbf{n}; \sigma, \sigma_*)\omega'(\mathbf{n}; \sigma, \sigma_*)], \tag{20}$$

and

$$I''(\sigma_*) = \mathbb{E}_*[g^2(\mathbf{n})\omega(\mathbf{n}; \sigma, \sigma_*)\omega''(\mathbf{n}; \sigma, \sigma_*)]. \tag{21}$$

According to the derived weight function for employing the IS scheme in (16), we can obtain $\omega'(\mathbf{n}; \sigma, \sigma_*)$ and $\omega''(\mathbf{n}; \sigma, \sigma_*)$ as

$$\omega'(\mathbf{n}; \sigma, \sigma_*) = \left(\frac{K}{\sigma^K}\sigma_*^{K-1} - \frac{\|\mathbf{n}_i\|_2^2}{\sigma^K}\sigma_*^{K-3}\right) \times \exp\left(\frac{\|\mathbf{n}_i\|_2^2}{2\sigma_*^2} - \frac{\|\mathbf{n}_i\|_2^2}{2\sigma^2}\right), \tag{22}$$

and

$$\omega''(\mathbf{n}; \sigma, \sigma_*) = \left(\frac{K^2 - K}{\sigma^K}\sigma_*^{K-2} - \frac{(2K-3)\cdot\|\mathbf{n}_i\|_2^2}{\sigma^K}\sigma_*^{K-4} + \frac{\|\mathbf{n}_i\|_2^4}{\sigma^K}\sigma_*^{K-6}\right) \times \exp\left(\frac{\|\mathbf{n}_i\|_2^2}{2\sigma_*^2} - \frac{\|\mathbf{n}_i\|_2^2}{2\sigma^2}\right). \tag{23}$$

Then, to facilitate the recursions in (19) and gradually converge to the optimal solution, the distribution of $g(\mathbf{n})$ involved in $I'(\sigma_*)$ and $I''(\sigma_*)$ need to be further figured out, which is intractable as illustrated in Section II without prior knowledge. Thus, we consider estimating the value of $I'(\sigma_*)$ and $I''(\sigma_*)$ through naive trials. More specifically, by generating some additional random samples and collecting the corresponding results, we can obtain the sample estimator to approximate $I(\sigma_*)$ as

$$\hat{I}(\sigma_*) = \frac{1}{N}\sum_{i=1}^{N} g^2(\mathbf{n}_i)\omega^2(\mathbf{n}_i; \sigma, \sigma_*), \tag{24}$$

where $\mathbf{n}_i \sim p(\mathbf{n}; \sigma_*)$ and N denotes the pre-set number of simulation trials for obtaining the sample estimator $\hat{I}(\sigma_*)$.

Accordingly, by differentiating $\hat{I}(\sigma_*)$, the sample estimator of $I'(\sigma_*)$ and $I''(\sigma_*)$ can be respectively given by

$$\hat{I}'(\sigma_*) = \frac{1}{N}\sum_{i=1}^{N} g^2(\mathbf{n}_i)\omega(\mathbf{n}_i; \sigma, \sigma_*)\omega'(\mathbf{n}_i; \sigma, \sigma_*), \tag{25}$$

and

$$\hat{I}''(\sigma_*) = \frac{1}{N}\sum_{i=1}^{N} g^2(\mathbf{n}_i)\omega(\mathbf{n}_i; \sigma, \sigma_*)\omega''(\mathbf{n}_i; \sigma, \sigma_*), \tag{26}$$

where $\mathbf{n}_i \sim p(\mathbf{n}; \sigma_*)$ and N denotes the number of trials for estimating $I'(\sigma_*)$ and $I''(\sigma_*)$ at each recursion step. Based on (25) and (26), the recursions in (19) can be then replaced by the following procedure as

$$\sigma_*^{(t+1)} = \sigma_*^{(t)} - \delta \frac{\hat{I}'(\sigma_*^{(t)})}{\hat{I}''(\sigma_*^{(t)})}, \quad t = 1, 2, \cdots, T. \tag{27}$$

Note that the initial value of $\sigma_*^{(0)}$ should be carefully selected. To make such issue more explicit, assume the considered SCMA system operates in high SNR regime and makes a large amount of random events with $g(\mathbf{n}) = 0$. Then, without a reasonable biased distribution, there may exist intermediate values of $\frac{\hat{I}'(\sigma_*^{(t)})}{\hat{I}''(\sigma_*^{(t)})} = \frac{0}{0}$. To guarantee the sustained feasibility of (27), we propose to initialize $\sigma_*^{(0)} = \tilde{\sigma}$, based on which the following is satisfied, i.e.,

$$\hat{P}(e) = \frac{1}{\tilde{N}}\sum_{i=1}^{\tilde{N}} g(\mathbf{n}_i) \geq \gamma, \quad \mathbf{n}_i \sim p(\mathbf{n}; \tilde{\sigma}), \tag{28}$$

where $\hat{P}(e)$ represents the estimated SER through $\tilde{N}$ naive simulations; and γ denotes the desired level for SER and is usually set to be $10^{-2} \leq \gamma \leq 10^{-1}$.

In summary, for the proposed SA-based algorithm, we firstly initialize the biased parameter $\sigma_*^{(0)} = \tilde{\sigma}$ through $\tilde{N}$ MC trials following (28). Then, by properly selecting the number of recursion steps T, the number of simulation trials N for approximating $I'(\sigma_*)$ and $I''(\sigma_*)$ in each recursion step, and the rate factor δ, we can adaptively update the biased parameter σ_* as the recursion procedure (27) proceeds. The original problem (17) can be finally solved once such recursion terminates.

IV. CROSS-ENTROPY BASED ALGORITHM

In this section, we will provide another adaptive importance-sampling approach to estimate the low detection error rate for the uplink SCMA system based on cross-entropy (CE) theory, which has been shown to perform well when the biased distribution belongs to an exponential family [27].

A. PROBLEM FORMULATION

Different from the SA-based approach invoked in the previous section that aims at finding the biased distribution with minimized estimated variance, the essential idea of the

CE-based approach is to select a biased distribution with minimized CE distance from the zero-variance biased distribution given in (12). In specific, the CE distance of distribution $p_1(\mathbf{n})$ from distribution $p_2(\mathbf{n})$ is defined as

$$\begin{aligned}\mathcal{D}\Big[p_2(\mathbf{n}), p_1(\mathbf{n})\Big] &= \mathbb{E}_{p_2}\Big[\ln\frac{p_2(\mathbf{n})}{p_1(\mathbf{n})}\Big]\\ &= \int \ln\frac{p_2(\mathbf{n})}{p_1(\mathbf{n})}p_2(\mathbf{n})\mathrm{d}\mathbf{n}\\ &= \int \ln p_2(\mathbf{n})p_2(\mathbf{n})\mathrm{d}\mathbf{n}\\ &\quad - \int \ln p_1(\mathbf{n})p_2(\mathbf{n})\mathrm{d}\mathbf{n}, \end{aligned}\tag{29}$$

where $\mathbb{E}_{p_2}[\cdot]$ denotes the expectation with respect to the distribution $p_2(\mathbf{n})$ and $\mathcal{D}[p_2(\mathbf{n}), p_1(\mathbf{n})] \geq 0$.

Based on the above definition, the CE distance of the biased distribution $p^*(\mathbf{n})$ from the optimal biased distribution $p^*_{\text{opt}}(\mathbf{n})$ can be expressed as

$$\begin{aligned}\mathcal{D}\Big[p^*_{\text{opt}}(\mathbf{n}), p^*(\mathbf{n})\Big] &= \int \ln p^*_{\text{opt}}(\mathbf{n})p^*_{\text{opt}}(\mathbf{n})\mathrm{d}\mathbf{n}\\ &\quad - \int \ln p^*(\mathbf{n})p^*_{\text{opt}}(\mathbf{n})\mathrm{d}\mathbf{n}, \end{aligned}\tag{30}$$

where $\mathcal{D}[p^*_{\text{opt}}(\mathbf{n}), p^*(\mathbf{n})] = 0$ leads to the optimal solution, i.e., $p^*(\mathbf{n}) = p^*_{\text{opt}}(\mathbf{n})$. Since $p^*(\mathbf{n})$ has no analytical form and $\int \ln p^*_{\text{opt}}(\mathbf{n})p^*_{\text{opt}}(\mathbf{n})\mathrm{d}\mathbf{n}$ is constant, the main objective becomes to search for the biased distribution satisfying:

$$p^*(\mathbf{n}) = \arg\max_{p^*} \int \ln p^*(\mathbf{n})p^*_{\text{opt}}(\mathbf{n})\mathrm{d}\mathbf{n}. \tag{31}$$

As mentioned in Section III.A, for ease of exposition and to maintain the simplicity of event generation, we extend the idea of restricting the biased distributions to Gaussian distributions parameterized by noise variance, i.e., $p^*(\mathbf{n}) = p(\mathbf{n}; \sigma_*)$. Then, referring to (12), we have

$$p^*_{\text{opt}} = \frac{g(\mathbf{n})p(\mathbf{n};\sigma)}{P(e)}, \tag{32}$$

where σ denotes the original noise variance in the considered uplink SCMA system.

From (31) and (32), the problem for determining the optimal biased distribution can be formulated as:

$$\begin{aligned}\sigma_*^{\text{opt}} &= \arg\max_{\sigma_*\in\mathcal{R}^+} \int g(\mathbf{n})p(\mathbf{n};\sigma)\ln p(\mathbf{n};\sigma_*)\mathrm{d}\mathbf{n}\\ &= \arg\max_{\sigma_*\in\mathcal{R}^+} \mathbb{E}_{\sigma}\Big[g(\mathbf{n})\ln p(\mathbf{n};\sigma_*)\Big], \end{aligned}\tag{33}$$

where $\mathbb{E}_{\sigma}[\cdot]$ denotes the expectation with respect to noise variance σ. In the following, we will discuss the method of solving the problem (33).

B. CROSS-ENTROPY (CE) BASED ALGORITHM

Since the objective in (33) also contains the unknown knowledge of $g(\mathbf{n})$, we consider to approximate it by taking the averaged results over additional random samples, i.e.,

$$\mathbb{E}_{\sigma}\Big[g(\mathbf{n})\ln p(\mathbf{n};\sigma_*)\Big] \simeq \frac{1}{N}\sum_{i=1}^{N} g(\mathbf{n}_i)\ln p(\mathbf{n}_i;\sigma_*), \tag{34}$$

where $\mathbf{n}_i \sim p(\mathbf{n};\sigma)$ and N denotes the pre-set number of trials for obtaining the sample estimator.

Upon obtaining the above sample estimator, the original optimization problem (33) can be transformed as

$$\hat{\sigma}_*^{\text{opt}} = \arg\max_{\sigma_*\in\mathcal{R}^+} \frac{1}{N}\sum_{i=1}^{N} g(\mathbf{n}_i)\ln p(\mathbf{n}_i;\sigma_*), \tag{35}$$

where $\hat{\sigma}_*^{\text{opt}}$ represents the approximation of σ_*^{opt}.

Recall the expression for $p(\mathbf{n};\sigma_*)$ in (14), we observe that $\ln p(\mathbf{n}_i;\sigma_*)$ is a concave function with respect to σ_*. Therefore, the optimal solution to the problem (35) can be achieved by solving the following equation:

$$\frac{1}{N}\sum_{i=1}^{N} g(\mathbf{n}_i)\ln' p(\mathbf{n}_i;\hat{\sigma}_*^{\text{opt}}) = 0, \tag{36}$$

where $\ln' p(\mathbf{n}_i;\sigma_*)$ denotes the first derivative of function $\ln p(\mathbf{n}_i;\sigma_*)$.

By further substituting (14) into (36), the optimal value of $\hat{\sigma}_*^{\text{opt}}$ can be mathematically derived in the closed form given by

$$\hat{\sigma}_*^{\text{opt}} = \sqrt{\frac{\sum_{i=1}^{N} g(\mathbf{n}_i)\|\mathbf{n}_i\|_2^2}{K\sum_{i=1}^{N} g(\mathbf{n}_i)}}, \tag{37}$$

where K represents the number of available orthogonal resources as defined in Section II.

However, as discussed in Section III.B, when the considered system operates in high SNR regimes, the detection output of $\sum_{i=1}^{N} g(\mathbf{n}_i)$ may always equal to zero even with large N. Therefore, (37) is not always feasible especially when the target event is rare. To avoid the infeasibility caused by the denominator in (37), we consider to introduce a tilting parameter $\tilde{\sigma}$, which corresponds to a biased distribution $p(\mathbf{n};\tilde{\sigma})$ that satisfies

$$\hat{P}(e) = \frac{1}{\tilde{N}}\sum_{i=1}^{\tilde{N}} g(\mathbf{n}_i) \geq \gamma, \quad \mathbf{n}_i \sim p(\mathbf{n};\tilde{\sigma}), \tag{38}$$

where $\tilde{N}$ denotes the pre-set number of trials for obtaining the tilting biased parameter $\tilde{\sigma}$ to guarantee (38); and γ is set to be a not too small number, e.g., $10^{-2} \leq \gamma \leq 10^{-1}$.

Given the tilting parameter and extend the idea from (8), it is readily seen that (34) is equivalent to the following:

$$\mathbb{E}_{\tilde{\sigma}}\Big[g(\mathbf{n})\omega(\mathbf{n};\sigma,\tilde{\sigma})\ln p(\mathbf{n};\sigma_*)\Big], \tag{39}$$

TABLE 1. Simulation parameters for different algorithms.

SNR / dB	0	1	2	3	4	5	6	7
T_1 (SA)	10	10	10	20	20	20	50	50
M_1 (SA)	100	100	100	500	500	500	1000	1000
T_2 (CE)	10	10	10	20	20	20	50	50
M_2 (CE)	100	100	100	500	500	500	1000	1000

where $\omega(\mathbf{n}; \sigma, \tilde{\sigma})$ is given by

$$\omega(\mathbf{n}; \sigma, \tilde{\sigma}) = \frac{p(\mathbf{n}; \sigma)}{p(\mathbf{n}; \tilde{\sigma})}. \tag{40}$$

Then, combining (35) and (39), we have

$$\hat{\sigma}_*^{\text{opt}} = \arg\max_{\sigma_* \in \mathcal{R}^+} \frac{1}{N} \sum_{i=1}^{N} g(\mathbf{n}_i)\omega(\mathbf{n}_i; \sigma, \tilde{\sigma}) \ln p(\mathbf{n}_i; \sigma_*), \tag{41}$$

where $\mathbf{n}_i \sim p(\mathbf{n}; \tilde{\sigma})$.

Similar as the solutions in (36), the sample-based optimal solution to the problem (41) is given by

$$\hat{\sigma}_*^{\text{opt}} = \sqrt{\frac{\sum_{i=1}^{N} g(\mathbf{n}_i)\omega(\mathbf{n}_i; \sigma, \tilde{\sigma})\|\mathbf{n}_i\|_2^2}{K\sum_{i=1}^{N} g(\mathbf{n}_i)\omega(\mathbf{n}_i; \sigma, \tilde{\sigma})}}. \tag{42}$$

For more accurate approximation, we propose to repeatedly implement the procedure (42) until T sample solutions are collected. Finally, the biased parameter σ_*^{opt} for the original problem (33) can be approximated as

$$\sigma_*^{\text{opt}} \simeq \frac{1}{T}\sum_{t=1}^{T} \sqrt{\frac{\sum_{i=1}^{N} g(\mathbf{n}_i)\omega(\mathbf{n}_i; \sigma, \tilde{\sigma})\|\mathbf{n}_i\|_2^2}{K\sum_{i=1}^{N} g(\mathbf{n}_i)\omega(\mathbf{n}_i; \sigma, \tilde{\sigma})}}. \tag{43}$$

In summary, for the proposed CE-based algorithm, we firstly choose a tilting biased parameter $\tilde{\sigma}$ by carrying out $\tilde{N}$ trials following (38). Then, based on the current SNR scenarios and the potential level of SER, we select the proper number of iterations T and the number of trials N for each sample solution. The suboptimal biased parameter based on CE theory can be finally determined from (43).

V. NUMERICAL RESULTS

In this section, simulations were carried out to investigate the performance of our proposed fast simulation schemes for the uplink SCMA systems. Assume the number of users $J = 16$ and the number of orthogonal resources $K = 12$. Further assume the 12×16 LDS signature matrix is designed for SCMA codebooks and binary phase-shift keying modulation (BPSK) is employed. Then, in an additive white Gaussian noise (AWGN) channel with $\sigma = 1$, the performance of MPA detection algorithm can be measured using SER. The other related parameters are consistent with [11].

To begin with, we compare the suboptimal biased parameters obtained from the stochastic approximation (SA) based algorithm and the cross-entropy (CE) based algorithm. As summarized in Section III.B and Section IV.B, both algorithms consist of two stages, i.e., the initialization of a tilting parameter $\tilde{\sigma}$ through $\tilde{N}$ standard MC trials and the

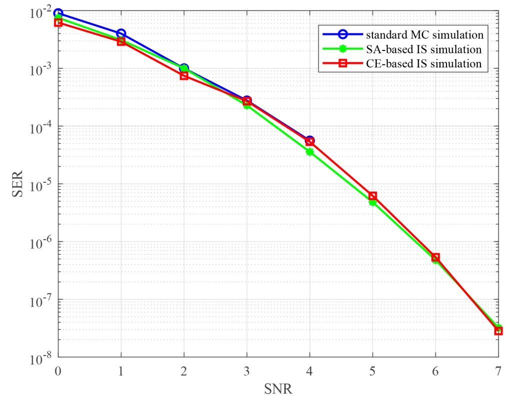

FIGURE 1. SER performance for the uplink SCMA system estimated by the MC and the proposed scheme.

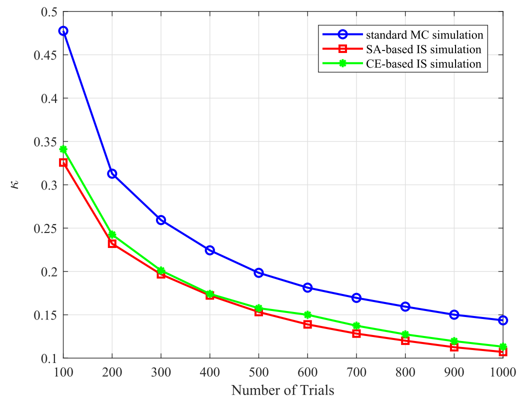

FIGURE 2. Relative error vs number of trials obtained by the MC and the proposed scheme.

approximation of the optimal biased parameter σ_*^{opt}. For the initialization stage, we assume $\tilde{N} = 100$ and $\gamma = 0.1$. For the approximation stage, let $T_1 = T, M_1 = N$ replace the related parameters in (27) for implementing the SA-based algorithm; and let $T_2 = T, M_2 = N$ replace the related parameters in (43) for implementing the CE-based algorithm. For fair comparison, their detailed values under different SNRs are set as in Table 1. Then, the corresponding resulted biased parameters σ_*^{opt} under different SNRs are given in Table 2. It can be observed that the biased parameters obtained by two algorithms are close to each other with the same computational complexity. In addition, it is also seen that the optimal biased parameter σ_*^{opt} will increase monotonously with SNR.

Given the biased parameters σ_*, the biased distribution $p(\mathbf{n}; \sigma_*)$ for event generation can be expressed according to (14). Then, we can apply the AIS-aided scheme to estimate the SER for the considered uplink SCMA system. Fig. 1 shows the efficiency of our proposed fast simulation compared to standard MC simulation in estimating the SER performance. Each curve is plotted at different SNRs with a stop condition on the relative error $\kappa = 0.1$. It is seen that the results obtained by our proposed scheme closely match with those of the MC scheme at low SNR regimes. Moreover, our proposed scheme is capable of estimating the extremely low detection error rate at high SNRs when the conventional MC scheme is not feasible due to the prohibitive simulation time.

TABLE 2. Suboptimal biased parameters with different algorithms.

SNR / dB	0	1	2	3	4	5	6	7
σ_*^{opt} (SA)	1.0986	1.1623	1.1913	1.2154	1.3285	1.5101	1.5757	1.7832
σ_*^{opt} (CE)	1.1140	1.1567	1.3119	1.2808	1.3531	1.4798	1.5655	1.7567

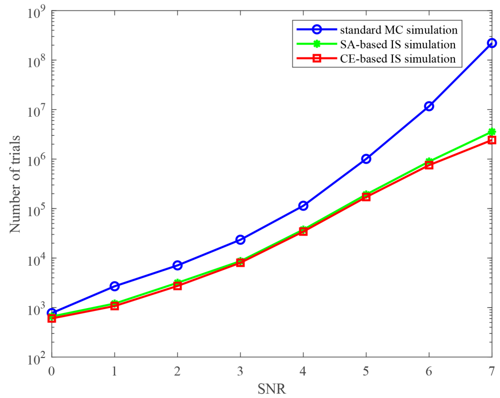

FIGURE 3. Simulation overhead caused by the MC and the proposed scheme.

To evaluate the accuracy of our proposed simulation scheme, we employ the relative error defined by κ_{MC} in (6) and κ_{IS} in (11) as the metric. By assuming the SCMA system is operated at SNR = 3dB, Fig.2 depicts the variations of the relative errors versus the number of simulation trials. From Fig.2, we observe that the relative error will decrease as the number of simulations increase for both MC and IS estimators. Moreover, our proposed IS estimator exhibits higher accuracy given the same number of trials.

The superiority of our proposed AIS-based scheme compared with the conventional MC scheme can be further reflected using the simulation overhead, which refers to the required number of trials for estimating the SER given $\kappa = 0.1$. The comparison result is reported in Fig. 3. It is evident that the number of trials that need to be performed using our proposed scheme can be significantly reduced compared with the MC scheme especially at high SNRs. Furthermore, such simulation gain brought by our proposed scheme will increase as the SNR increases.

VI. CONCLUSION

In this paper, we have proposed a fast simulation scheme based on adaptive importance-sampling (AIS) theory to estimate the detection error rate in the uplink SCMA systems. Different from the conventional Monte Carlo (MC) simulation, we have proposed to modify the original density distribution of random events, which aims at increasing the occurrence frequency of the rare event. Particularly, we have considered to restrict the biased distributions to Gaussian distributions with alterable noise variance. Toward this end, two algorithms have been offered to determine the suboptimal biased parameters such that the estimated variance or the cross-entropy is minimized respectively. Simulation results have shown that our proposed fast simulation scheme is feasible of estimating the extremely low detection error rate. Moreover, it is also shown that our proposed scheme outperforms the standard MC scheme in term of both accuracy and simulation overhead.

REFERENCES

[1] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A survey on Internet of Things: Architecture, enabling technologies, security and privacy, and applications," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1125–1142, Oct. 2017.
[2] M. Chiang and T. Zhang, "Fog and IoT: An overview of research opportunities," *IEEE Internet Things J.*, vol. 3, no. 6, pp. 854–864, Dec. 2016.
[3] H. Huang, S. Guo, G. Gui, Z. Yang, J. Zhang, H. Sari, and F. Adachi, "Deep learning for physical-layer 5G wireless techniques: Opportunities, challenges and solutions," *IEEE Wireless Commun.*, to be published.
[4] M. Moltafet, N. Mokari, M. R. Javan, H. Saeedi, and H. Pishro-Nik, "A new multiple access technique for 5G: Power domain sparse code multiple access (PSMA)," *IEEE Access*, vol. 6, pp. 747–759, 2018.
[5] S. Hu, H. Guo, C. Jin, Y. Huang, B. Yu, and S. Li, "Frequency-domain oversampling for cognitive CDMA systems: Enabling robust and massive multiple access for Internet of Things," *IEEE Access*, vol. 4, pp. 4583–4589, 2016.
[6] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE 77th Veh. Technol. Conf. (VTC Spring)*, Jun. 2013, pp. 1–5.
[7] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channel," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1616–1626, Apr. 2008.
[8] B. Lyu, Z. Yang, and G. Gui, "Non-orthogonal multiple access in wireless powered communication networks with SIC constraints," *IEICE Trans. Commun.*, vol. E101.B, no. 4, pp. 1094–1101, 2018.
[9] H. Sari, A. Maatouk, E. Caliskan, M. Assaad, M. Koca, and G. Gui, "On the foundation of NOMA and its application to 5G cellular networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2018, pp. 1–6.
[10] H. Nikopour and H. Baligh, "Sparse code multiple access," in *Proc. IEEE 24th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Sep. 2013, pp. 332–336.
[11] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh, "SCMA codebook design," in *Proc. IEEE 80th Veh. Technol. Conf.*, Sep. 2014, pp. 1–5.
[12] L. Yang, Y. Liu, and Y. Siu, "Low complexity message passing algorithm for SCMA system," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2466–2469, Dec. 2016.
[13] J. Dai, K. Niu, C. Dong, and J. Lin, "Improved message passing algorithms for sparse code multiple access," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 9986–9999, Nov. 2017.
[14] C. Zhang, Y. Luo, and Y. Chen, "A low-complexity SCMA detector based on discretization," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2333–2345, Apr. 2018.
[15] W. H. Tranter, T. S. Rappaport, K. L. Kosbar, and K. S. Shanmugan, *Principles of Communication Systems Simulation with Wireless Applications*, vol. 1. Upper Saddle River, NJ, USA: Prentice-Hall, 2004.
[16] R. Y. Rubinstein and D. P. Kroese, *Simulation and the Monte Carlo Method*. Hoboken, NJ, USA: Wiley, 2016, vol. 10.
[17] S. Juneja and P. Shahabuddin, "Rare-event simulation techniques: An introduction and recent advances," in *Handbooks in Operations Research and Management Science*, vol. 13. Amsterdam, The Netherlands: Elsevier B.V., 2006, pp. 291–350.
[18] S. Chakraborty and R. Chowdhury, "Hybrid framework for the estimation of rare failure event probability," *J. Eng. Mech.*, vol. 143, no. 5, May 2017, Art. no. 04017010.

[19] Z. I. Botev, P. L'Ecuyer, and B. Tuffin, "Markov chain importance sampling with applications to rare event probability estimation," *Stat Comput*, vol. 23, no. 2, pp. 271–285, Mar. 2013.
[20] R. Srinivasan, *Importance Sampling: Applications in Communications and Detection*. Berlin, Germany: Springer, 2013.
[21] S.-K. Ahn and K. Yang, "Importance sampling for performance estimation of LDPC codes over Rayleigh fading channels," in *Proc. 7th Int. Symp. Turbo Codes Iterative Inf. Process. (ISTC)*, Aug. 2012, pp. 86–90.
[22] V. Elvira and I. Santamaria, "Multiple importance sampling for efficient symbol error rate estimation," *IEEE Signal Process. Lett.*, vol. 26, no. 3, pp. 420–424, Mar. 2019.
[23] D. Remondo, R. Srinivasan, V. Nicola, W. Van Etten, and H. Tattje, "Adaptive importance sampling for performance evaluation and parameter optimization of communication systems," *IEEE Trans. Commun.*, vol. 48, no. 4, pp. 557–565, Apr. 2000.
[24] C. Feng and Y. M. Marzouk, "A layered multiple importance sampling scheme for focused optimal Bayesian experimental design," 2019, *arXiv:1903.11187*. [Online]. Available: https://arxiv.org/abs/1903.11187
[25] H. Roelofs, J. Thijs, and R. Srinivasan, "Performance and cellular capacity of M-ary PSK in co-channel interference," in *Proc. 11th IEEE Medit. Electrotech. Conf.*, Jun. 2003, pp. 204–208.
[26] S. Dey and S. Juneja, "Efficient estimation of density and probability of large deviations of sum of IID random variables," in *Proc. Winter Simulation Conf. (WSC)*, Dec. 2011, pp. 3800–3811.
[27] Z. Wang and J. Song, "Cross-entropy-based adaptive importance sampling using von Mises-Fisher mixture for high dimensional reliability analysis," *Struct. Saf.*, vol. 59, pp. 42–52, Mar. 2016.
[28] G. Romano and D. Ciuonzo, "Minimum-variance importance-sampling bernoulli estimator for fast simulation of linear block codes over binary symmetric channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 486–496, Jan. 2014.
[29] S.-K. Ahn, K. Yang, and D. Har, "Evaluation of the low error-rate performance of LDPC codes over rayleigh fading channels using importance sampling," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2166–2177, Jun. 2013.
[30] J. Míguez, I. P. Mariño, and M. A. Vázquez, "Analysis of a nonlinear importance sampling scheme for Bayesian parameter estimation in state-space models," *Signal Process.*, vol. 142, pp. 281–291, Jan. 2018.
[31] M. F. Bugallo, L. Martino, and J. Corander, "Adaptive importance sampling in signal processing," *Digit. Signal Process.*, vol. 47, pp. 36–49, Dec. 2015.
[32] J.-M. Cornuet, J.-M. Marin, A. Mira, and C. P. Robert, "Adaptive multiple importance sampling," *Scandin. J. Statist.*, vol. 39, no. 4, pp. 798–812, Dec. 2012.
[33] R. Holzlohner, A. Mahadevan, C. Menyuk, J. Morris, and J. Zweck, "Evaluation of the very low BER of FEC codes using dual adaptive importance sampling," *IEEE Commun. Lett.*, vol. 9, no. 2, pp. 163–165, Feb. 2005.
[34] Y. Liu, W. Chen, J. Yang, and T. Dong, "Improved dual adaptive importance sampling method for LDPC codes," in *Proc. 18th Asia–Pacific Conf. Commun. (APCC)*, Oct. 2012, pp. 975–979.
[35] L. Dai, B. Wang, Y. Yuan, S. Han, C.-L. I, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.

FENGXIA HAN received the B.S. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2015. She is currently pursuing the Ph.D. degree with Tongji University, Shanghai, China. From 2017 to 2019, she was a Visiting Student with the Department of Electrical Engineering, Columbia University, New York City, NY, USA. Her research interests include machine-type communications, multiple access, and massive MIMO.

SHENGJIE ZHAO (Senior Member, IEEE) received the B.S. degree in electrical engineering from the University of Science and Technology of China, Hefei, China, in 1988, the M.S. degree in electrical and computer engineering from the China Aerospace Institute, Beijing, China, in 1991, and the Ph.D. degree in electrical and computer engineering from Texas A&M University, College Station, TX, USA, in 2004. He is currently the Dean of the College of Software Engineering, a Professor with the College of Software Engineering and the College of Electronics and Information Engineering, Tongji University, Shanghai, China. In previous postings, he has conducted research with Lucent Technologies, Whippany, NJ, USA, and the China Aerospace Science and Industry Corporation, Beijing. His research interests include artificial intelligence, big data, wireless communications, image processing, and signal processing. He is a Fellow of the Thousand Talents Program of China.

HAO JIANG (Member, IEEE) received the B.S. and M.S. degrees in electrical and information engineering from the Nanjing University of Information Science and Technology, Nanjing, China, in 2012 and 2015, respectively, and the Ph.D. degree from the National Mobile Communications Research Laboratory, Southeast University, Nanjing, in 2019. From 2017 to 2018, he was a Visiting Student with the Department of Electrical Engineering, Columbia University, New York City, NY, USA. Since April 2019, he has been a Professor with the College of Artificial Intelligence, Nanjing University of Information Science and Technology. His current research interests include general areas of vehicle-to-vehicle communications, massive multiple-input and multiple-output channel modeling, signal processing, communications, machine learning, and AI-driven technologies.

HONG CHEN (Senior Member, IEEE) received the B.S. and M.S. degrees in process control from Zhejiang University, Zhejiang, China, in 1983 and 1986, respectively, and the Ph.D. degree in system dynamics and control engineering from the University of Stuttgart, Stuttgart, Germany, in 1997. She is currently a Professor with the School of Automotive Studies, Tongji University. Her current research interests include model predictive control, optimal and robust control, and nonlinear control and its applications in mechatronic systems focusing on automotive systems.

CHENXI ZHANG received the B.S. and Ph.D. degrees from the National University of Defense Technology. He is currently a Professor with the School of Software Engineering, Tongji University. His research interests include computer architecture and distributed computing.