RESEARCH　　　　　　　　　　　　　　　　　　　　　　**Open Access**

A simple importance sampling technique for orthogonal space-time block codes on Nakagami fading channels

Lennert Jacobs[*] and Marc Moeneclaey

Abstract

In this contribution, we present a simple importance sampling technique to considerably speed up Monte Carlo simulations for bit error rate estimation of orthogonal space-time block coded systems on spatially correlated Nakagami fading channels.

Keywords: Orthogonal space-time block codes; Bit error rate; Nakagami fading; Importance sampling

1　Introduction

In digital communications, the bit error rate (BER) serves as a fundamental performance measure. For complicated systems, analytical assessment of the BER becomes very difficult, if not impossible, and Monte Carlo (MC) simulations must be applied to estimate the BER. In the case of low BER, however, the computation time associated with MC simulations can be very long, since many bits must be sent to generate a sufficient number of bit errors. In order to speed up these simulations without losing precision, importance sampling (IS) [1] can be employed. The latter technique aims to reduce the variance of the BER estimator by using a biased distribution for the input variables that increases the bit error probability during simulation. The use of the biased distribution is corrected for by weighting the results with the ratios of the actual to the biased probability density functions (PDFs), yielding an unbiased BER estimate with lower variance.

Although the application of IS is well documented, the search for a convenient biased distribution remains a challenging task. In conventional importance sampling (CIS) [2], the variance of the additive Gaussian noise is increased, resulting in a scaled noise distribution and more errors. The improved importance sampling (IIS) technique proposed in [3], on the other hand, shifts the

mean of the actual PDF of the input random variables (RVs) directly towards the error region. Both techniques are combined in [4] for efficient simulation in Rayleigh fading. In order to speed up the simulation of orthogonal space-time block codes (OSTBCs) on Rayleigh-faded multiple-input multiple-output (MIMO) channels, it is shown in [5] that scaling the channel distribution is more efficient than scaling the noise distribution. However, an optimal biased channel distribution was not derived. An IS methodology based on the stochastic gradient descent (SGD) algorithm is presented in [6] for simulation of adaptive systems in frequency-flat Rayleigh fading channels.

In spite of the numerous papers on IS, an easy-to-use biased distribution for simulation over non-Rayleigh fading channels is not yet available from the literature to the best of our knowledge. In this contribution, we propose an IS technique for simulation over a frequency-flat MIMO fading channel, where we maintain the actual distribution for the channel noise and the data symbols, but use a biased channel distribution. Furthermore, we illustrate how the presented technique enables to derive a practical biased channel distribution for the case of OSTBCs transmitted over a spatially correlated Nakagami-m fading channel. Although perfect channel knowledge (PCK) is assumed for the derivation of the biased distribution, this distribution allows accurate BER estimation in a wide range of scenarios, such as pilot-based [7] or blind channel estimation, and in the presence of residual frequency

---

*Correspondence: lennert.jacobs@telin.ugent.be
Department of Telecommunications and Information Processing, Ghent University, St.-Pietersnieuwstraat 41, Ghent 9000, Belgium

offset, IQ imbalance, and phase noise. Moreover, the technique can easily be extended to the case of multicarrier communication on dispersive MIMO channels, using for instance OSTBCs or orthogonal space-frequency block codes.

2 Importance sampling

Let us assume that a vector $\mathbf{s}$ consisting of N_s information symbols is coded and transmitted over a frequency-flat MIMO channel, the coefficients of which are comprised in the vector $\mathbf{h}$. Collecting both $\mathbf{s}$ and $\mathbf{h}$ and the vector $\mathbf{w}$ consisting of the additive Gaussian noise samples into a single input vector $\mathbf{x} = [\mathbf{s}, \mathbf{w}, \mathbf{h}]$, the average BER can be written as

$$P_{\mathrm{b}} = \mathbb{E}\left[F(\mathbf{x})\right], \tag{1}$$

where $F(\mathbf{x})$ is the fraction of bit errors corresponding to given $\mathbf{x}$ and $\mathbb{E}[.]$ denotes expectation over the PDF $p(\mathbf{x})$ of the vector $\mathbf{x}$. The fraction $F(\mathbf{x})$ is given by

$$F(\mathbf{x}) = \frac{1}{N_{\mathrm{b}}} \sum_{n=1}^{N_{\mathrm{b}}} I_n(\mathbf{x}), \tag{2}$$

where N_{b} is the number of bits contained in the symbol vector $\mathbf{s}$ and $I_n(\mathbf{x})$ equals 1 when the decision about the nth bit is wrong and zero otherwise. When analytical averaging over $\mathbf{x}$ is too complex, a closed-form BER expression cannot be obtained. Using the MC method, however, the average BER can be easily estimated by independently generating a set of N realizations $\{\mathbf{x}_i\}$ of the input vector $\mathbf{x}$ according to the actual PDF $p(\mathbf{x})$ and simulating for each realization the system operations that yield the bit decisions at the receiver. An estimate of the BER is obtained as the ratio of the number of counted bit errors to the total number of bits transmitted

$$\hat{P}_{\mathrm{b}} \triangleq \frac{1}{N} \sum_{i=1}^{N} F(\mathbf{x}_i). \tag{3}$$

Note that $\mathbb{E}\left[\hat{P}_{\mathrm{b}}\right] = P_{\mathrm{b}}$. As the vectors $\mathbf{x}_i$ in (3) are independently generated, the variance of the BER estimator is given by

$$\mathbb{E}\left[\left(\hat{P}_{\mathrm{b}} - P_{\mathrm{b}}\right)^2\right] = \frac{1}{N}\left(\mathbb{E}\left[F^2(\mathbf{x})\right] - P_{\mathrm{b}}^2\right), \tag{4}$$

which can be reduced by increasing the number of simulation runs N. Nevertheless, when differences in the vectors $\{\mathbf{x}_i\}$ have a great impact on $F(\mathbf{x}_i)$, N must be prohibitively large (especially for small P_{b}) in order to obtain a sufficient estimation accuracy. When using IS, N^* vectors $\{\mathbf{x}_i\}$ are generated independently according to a biased distribution $q(\mathbf{x})$, and the BER estimate is computed as

$$\hat{P}_{\mathrm{b}}^* \triangleq \frac{1}{N^*} \sum_{i=1}^{N^*} F(\mathbf{x}_i) \frac{p(\mathbf{x}_i)}{q(\mathbf{x}_i)}, \tag{5}$$

where the correction factors $p(\mathbf{x}_i)/q(\mathbf{x}_i)$ guarantee an unbiased BER estimate. Note that the biased distribution $q(\mathbf{x})$ provides us with an additional degree of freedom which can be used to reduce the variance σ^{*2} of the BER estimate, which is given by

$$\sigma^{*2} = \frac{1}{N^*}\left(\mathbb{E}^*\left[\left(\frac{F(\mathbf{x})p(\mathbf{x})}{q(\mathbf{x})}\right)^2\right] - P_{\mathrm{b}}^2\right), \tag{6}$$

where $\mathbb{E}^*[\cdot]$ denotes expectation over the biased distribution $q(\mathbf{x})$. Using a proper biased distribution, the simulation time to estimate the BER with a given precision can be reduced substantially as compared to conventional MC simulation, i.e., $N^* \ll N$. The variance σ^{*2} is minimized when the expectation in (6) is minimized; clearly, this occurs for

$$q(\mathbf{x}) = \frac{F(\mathbf{x})p(\mathbf{x})}{P_{\mathrm{b}}}, \tag{7}$$

which yields $\sigma^{*2} = 0$. However, the biased distribution from (7) is impractical, as it depends on the unknown bit error rate P_{b} that is to be estimated by simulation. Nevertheless, (7) indicates that an efficient biased distribution should be proportional to an approximation of $F(\mathbf{x})p(\mathbf{x})$.

In this contribution, we propose an IS approach where we keep the actual PDFs for the data symbols $\mathbf{s}$ and the additive channel noise $\mathbf{w}$ unchanged and search for a convenient biased distribution $q(\mathbf{h})$ for the channel $\mathbf{h}$. Hence, we have

$$p(\mathbf{x}) = p(\mathbf{s}, \mathbf{w}, \mathbf{h}) = p(\mathbf{s}, \mathbf{w}|\mathbf{h})p(\mathbf{h}), \tag{8}$$

$$q(\mathbf{x}) = q(\mathbf{s}, \mathbf{w}, \mathbf{h}) = p(\mathbf{s}, \mathbf{w}|\mathbf{h})q(\mathbf{h}). \tag{9}$$

Using (8) and (9), it can be shown that (6) reduces to

$$\sigma^{*2} = \frac{1}{N^*}\left(\mathbb{E}^*\left[\left(\tilde{F}(\mathbf{h})\frac{p(\mathbf{h})}{q(\mathbf{h})}\right)^2\right] - P_{\mathrm{b}}^2\right), \tag{10}$$

where $\mathbb{E}^*[\cdot]$ reduces to averaging over the biased channel distribution $q(\mathbf{h})$ and $\tilde{F}(\mathbf{h})$ is defined as

$$\tilde{F}(\mathbf{h}) = \sqrt{\mathbb{E}_{\mathbf{s},\mathbf{w}|\mathbf{h}}\left[F^2(\mathbf{s}, \mathbf{w}, \mathbf{h})\right]}, \tag{11}$$

with $\mathbb{E}_{\mathbf{s},\mathbf{w}|\mathbf{h}}[\cdot]$ denoting expectation over the conditional PDF $p(\mathbf{s}, \mathbf{w}|\mathbf{h})$. Considering the similarity of (6) and (10), it follows that σ^{*2} from (10) is minimum for

$$q(\mathbf{h}) \propto \tilde{F}(\mathbf{h})p(\mathbf{h}), \tag{12}$$

where $\propto$ denotes proportionality. However, a closed-form expression for $\tilde{F}(\mathbf{h})$ from (11) is usually not available or too complex to yield a practical biased distribution.

In order to find a suitable biased distribution, we rearrange (5) as

$$\hat{P}_{\mathrm{b}}^* = \frac{1}{N_{\mathrm{b}}} \sum_{n=1}^{N_{\mathrm{b}}} \hat{P}_{\mathrm{b},n}^*, \tag{13}$$

where

$$\hat{P}_{\mathrm{b},n}^* = \frac{1}{N^*} \sum_{i=1}^{N^*} I_n(\mathbf{x}_i) \frac{p(\mathbf{x}_i)}{q(\mathbf{x}_i)} \tag{14}$$

is the IS estimate of the probability that a detection error for the nth bit occurs. As the variance of $\hat{P}_{\mathrm{b}}^*$ is hard to compute because of the correlation of the quantities $\hat{P}_{\mathrm{b},n}^*$, we look for the biased distribution of the form (9) that minimizes the variance of the individual terms $\hat{P}_{\mathrm{b},n}^*$ rather than the variance of $\hat{P}_{\mathrm{b}}^*$. Using the same reasoning that led to (12), this biased distribution that corresponds to the bit index n is

$$q_n(\mathbf{h}) \propto \tilde{I}_n(\mathbf{h}) p(\mathbf{h}), \tag{15}$$

where

$$\tilde{I}_n(\mathbf{h}) = \sqrt{\mathbb{E}_{\mathbf{s},\mathbf{w}|\mathbf{h}}\left[I_n^2(\mathbf{s},\mathbf{w},\mathbf{h})\right]} = \sqrt{\mathbb{E}_{\mathbf{s},\mathbf{w}|\mathbf{h}}\left[I_n(\mathbf{s},\mathbf{w},\mathbf{h})\right]}. \tag{16}$$

Note that $\mathbb{E}_{\mathbf{s},\mathbf{w}|\mathbf{h}}\left[I_n(\mathbf{s},\mathbf{w},\mathbf{h})\right]$ represents the conditional error probability of the nth bit, conditioned on $\mathbf{h}$. By introducing $P_{\mathrm{b},n}(\mathbf{h}) \triangleq \mathbb{E}_{\mathbf{s},\mathbf{w}|\mathbf{h}}\left[I_n(\mathbf{s},\mathbf{w},\mathbf{h})\right]$, the biased distribution (15) reduces to

$$q_n(\mathbf{h}) \propto \sqrt{P_{\mathrm{b},n}(\mathbf{h})} p(\mathbf{h}). \tag{17}$$

The exact expression of the conditional bit error probability $P_{\mathrm{b},n}(\mathbf{h})$ depends on the observation model and the type of receiver considered and is often unknown. Hence, a suitable approximation of $P_{\mathrm{b},n}(\mathbf{h})$ is needed to obtain a biased distribution (17) that adequately reduces the variance of the estimate $\hat{P}_{\mathrm{b},n}^*$.

3 Application to OSTBCs over Nakagami-m fading channels

In this section, we derive a proper biased channel distribution for OSTBC systems operating over spatially correlated Nakagami-m fading channels. As an approximation of the actual conditional bit error probability $P_{\mathrm{b},n}(\mathbf{h})$ of the system affected by channel estimation errors, we take the conditional bit error probability $P_{\mathrm{b},n}^{(\mathrm{ML})}(\mathbf{h})$ of a maximum-likelihood (ML) receiver with PCK, which in turn is well approximated by [8]

$$P_{\mathrm{b},n}^{(\mathrm{ML})}(\mathbf{h}) \propto Q\left(\sqrt{\beta \frac{E_s}{N_0} \|\mathbf{h}\|^2}\right), \tag{18}$$

where $Q(\cdot)$ is the Gaussian Q-function, $\beta = 2\lambda$ in the case of binary phase-shift keying (BPSK), $\beta = \frac{3\lambda}{M-1}$ in the case of M-quadrature amplitude modulation (QAM) transmission with Gray mapping, and the parameter λ depends on the specific OSTBC [7]. A further approximation involves replacing $Q(x)$ by the Chernoff upper bound $(1/2)\exp(-x^2/2)$ [8]. It follows from (17) that the resulting biased distribution is given by

$$q_n(\mathbf{h}) \propto \exp\left(-\frac{\beta}{4}\frac{E_s}{N_0}\|\mathbf{h}\|^2\right) p(\mathbf{h}), \tag{19}$$

which is independent of the bit index n, so that the bit index can be dropped. Hence, the same biased distribution can be used to reduce the variance of the individual bit error probability estimates $\hat{P}_{\mathrm{b},n}^*$ for $n = 1, \ldots, N_{\mathrm{b}}$, and therefore, it is expected to efficiently reduce the variance of the BER estimate $\hat{P}_{\mathrm{b}}^*$ as well.

The Nakagami-m distribution is a versatile statistical distribution able to accurately model a variety of fading environments by selecting a proper value for the fading parameter $m \geq 1/2$ [8]. It also includes the Rayleigh distribution for $m = 1$. Assuming a MIMO channel vector $\mathbf{h} = [h_1, \ldots, h_L]$ with L elements, the correlation between the channel coefficients is represented by the $L \times L$ power correlation matrix $\boldsymbol{\Sigma}$, the entries of which are defined as ([8] Eq. 9.195)

$$(\boldsymbol{\Sigma})_{i,n} \triangleq \frac{\mathrm{cov}(\alpha_i^2, \alpha_n^2)}{\sqrt{\mathrm{var}(\alpha_i^2)\,\mathrm{var}(\alpha_n^2)}}, \tag{20}$$

where $\alpha_\ell = |h_\ell|$ and $i, n = 1, 2, \ldots, L$. For integer and identical fading parameters, i.e., $m_\ell = m$, $\forall \ell$, it is shown in [9] that L correlated Nakagami-m RVs α_ℓ can be obtained from $2m$ i.i.d. real-valued zero-mean (ZM) Gaussian random vectors $\{\mathbf{y}_1, \ldots, \mathbf{y}_{2m}\}$, where $\mathbf{y}_k = \left[y_{k,1}, y_{k,2}, \ldots, y_{k,L}\right]^{\mathrm{T}}$ and $y_{k,l}$ with $k = 1, 2, \ldots, 2m$ and $l = 1, 2, \ldots, L$ are called auxiliary variables. In particular, by defining

$$\alpha_\ell^2 \triangleq \sum_{k=1}^{2m} y_{k,\ell}^2, \tag{21}$$

it is readily verified that α_ℓ's are correlated Nakagami-m RVs with $\mathbb{E}\left[\alpha_\ell^2\right] = \Omega_\ell$ and power correlation matrix $\boldsymbol{\Sigma}$, if the covariance matrix $\mathbf{Q} = \mathbb{E}\left[\mathbf{y}_k\,\mathbf{y}_k^{\mathrm{T}}\right]$ of the column vectors $\mathbf{y}_k$ is given by

$$\mathbf{Q} = \frac{1}{2m}\,\boldsymbol{\Omega}^{\circ\frac{1}{2}}\,\boldsymbol{\Sigma}_{\mathbf{G}}\,\boldsymbol{\Omega}^{\circ\frac{1}{2}}, \tag{22}$$

where the $L \times L$ diagonal matrix $\boldsymbol{\Omega}$ is given by $\boldsymbol{\Omega} = \mathrm{diag}\{\Omega_1, \Omega_2, \ldots, \Omega_L\}$ and $\boldsymbol{\Sigma}_{\mathbf{G}} = \boldsymbol{\Sigma}^{\circ\frac{1}{2}}$, with $\mathbf{X}^{\circ\frac{1}{2}}$ denoting the element-wise square root of a matrix $\mathbf{X}$.

Since the channel coefficients are obtained from auxiliary variables, we derive the biased distribution of the auxiliary variables rather than of the channel coefficients. Introducing the vector $\mathbf{y} = \left[\mathbf{y}_1^T, \mathbf{y}_2^T, \ldots, \mathbf{y}_{2m}^T\right]^T$ and taking into account that $||\mathbf{h}||^2 = ||\mathbf{y}||^2 = \sum_{k=1}^{2m} \mathbf{y}_k^T \mathbf{y}_k$ and the vectors $\{\mathbf{y}_k\}$ are i.i.d., the biased distribution $q(\mathbf{y})$ of the auxiliary random variables reduces to $q(\mathbf{y}) = \prod_{k=1}^{2m} q_0(\mathbf{y}_k)$, where

$$q_0(\mathbf{y}_k) \propto \exp\left(-\frac{\beta}{4}\frac{E_s}{N_0}\mathbf{y}_k^T\mathbf{y}_k\right) p_0(\mathbf{y}_k), \qquad (23)$$

and the PDF $p_0(\mathbf{y}_k)$ of the random vector $\mathbf{y}_k$ is ZM Gaussian with covariance matrix $\mathbf{Q}$, i.e.,

$$p_0(\mathbf{y}_k) = \frac{1}{(2\pi)^{\frac{L}{2}}\sqrt{\det(\mathbf{Q})}} \exp\left(-\frac{1}{2}\mathbf{y}_k^T\mathbf{Q}^{-1}\mathbf{y}_k\right). \qquad (24)$$

It follows from (23) and (24) that $q(\mathbf{y}_k)$ is the joint PDF of L correlated ZM Gaussian RVs with a covariance matrix $\mathbf{Q}'$ given by

$$\mathbf{Q}' = \left(\mathbf{Q}^{-1} + \frac{\beta}{2}\frac{E_s}{N_0}\mathbf{I}_L\right)^{-1}. \qquad (25)$$

Taking the PDFs $p_0(\mathbf{y}_k)$ and $q_0(\mathbf{y}_k)$ into account, it follows that the correction factor $p(\mathbf{x})/q(\mathbf{x})$ in (5) depends only on $\mathbf{y}$ and is given by

$$\frac{p(\mathbf{x})}{q(\mathbf{x})} = \frac{p(\mathbf{s}, \mathbf{w}, \mathbf{y})}{q(\mathbf{s}, \mathbf{w}, \mathbf{y})} = \frac{\exp\left(\frac{\beta}{4}\frac{E_s}{N_0}\|\mathbf{y}\|^2\right)}{\left[\det\left(\mathbf{I}_L + \frac{\beta}{2}\frac{E_s}{N_0}\mathbf{Q}\right)\right]^m}. \qquad (26)$$

It is important to note that, although (18) is obtained under the assumption of PCK, the resulting biased distribution is convenient also in the case of imperfect channel estimation (ICE), as demonstrated in Section 4, or in the presence of impairments, such as residual frequency offset, IQ imbalance, and phase noise, in which case $P_{\mathrm{b},n}^{(\mathrm{ML})}(\mathbf{h})$ in (18) is regarded as an approximation of the actual $P_{\mathrm{b},n}(\mathbf{h})$. In addition, the proposed IS technique can easily be extended to the case of multicarrier communication on dispersive MIMO channels, using for instance OSTBCs or orthogonal space-frequency block codes. In multicarrier systems, the relevant channel distribution is the joint distribution of the channel transfer function values at the subcarrier frequencies, which is to be derived from the joint distribution of the channel impulse response samples. In the important case of jointly Gaussian channel impulse response samples (i.e., Rayleigh or Rice fading), the channel transfer function values are also jointly Gaussian.

4 Numerical results

In this section, we illustrate the computation time savings that can be achieved by using the biased sampling distribution proposed in Section 3. In order to guarantee a certain accuracy for the simulated BER results, we require that the ratio of the variance of the simulated BER to the square of its expectation does not exceed a prescribed value ϵ^2

$$\frac{\mathrm{var}\left[\hat{P}_{\mathrm{b}}\right]}{\left(\mathbb{E}\left[\hat{P}_{\mathrm{b}}\right]\right)^2} \le \epsilon^2, \qquad (27)$$

where $\hat{P}_{\mathrm{b}}$ is given by (3) or (5). Since the N realizations $\{\mathbf{x}_i\}$ of the vector $\mathbf{x}$ are independently generated, we have $\mathrm{var}\left[\hat{P}_{\mathrm{b}}\right] = \frac{1}{N}\mathrm{var}\left[F(\mathbf{x})\right]$, so that for a given accuracy ϵ^2, the number of simulation runs N needs to satisfy

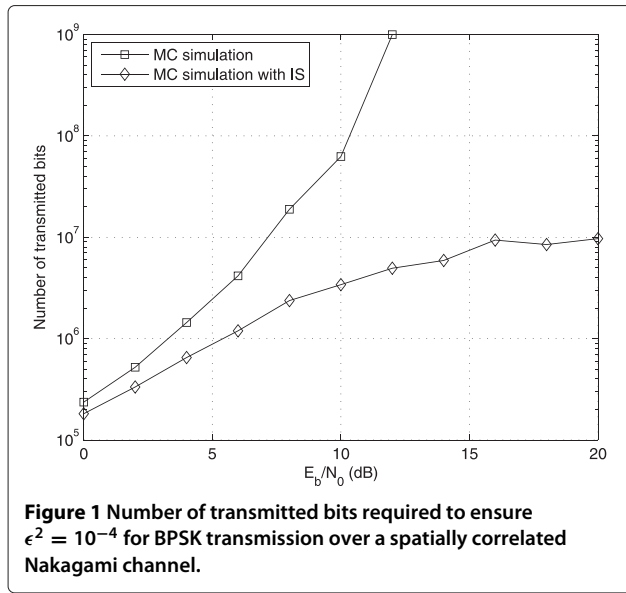$$N \ge \frac{\mathrm{var}\left[F(\mathbf{x})\right]}{\epsilon^2 P_{\mathrm{b}}^2}. \qquad (28)$$

In the following, the lower bound (28) will be computed from the simulations by substituting P_{b} and $\mathrm{var}\left[F(\mathbf{x})\right]$ by the sample mean and variance resulting from $\{F(\mathbf{x}_i)\}$, respectively.

First, we consider BPSK transmission over a 1×2 SIMO channel affected by Nakagami-m fading with fading parameter $m = 2$ and power correlation matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix}. \qquad (29)$$

The receiver has perfect channel knowledge and performs ML detection. In the simulations, for each BPSK symbol, a new realization of $\mathbf{x}$ is generated. For MC simulations with and without IS, Figure 1 shows the minimal number of realizations N required to ensure a given accuracy associated with $\epsilon^2 = 10^{-4}$ as a function of E_b/N_0, where $E_b = E_s/\log_2(M)$ denotes the energy per information bit (for BPSK, $M = 2$). Although in both cases the required simulation time increases with the signal-to-noise ratio (SNR), it is observed that huge run-time savings of several orders of magnitude can be achieved by using the proposed biased sampling distribution, especially for high SNR values where the BER is low.

Now we consider a 2×2 MIMO system employing Alamouti's code [10]. The receiver performs pilot-based linear minimum mean-square error (LMMSE) channel estimation, followed by mismatched ML detection [11]. To this end, the transmission is organized in frames containing 100 coded data symbols and 14 pilot symbols. The channel coefficients, which are assumed to remain constant during a frame, are distributed according to the

Figure 1 Number of transmitted bits required to ensure $\epsilon^2 = 10^{-4}$ for BPSK transmission over a spatially correlated Nakagami channel.

Nakagami-m distribution with $m = 2$ and have a power correlation matrix given by $\mathbf{\Sigma} = \mathbf{\Sigma}_t \otimes \mathbf{\Sigma}_r$, with $\mathbf{\Sigma}_r = \mathbf{I}_2$ and

$$\mathbf{\Sigma}_t = \begin{bmatrix} 1 & 0.64 \\ 0.64 & 1 \end{bmatrix}. \tag{30}$$

In the simulations, a fixed number of $N = 10^4$ data frames is used, and for each data frame, a new realization of $\mathbf{x}$ is generated. Figure 2 displays the simulated BER, for both 4-QAM and 16-QAM transmissions. With IS, $N = 10^4$ frames yield accurate results for BER values down to 10^{-10}. In particular, an accuracy of $\epsilon^2 = 0.025$ is obtained for $E_b/N_0 = 20$ dB and 4-QAM transmission. When no IS is used, a substantial
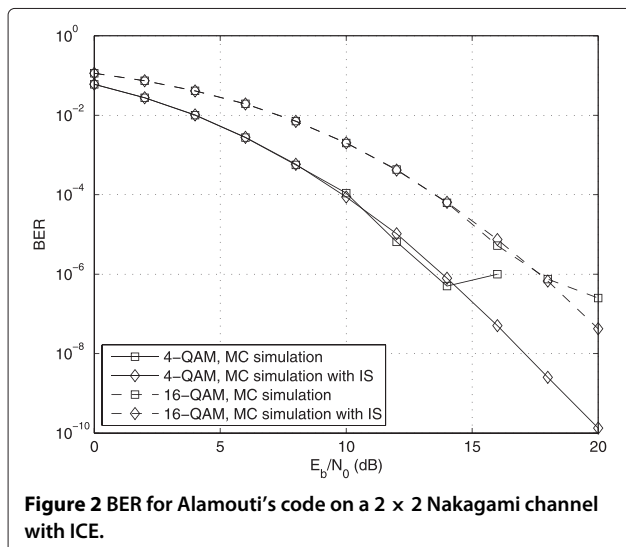
deterioration of the accuracy can be observed for BER values below 10^{-4}; for BER $\in (10^{-6}, 10^{-4})$, the deviation w.r.t. IS is up to 37.8%, and for BER values below 10^{-6}, the results without IS are highly unreliable. The number of frames required to obtain an accuracy of $\epsilon^2 = 0.025$ with and without using IS is shown in Figure 3 for 4-QAM and 16-QAM transmissions. Note that for given E_b/N_0, 16-QAM requires fewer frames than does 4-QAM to achieve a given estimation accuracy, because the formed constellation yields a larger BER than the latter. Again, it is observed from the figure that substantial run-time savings can be achieved using the proposed IS technique.

Let us now define the efficiency gain γ_{IS} that is attained by applying IS as the ratio of the number of frames required to achieve a certain accuracy ϵ^2 by using straightforward MC simulations to the number of frames required to achieve the same accuracy when using IS. Obviously, this efficiency gain is directly related to the run-time savings achieved by IS. From (28), it follows that

$$\gamma_{IS} = \frac{\text{var}\left[F(\mathbf{x})\right]|_{MC}}{\text{var}\left[F(\mathbf{x})\right]|_{IS}}, \tag{31}$$

where the variances $\text{var}\left[F(\mathbf{x})\right]|_{IS}$ and $\text{var}[F(\mathbf{x})]|_{MC}$ for MC simulations with and without IS, respectively, are obtained from the MC simulation results $\{F(\mathbf{x}_i)\}$ conducted with and without using IS. In order to investigate the impact of spatial correlation and the number of receive antennas on the efficiency gain, we consider Alamouti's code transmitted on a Nakagami-m MIMO channel with fading parameter $m = 2$ and captured by an ML receiver equipped with 1, 2, or 3 antennas. Taking into account that the power correlation matrix of the MIMO channel is given by $\mathbf{\Sigma} = \mathbf{\Sigma}_t \otimes \mathbf{\Sigma}_r$, the diagonal elements of $\mathbf{\Sigma}_t$ and $\mathbf{\Sigma}_r$ are assumed to be given by 1, whereas

Figure 2 BER for Alamouti's code on a 2 × 2 Nakagami channel with ICE.

Figure 3 Number of transmitted frames required to ensure $\epsilon^2 = 0.025$.
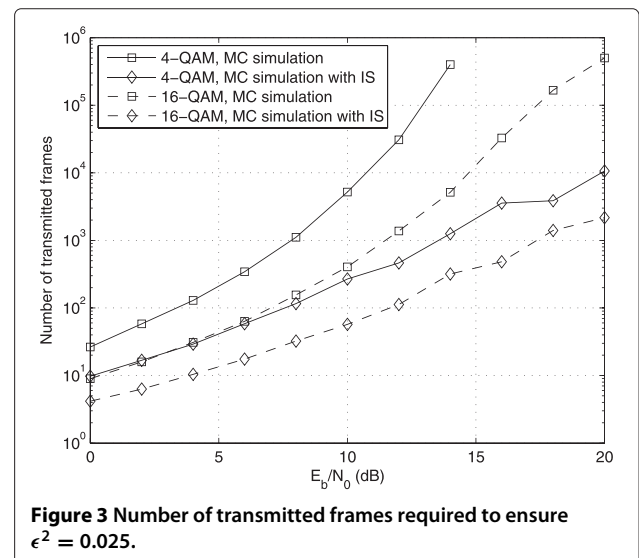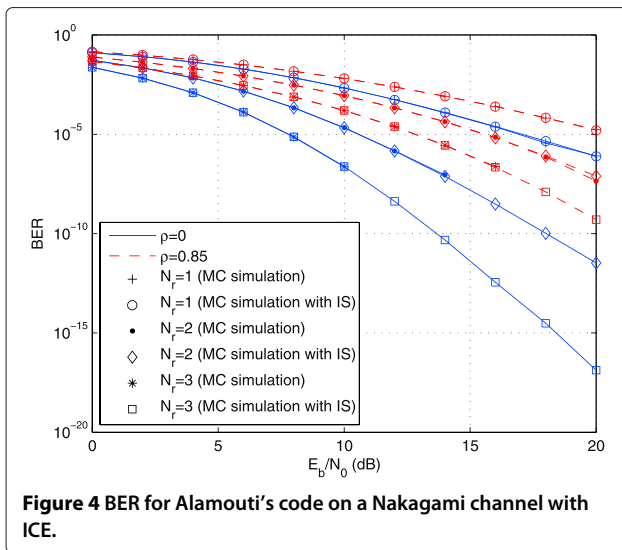
Figure 4 BER for Alamouti's code on a Nakagami channel with ICE.

the non-diagonal elements are given by ρ. Furthermore, we consider 4-QAM transmission, LMMSE channel estimation, and data frames consisting of 100 coded data symbols and 14 pilot symbols per transmit antenna. In Figure 4, the simulated BER is shown for a fixed number of $N = 10^6$ data frames transmitted. Obviously, straightforward MC simulation allows accurate BER estimation down to about 10^{-7}, whereas employing IS enables accurate BER values down to 10^{-17} with the same number of data frames. In addition, it is observed from the figure that spatial correlation with $\rho = 0.85$ results in a substantial degradation of the BER as compared to the case where all spatial channels are uncorrelated ($\rho = 0$). In Figure 5, we display the efficiency gain γ_{IS} achieved by IS corresponding to the BER curves from Figure 4. For SNR
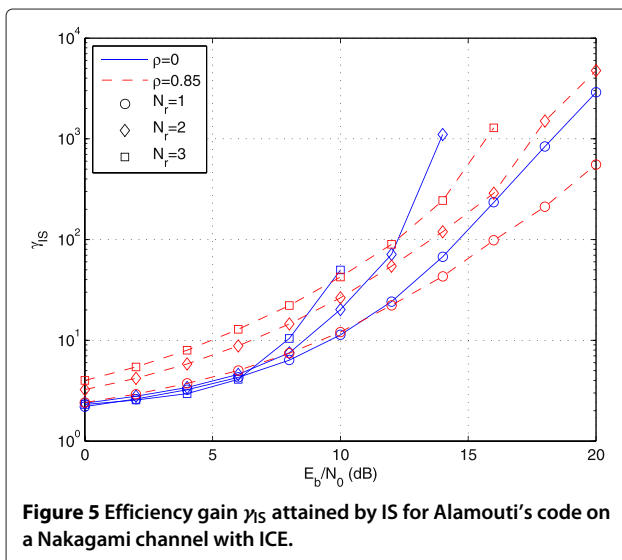
values below 10 dB, efficiency gains between 2 and 50 can be observed. For SNR values above 10 dB, however, rapidly increasing efficiency gains up to $5 * 10^3$ can be appreciated. Moreover, even higher efficiency gains are to be expected for the receiver with three antennas and, in the case of uncorrelated channels, the dual-antenna receiver. In general, it follows from the figure that for high SNR, the efficiency gain grows when the BER decreases, i.e., for increasing number of receive antennas and decreasing level of spatial correlation. As a result, the proposed IS sampling distribution allows practical BER simulation for OSTBC systems on spatially uncorrelated and correlated Nakagami fading channels, even when very low BERs are considered.

5 Conclusions

In this contribution, we presented a simple but efficient IS technique to speed up by orders of magnitude the BER simulations of OSTBC systems operating on spatially correlated frequency-flat Nakagami-m fading channels. While maintaining the actual distributions for the channel noise and the data symbols, we proposed a suitable biased distribution for the fading channel.

This IS technique can be extended to the case of multicarrier communication on dispersive MIMO channels, using for instance OSTBCs or orthogonal space-frequency block codes (OSFBCs). In multicarrier systems, the relevant channel distribution is the joint distribution of the channel transfer function values at the subcarrier frequencies, which is to be derived from the joint distribution of the channel impulse reponse samples. In the important case of jointly Gaussian channel impulse response samples, i.e., Rayleigh or Rice fading, the channel transfer function values are also jointly Gaussian.

Figure 5 Efficiency gain γ_{IS} attained by IS for Alamouti's code on a Nakagami channel with ICE.

References
1. PW Glynn, DL Iglehart, Importance sampling for stochastic simulation. Manage. Sci. **35**, 1367–1392 (1989)
2. K Shanmugam, P Balaban, A modified Monte-Carlo simulation technique for the evaluation of error rate in digital communication systems. IEEE Trans. Commun. **28**(11), 1916–1924 (1980)
3. D Lu, K Yao, Improved importance sampling technique for efficient simulation of digital communication systems. IEEE J. Select. Areas Commun. **6**(1), 67–75 (1988)
4. DI Kim, DI Lee, CQ importance sampling technique for efficient simulation in Rayleigh-faded channels, in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)* (Cambridge, MA, USA, 16–21 Aug 1998), p. 158
5. KC Nguyen, U Gunawardanam, R Liyana-Pathirana, Application of importance sampling for simulation of orthogonal space-time block

coded systems, in *Asia-Pacific Conference on Communications (APCC)* (Perth, Australia, 3–5 Oct 2005), pp. 986–990

6. WA Al-Qaq, JK Townsend, A stochastic importance sampling methodology for the efficient simulation of adaptive systems in frequency nonselective Rayleigh fading channels. IEEE J. Sel. Area Comm. **15**(4), 614–625 (1997)

7. L Jacobs, GC Alexandropoulos, M Moeneclaey, H Bruneel, PT Mathiopoulos, Analysis and efficient evaluation of the BER of OSTBCs with imperfect channel estimation in arbitrarily correlated fading channels. IEEE Trans. Signal Process. **59**(6), 2720–2733 (2011)

8. MK Simon, M-S Alouini, *Digital Communication over Fading Channels*, 2nd edn. (Wiley, New York, 2005)

9. GC Alexandropoulos, NC Sagias, FI Lazarakis, K Berberidis, New results for the multivariate Nakagami-*m* fading model with arbitrary correlation matrix and applications. IEEE Trans. Wireless Commun. **8**(1), 245–255 (2009)

10. V Tarokh, H Jafarkhani, AR Calderbank, Space-time block codes from orthogonal designs. IEEE Trans. Inform. Theory **45**(5), 1456–1467 (1999)

11. L Jacobs, M Moeneclaey, BER analysis of square OSTBCs with LMMSE channel estimation in arbitrarily correlated Rayleigh fading channels. IEEE Commun. Lett. **14**(7), 626–628 (2010)