Journal of the Korean Statistical Society (2006), **35**: *2, pp 157–166*

PERFORMANCE EVALUATION VIA MONTE CARLO IMPORTANCE SAMPLING IN SINGLE USER DIGITAL COMMUNICATION SYSTEMS†

MAN-SUK OH[1]

ABSTRACT

This research proposes an efficient Monte Carlo algorithm for computing error probability in high performance digital communication systems. It characterizes special features of the problem and suggests an importance sampling algorithm specially designed to handle the problem. It uses a shifted exponential density as the importance sampling density, and shows an adaptive way of choosing the rate and the origin of the shifted exponential density. Instead of equal allocation, an intelligent allocation of the samples is proposed so that more samples are allocated to more important part of the error probability. The algorithm uses the nested feature of the error space and avoids redundancy in estimating the probability. The algorithm is applied to an example data set and shows a great improvement in accuracy of the error probability estimation.

AMS 2000 subject classifications. Primary 65C05; Secondary 65D30.
Keywords. Monte Carlo, adaptive importance sampling, stratification.

1. INTRODUCTION

Performance evaluation in high performance digital communication system is closely related to estimation of the probability of bit error which can be converted into estimation of the integral

$$I = \int_E f(x)dx,$$

Received February 2006; accepted March 2006.
†This work was supported by the Korea Research Foundation Grant KRF-2004-015-C00079.
[1]Department of Statistics, Ewha Womans University, Seoul 120-750, Korea. (e-mail: msoh@ewha.ac.kr.)

where the area E is the error space in which communication error occurs and f is the probability density function of a random noise in the system (Stadler and Roy, 1993).

If we assume a Gaussian distribution for the noise in the system, the probability of bit error is the probability of tail area under a normal distribution. Specifically, in high performance digital communication systems the error space E is an extreme tail area under the standard normal distribution and the error probability is extremely small, often less than 10^{-4}. In addition, when the total of K binary sequences are sent, there are 2^K possible combinations of K binary data and hence we need to compute the error probability for each of the 2^K cases.

Due to the above special characteristics of the error probability in digital communication system, general purpose numerical methods can be very inefficient in estimating the error probability and hence development of an efficient algorithm taking account of the above features is desired. In this paper, we attack the problem by using a Monte Carlo method, specifically an importance sampling scheme.

In Section 2, we present the single user digital communication system and describe the special features of the problem of estimating the error probability. In Section 3, we propose an efficient importance sampling algorithm taking into account of the features of the problem. We apply the proposed algorithm in an example data set and compare the results with those from Stadler and Roy (1993) in Section 4. Summary and discussion are given in Section 5.

2. THE SINGLE USER DIGITAL COMMUNICATION SYSTEM

Consider the single user digital communication system model shown in Figure 2.1. The user symbols or data is converted into a binary sequence ($d_k = \pm A$) and sent to the channel. The channel is a linear filter which is a MA($p+1$) process. A Gaussian noise w_k is added to the output from the channel and a linear combination of the noise and the output from the channel is passed through the equalizer which is a MA($q+1$) process, and the output from the equalizer y_k is given to the receiver. The data received by the receiver is continuous due to the Gaussian noise, hence is converted into a binary sequence $\hat{d}_k$ depending on whether the continuous data is above a threshold χ_0 or not. The converted binary data is a final output from the digital communication system. Here we assume that the binary sequence d_k are *iid* and $P(d_k = A) = P(d_k = -A)$, and that the Gaussian noise w_k are *iid* and follow $N(0, \sigma^2)$ distribution. The coefficient $b_0, \ldots, b_p$ of the MA($p+1$) channel and the coefficient $h_0, \ldots, h_q$ of the MA($q+1$) equalizer can be estimated in a training stage and hence are assumed to be known constants.

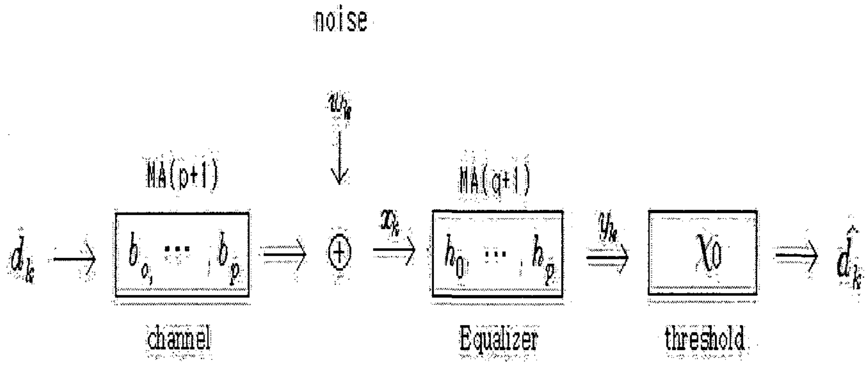

FIGURE 2.1 *The single user digital communication system.*

Given $\{b_0, \ldots, b_p\}$, $\{h_0, \ldots, h_q\}$, $\{d_1, \ldots, d_k\}$, x_k and y_k can be represented as

$$x_k = \sum_{i=0}^{p} b_i d_{k-i} + w_k = \mathbf{b}^t \mathbf{d}_k + w_k,$$
$$y_k = \sum_{i=0}^{q} h_i x_{k-i} = \mathbf{h}^t \mathbf{x}_k,$$

where

$$\mathbf{b} = (b_0, \ldots, b_p), \qquad \mathbf{d}_k = (d_k, \ldots, d_{k-p}),$$
$$\mathbf{h} = (h_0, \ldots, h_q), \qquad \mathbf{x}_k = (x_k, \ldots, x_{k-q}),$$

and $w_k \sim N(0, \sigma^2)$. If we let $c_j = \sum_{l=0}^{j} h_l b_{j-l}, \;\; j = 0, \ldots, p$ and let

$$\tilde{w_k} = \mathbf{h}^t \mathbf{w}_k = \sum_{i=0}^{q} h_i w_{k-i}, \quad \mathbf{w}_k = (w_k, \ldots, w_{k-q}),$$

then $y_k = \mathbf{c}^t \mathbf{d}_k + \tilde{w_k}$. Here $\tilde{w_1}, \ldots, \tilde{w_k}$ are not independent but are correlated.

For the k^{th} signal d_k, the error probability is $P(error) = 0.5P(\hat{d}_k = -A|d_k = A) + 0.5P(\hat{d}_k = A|d_k = -A)$. By symmetry $P(\hat{d}_k = -A|d_k = A) = P(\hat{d}_k = A|d_k = -A)$. Hence, it suffices to consider the case that $d_k = A$ is sent. Specifically, the problem of interest is computation of

$$\begin{aligned} P(\hat{d}_k = -A|d_k = A) &= P(y_k \leq 0|d_k = A) = P(\mathbf{c}^t\mathbf{d}_k + \mathbf{h}^t\mathbf{w}_k \leq 0|d_k = A) \\ &= P(\mathbf{h}^t\mathbf{w}_k \leq -\mathbf{c}^t\mathbf{d}_k|d_k = A), \end{aligned}$$

where $\chi_0 = 0$ is used as the threshold.

Given $\mathbf{h}$, $\mathbf{c}$, and $\mathbf{d}_k = (d_k, \ldots, d_{k-p})$, $\mathbf{h}^t\mathbf{w}_k$ is a linear combination of normal variables and follows a normal distribution. Hence the problem converts to a generic problem of computing $P(Z > \gamma_k)$, where Z is a standard normal variable and $\gamma_k = (\mathbf{c}^t\mathbf{d}_k + \mu_k)/\sigma_k^2$ and μ_k, σ_k^2 are respectively the mean and variance of $\mathbf{h}^t\mathbf{w}_k$. Since there are $m = 2^p$ possible cases of $\mathbf{d}_k = (d_k, \ldots, d_{k-p})$, given a value of d_k, which are assumed to be equally probable, $P(Z > \gamma_k) = m^{-1}\sum_{i=1}^m P(Z > a_i)$, where a_i corresponds to the i^{th} possible value of γ_k. Therefore, it requires $m = 2^p$ computations of standard normal probabilities.

There are several special features of the problem. First, the value a_i are very large in high performance digital communication systems so that $p_i = P(Z > a_i)$'s are extreme tail probabilities of the standard normal distribution and are very small. The cumulative normal probabilities provided by typical statistical software packages can provide inaccurate results. Therefore, an efficient numerical method specially designed for our problem is required. It has been shown that Monte Carlo methods, especially importance sampling methods, are effective methods for accurate estimation of the probabilities and significant effort has been given to improve the efficiency of the importance sampling methods (Shanmugam and Balaban, 1980; Jeruchim, 1984; Davis, 1986; Beaulieu, 1990a, b; Schlebusch, 1990; Stadler and Roy, 1993). However, the complexity of choosing an appropriate sample generating density, called the importance sampling density, has prohibited widespread use of the importance sampling methods.

Second, without loss of generality we may let $a_1 < a_2 < \cdots < a_m$, hence $p_i = P(Z > a_i) = P(Z > a_{i+1}) + P(a_{i+1} > Z > a_i) = p_{i+1} + P(a_{i+1} > Z > a_i)$. Thus, if we ignore the above nested feature of p_i's and compute the p_i's separately then it would yield redundant estimation of p_j, $j = i+1, \ldots, m$ for each i, yielding inefficiency of the algorithm.

Third, due to the order in a_i's, it holds that $p_1 > p_2 > \cdots > p_m$. Since the standard normal density has a fast convergence rate and a_i's are all very large, the first few p_i's are much larger, usually in the order of magnitudes, than the others so that the total error probability is dominated by the first few p_i's. Therefore, instead of dealing with all p_i's with the same effort it would be desirable to use most of the computing cost to first few p_i's and use relatively small computing cost to the rest p_i's.

3. AN IMPORTANCE SAMPLING ALGORITHM

In this section we propose an efficient importance sampling algorithm, taking account of each of the special features described in Section 2. First, the pdf of the standard normal distribution is monotone decreasing in the tail region $Z \in (a, \infty)$ with rate $e^{-0.5z^2}$. Thus, an importance sampling density which is monotone decreasing in the given region would be desirable. Also, it is desirable that the decreasing rate of the importance sampling density is as close as possible to the rate of the standard normal density in the given region. Note that the decreasing rate of the standard normal pdf in (a, ∞) is so fast that the region very close to a, say $(a, a + \varepsilon)$ for very small $\varepsilon > 0$, is dominant in computing the probability, and this phenomenon is more pronounced for a larger a. Thus, for a large a it would be sufficient to consider the decreasing rate near a when choosing the importance sampling density. In addition, the tails of the importance sampling density should be heavier than those of $\phi(z)$, the standard normal pdf, in the given region to avoid a large variance in importance sampling estimates (Oh and Berger, 1992).

Keeping these considerations in mind, we suggest to use the pdf of a shifted exponential distribution, $h_\theta(z) = \theta e^{-\theta(z-a)} I(a < z < \infty)$, as an importance sampling density. The rate θ should be chosen so that $h_\theta(z)$ mimics $\phi(z)$, the standard normal pdf, in $Z \in (a, \infty)$ as much as possible. For this, we note $\phi(a)/\phi(a+\varepsilon) = e^{a\varepsilon+\varepsilon^2}$ and $h_\theta(a)/h_\theta(a+\varepsilon) = e^{\theta\varepsilon}$. If we let $\theta = a$ then $\phi(a)/\phi(a+\varepsilon) \approx h_a(a)/h_a(a+\varepsilon)$ for small ε but also $h_a(z)$ has heavier tails than $\phi(z)$ since $\phi(a)/\phi(a+\varepsilon) > h_a(a)/h_a(a+\varepsilon)$ for any $\varepsilon > 0$. Thus, we choose $h_a(z) = ae^{-a(z-a)} I(a < z < \infty)$ as the importance sampling density for estimating $P(Z > a)$. Note that Oh and Kim (1994) also chose $h_a(z)$ for estimating the extreme tail probability of t-distributions by importance sampling method.

Second, if we let $C_i = P(a_i < Z < a_{i+1})$ and $a_{m+1} = \infty$ then $p_i = p_{i+1} + C_i$ for $i = 1, \ldots, m$. Since, given p_{i+1}, we only need to estimate C_i for p_i, we suggest to estimate C_i's separately and use them to estimate p_i. Note that, the regions $(a_i < Z < a_{i+1})$'s are disjoint and C_i needs to be estimated only once. As the importance sampling density for estimation of C_i, we use the pdf of the shifted exponential distribution restricted to $(a_i < Z < a_{i+1})$, *i.e.*, $g_i(z) = h_{a_i}(z)I(a_i < z < a_{i+1})/A_i$, where $A_i = \int_{a_i}^{a_{i+1}} h_{a_i}(z)dz = e^{-a_i(a_{i+1}-a_i)}$.

Third, since a few p_i's, *i.e.*, a few C_i's are dominant in computing the total error probability, it would be desirable to use a large number of samples for the dominant C_i's and a small number of samples for the rest. Accurate estimation of the dominant C_i's are important while relatively inaccurate estimation of negligible C_i's does not significantly affect the accuracy of the total error probability. This idea of intelligent stratification (allocation) of samples can be achieved as follows. Given the total of N samples, allocate $N_i = N \cdot A_i\phi(a_i)/\sum_{i=1}^{m} A_i\phi(a_i)$ samples for estimating C_i. Note that $A_i\phi(a_i)$ is an approximation to $C_i = P(a_i < Z < a_{i+1})$ and the probability C_i can be considered as a measure of contribution of the region (a_i, a_{i+1}) to the total error probability. For more improvement of the allocation, one may run a preliminary importance sampling with $N_i \propto A_i\phi(a_i)$, estimate C_i, and then run the actual importance sampling with N_i proportional to the estimated C_i.

Allocation of $N_i \propto A_i\phi(a_i)$ or $N_i \propto \hat{C}_i$ can make some of N_i's equal to zero. This often happens in practice since there are negligibly small C_i's. Deletion of those C_i's in computation would save computing cost but it would bring a bias in estimating the total error probability. Moreover, very small N_i's, though not zero, may yield inaccurate estimates of C_i's and cumulation of inaccurate C_i's may affect estimation of the total error probability. Thus, we suggest to equally allocate a small portion, say $r \in (0,1)$, of N to all C_i's and allocate the rest proportional to C_i, taking compromise between equal allocation and stratified allocation.

We summarize the idea in the following algorithm.

Importance sampling algorithm for the error probability

1. Input $N, r, \sigma^2, \mathbf{b}, \mathbf{h}, p, d_k$.

2. From possible 2^p combinations of $d_{k-1} = \pm A, \ldots, d_{k-p} = \pm A$, compute $a_i^* = -\mathbf{h}^t\mathbf{d}$, where $\mathbf{d} = (d_k, d_{k-1}, \ldots, d_{k-p})^t$, for $i = 1, \ldots, 2^p$.

3. Sort $\{a_i^*\}$ such that $a_1^* < \cdots < a_m^*$, where $m = 2^p$.

4. Let $a_i = \frac{a_i^*}{\sigma}$ and $A_i = e^{-a_i(a_{i+1}-a_i)}$.

5. For $i = m, m-1, \ldots, 1$, do;

5.1. Let $N_i = \frac{rN}{m} + (N - rN)\frac{A_i\phi(a_i)}{\sum_{i=1}^{m} A_i\phi(a_i)}$.

5.2. Generate $X_{ij} \sim g_i(x) = \frac{a_i e^{-a_i(x-a_i)}}{A_i} \frac{}{I(a_i<x<a_{i+1})}$, $j = 1, \ldots, N_i$.

5.3. Compute $w_{ij} = \frac{\phi(X_{ij})}{g_i(X_{ij})}$, $j = 1, \ldots, N_i$.

5.4. Let

$$\hat{C}_i = \frac{1}{N_i}\sum_{j=1}^{N_i} w_{ij},$$

$$\hat{Var}(\hat{C}_i) = \frac{1}{N_i}(\frac{1}{N_i}\sum_{j=1}^{N_i} w_{ij}^2 - \hat{C}_i^2).$$

5.5. Let

$$\hat{p}_i = \hat{C}_i + \hat{p}_{i+1}, \text{ where } \hat{p}_{m+1} \equiv 0$$
$$\hat{Var}(\hat{p}_i) = \hat{Var}(\hat{C}_i) + \cdots + \hat{Var}(\hat{C}_m).$$

6. Estimate the total error probability and its variance by

$$\hat{P}(error) = \frac{1}{m}\sum_{i=1}^{m} \hat{p}_i,$$
$$\hat{Var}(\hat{P}(error)) = \frac{1}{m^2}[\hat{Var}(\hat{C}_1) + 2^2\hat{Var}(\hat{C}_2) + 3^2\hat{Var}(\hat{C}_3) + \cdots + m^2\hat{Var}(\hat{C}_m)].$$

4. AN EXAMPLE

To see the efficiency of the proposed algorithm we apply the algorithm to the case of $p = 2$, $A = 5.77$, $\mathbf{b} = (1, 0.2, 0.1)$, $\mathbf{h} = (1, 0.2, 0.1)$, $\sigma^2 = 1$, $d_k = -A$, $r = 0.4$. We use the threshold $\chi_0 = 0$ and try the total number of samples $N = 5,000, 10,000, 15,000, \ldots, 50,000$. We compare the probabilities and their SE's from the proposed algorithm with those from Stadler and Roy (1993) in which the pdf of a shifted normal distribution $N(0,1)I(x > a)$ is used as the importance sampling density and stratification is not used. The results given in Table 4.1 show that the estimates from the proposed algorithm are significantly more accurate than those from the importance sampling with the shifted normal density. The standard errors are about 30 times smaller in the proposed method, implying that the accuracy is improved significantly.

TABLE 4.1 *Estimates of the error probability and their SE from the proposed importance sampling algorithm and the importance sampling by Stadler and Roy (1993).*

N	N_i	d_{k-1}	d_{k-2}	$\hat{P}_{proposed}$	$SE_{proposed}$	$\hat{P}_{S\&R}$	$SE_{S\&R}$	$\frac{SE_{proposed}}{SE_{S\&R}}$
5,000	3,485	+5.77	+5.77	0.269E-04	0.360E-07	0.273E-04	0.100E-05	0.0358
	514	+5.77	−5.77	0.102E-06	0.301E-09	0.105E-06	0.454E-08	0.0663
	500	−5.77	+5.77	0.109E-09	0.219E-12	0.111E-09	0.544E-11	0.0403
	500	−5.77	−5.77	0.316E-13	0.576E-16	0.319E-13	0.173E-14	0.0332
total	5,000			0.270E-04	0.360E-07	0.274E-04	0.100E-05	0.0358
10,000	6,970	+5.77	+5.77	0.268E-04	0.262E-07	0.252E-04	0.693E-06	0.0378
	1,029	+5.77	−5.77	0.103E-06	0.204E-09	0.966E-07	0.312E-08	0.0653
	1,000	−5.77	+5.77	0.109E-09	0.166E-12	0.101E-09	0.374E-11	0.0445
	1,000	−5.77	−5.77	0.316E-13	0.382E-16	0.292E-13	0.119E-14	0.0320
total	10,000			0.269E-04	0.262E-07	0.253E-04	0.693E-06	0.0378
15,000	10,456	+5.77	+5.77	0.268E-04	0.210E-07	0.271E-04	0.584E-06	0.0360
	1,543	+5.77	−5.77	0.103E-06	0.154E-09	0.104E-06	0.264E-08	0.0583
	1,500	−5.77	+5.77	0.109E-09	0.134E-12	0.110E-09	0.316E-11	0.0424
	1,500	−5.77	−5.77	0.317E-13	0.253E-16	0.319E-13	0.100E-14	0.0251
total	15,000			0.269E-04	0.210E-07	0.272E-04	0.584E-06	0.0360
20,000	13,941	+5.77	+5.77	0.268E-04	0.182E-07	0.263E-04	0.506E-06	0.0360
	2,058	+5.77	−5.77	0.103E-06	0.134E-09	0.101E-06	0.230E-08	0.0585
	2,000	−5.77	+5.77	0.109E-09	0.107E-12	0.108E-09	0.278E-11	0.0387
	2,000	−5.77	−5.77	0.316E-13	0.238E-16	0.312E-13	0.891E-15	0.0267
total	20,000			0.269E-04	0.182E-07	0.264E-04	0.506E-06	0.0360
25,000	17,426	+5.77	+5.77	0.268E-04	0.164E-07	0.268E-04	0.450E-06	0.0365
	2,573	+5.77	−5.77	0.103E-06	0.115E-09	0.103E-06	0.203E-08	0.0568
	2,500	−5.77	+5.77	0.109E-09	0.983E-13	0.109E-09	0.244E-11	0.0402
	2,500	−5.77	−5.77	0.316E-13	0.225E-16	0.315E-13	0.777E-15	0.0290
total	25,000			0.164E-07	0.269E-04	0.450E-06	0.365E-01	
30,000	20,912	+5.77	+5.77	0.268E-04	0.152E-07	0.273E-04	0.417E-06	0.0363
	3,087	+5.77	−5.77	0.103E-06	0.107E-09	0.105E-06	0.189E-08	0.0564
	3,000	−5.77	+5.77	0.109E-09	0.958E-13	0.112E-09	0.228E-11	0.0419
	3,000	−5.77	−5.77	0.317E-13	0.187E-16	0.325E-13	0.729E-15	0.0257
total	30,000			0.269E-04	0.152E-07	0.274E-04	0.417E-06	0.0363
35,000	24,397	+5.77	+5.77	0.268E-04	0.138E-07	0.265E-04	0.381E-06	0.0362
	3,602	+5.77	−5.77	0.103E-06	0.103E-09	0.102E-06	0.172E-08	0.0597
	3,500	−5.77	+5.77	0.109E-09	0.858E-13	0.108E-09	0.207E-11	0.0414
	3,500	−5.77	−5.77	0.316E-13	0.187E-16	0.312E-13	0.661E-15	0.0283
total	35,000			0.269E-04	0.138E-07	0.266E-04	0.381E-06	0.0362

(*continued*)

N	N_i	d_{k-1}	d_{k-2}	$\hat{P}_{proposed}$	$SE_{proposed}$	$\hat{P}_{S\&R}$	$SE_{S\&R}$	$\frac{SE_{proposed}}{SE_{S\&R}}$
40,000	27,882	+5.77	+5.77	0.268E-04	0.131E-07	0.272E-04	0.360E-06	0.0364
	4,116	+5.77	−5.77	0.103E-06	0.988E-10	0.105E-06	0.163E-08	0.0604
	4,000	−5.77	+5.77	0.109E-09	0.780E-13	0.111E-09	0.196E-11	0.0397
	4,000	−5.77	−5.77	0.316E-13	0.166E-16	0.323E-13	0.627E-15	0.0265
total	40,000			0.269E-04	0.131E-07	0.273E-04	0.360E-06	0.0364
45,000	31,368	+5.77	+5.77	0.268E-04	0.121E-07	0.271E-04	0.336E-06	0.0361
	4,631	+5.77	−5.77	0.103E-06	0.945E-10	0.104E-06	0.152E-08	0.0621
	4,500	−5.77	+5.77	0.109E-09	0.744E-13	0.110E-09	0.182E-11	0.0407
	4,500	−5.77	−5.77	0.316E-13	0.176E-16	0.318E-13	0.581E-15	0.0303
total	45,000			0.269E-04	0.121E-07	0.272E-04	0.336E-06	0.0361
50,000	34,853	+5.77	+5.77	0.268E-04	0.116E-07	0.269E-04	0.321E-06	0.0361
	5,146	+5.77	−5.77	0.103E-06	0.863E-10	0.103E-06	0.145E-08	0.0593
	5,000	−5.77	+5.77	0.109E-09	0.710E-13	0.110E-09	0.174E-11	0.0406
	5,000	−5.77	−5.77	0.316E-13	0.165E-16	0.318E-13	0.558E-15	0.0297
total	50,000			0.269E-04	0.116E-07	0.270E-04	0.321E-06	0.0361

Next, to see the effect of stratification, we estimate the probabilities without stratification (equal allocation for each of the four cases) and compare the results with those from with stratification. Note that we use the same shifted exponential density as the importance sampling density in both procedures. The ratio of SE with stratification versus SE without stratification is about 60%, so there is about 40% reduction solely by the intelligent stratification.

5. SUMMARY AND DISCUSSION

We have proposed an efficient importance sampling algorithm for estimating the bit error probability in single user digital communication systems. Specifically, we have proposed a shifted exponential density as the importance sampling density and derived an appropriate rate of the exponential distribution. This choice of the importance sampling density is adaptive in that it adjusts the rate and the origin of the shifted exponential density according to the error space. The adaptive shifted exponential importance sampling density improves the accuracy of the estimates astonishingly as shown in the example. We also have proposed a stratified allocation of the samples such that more samples are allocated to more important part of the error probability. The stratification requires almost no extra cost but improves the accuracy significantly. Finally, the algorithm uses the nested feature of the error space and avoids redundancy in estimating the probability.

Future research interest is the computation of error probability in multi-user digital communication systems. This requires computation of tail probabilities in multivariate normal distributions in which the variables are correlated. Extension of the univariate shifted exponential distribution to a multivariate distribution which mimics the given covariance structure of the multivariate normal distribution in the tail area would be the key issue for the multi-user systems.

REFERENCES

BEAULIEU, N. C. (1990a). "An investigation of Gaussian tail and Rayleigh tail density functions for importance sampling digital communication system simulation", *IEEE Transactions on Communications*, **38**, 1288–1292.

BEAULIEU, N. C. (1990b). "A composite importance sampling technique for digital communication system simulation", *IEEE Transactions on Communications*, **38**, 393–396.

DAVIS, B. R. (1986). "An improved importance sampling method for digital communication system simulation", *IEEE Transactions on Communications*, **34**, 715–719.

JERUCHIM, M. C. (1984). "Techniques for estimating the bit error rate in the simulation of digital communication systems", *IEEE Journal on Selected Areas in Communications*, **2**, 153–170.

OH, M-S. AND BERGER, J. O. (1992). "Adaptive importance sampling in Monte Carlo integration", *Journal of Statistical Computation and Simulation*, **41**, 143–168.

OH, M-S. AND KIM, N. Y. (1994). "Efficient random generation from extreme tail areas of t distribution", *Korean Journal of Applied Statistics*, **9**, 165–177.

SCHLEBUSCH, H. J. (1990). "Nonlinear importance sampling techniques for efficient simulation of communication systems", In *Proceeding of IEEE International Conference on Communications*, 631–635.

SHANMUGAM, K. S. AND BALABAN, P. (1980). "A modified Monte-Carlo simulation technique for the evaluation of error rate in digital communication systems", *IEEE Transactions on Communications*, **28**, 1916–1924.

STADLER, J. S. AND ROY, S. (1993). "Adaptive Importance Sampling", *IEEE Journal on Selected Areas in Communications*, **11**, 309–316.