Importance Sampling for Analysis of Direct Detection Optical Communication Systems

Narayan B. Mandayam, Member, and Behnaam Aazhang, Senior Member, IEEE

Abstract—Analytical solutions of the performance of optical communication systems are difficult to obtain and often. Monte Carlo simulations are used to achieve realistic estimates of the performance of such systems. However, for high performance systems, this technique requires a large number of simulation trials for the estimates to be in a reasonable interval of confidence, with the number of trials increasing linearly with the performance of the system. We apply an Importance Sampling technique to estimate the performance of direct detection optical systems, where the “gain” of Importance Sampling over Monte Carlo simulations is shown to increase linearly with the system performance. Further, we use this technique to study the performance of optical communication systems employing avalanche photodiodes as well as fiber-optic code division multiple access systems (FO-CDMA). We also show that the quick simulation technique developed can be used for a wide variety of networking schemes, and for the first time, we present a comparative analysis of the performance of FO-CDMA systems employing optical orthogonal codes and prime sequences. In all cases, it is shown that Importance Sampling simulations require less than 50–100 trials for estimating error probabilities of $10^{-10}$ and below.

Keywords—Optical Code Division Multiple Access, Importance Sampling, Efficient Simulation Techniques, Exponential Tilt, Avalanche Photodetector, Fiber Optic Communications.

I. INTRODUCTION

In most optical communication systems in existence today, the statistics of the underlying random processes and the structure of the detection systems are quite complicated, making it very difficult to achieve analytical performance estimates. Therefore, efforts have been concentrated on obtaining asymptotic approximations and bounds for estimating system error rates. A widely used method for estimating systems performance is the Chernoff-bound [1], [2]. However, the Chernoff bound has been shown to be loose for direct-detection optical systems with avalanche photodiodes [3]. Other bounds [4] and approximations [5] have been derived, but due to the inherent complexity of such systems, these techniques are often highly inaccurate.

Another, often more tractable and versatile method for obtaining performance estimates is Monte Carlo simulation. However, the drawback of this approach is that, for systems with a very small probability of error, $P_e$, a large number of simulation trials are required to achieve a meaningful estimate of the system performance. As an example, for a 95% confidence interval of \( \frac{2P_e}{3}, \frac{8P_e}{5} \) we require approximately $10/P_e$ simulation trials [6]. This can be quite large for $P_e \leq 10^{-7}$, thus making conventional Monte Carlo methods impractical for simulating optical detection systems.

A technique known as Importance Sampling has been employed to greatly reduce the number of simulation trials [7], [8], [9]. This technique is based on biasing the noise distribution such that more samples are taken from the important regions in the observation space (regions that cause errors) and then unbiased the estimate of system performance with a weighting function. The effectiveness of this method applied to communications systems problems has been demonstrated [7], [8], [9], [10], [11], [12], [13], [14], but in general, Importance Sampling has been rarely applied to optical systems with direct detection and more so for fiber based optical networks [15].

The technique of fiber-optic code division multiple access (FO-CDMA) provides random, asynchronous communications access, free of network control among many users. A class of optical codes called optical orthogonal codes (OOC's) [16], [17] and Prime Sequence codes [18] have been applied to FO-CDMA systems, but the Poisson characteristics of the photodetectors have not been taken into consideration for analysis. An analysis of the error rate for optical CDMA systems has been proposed [19] and bounds have been developed. In all the above works, the efforts have been concentrated on obtaining asymptotic approximations and bounds for estimating the error rates and therefore, there is a great need for developing Importance Sampling techniques of Monte Carlo simulations.

In this paper, we analyze three different optical communication systems and arrive at estimates of the performance of these systems through both Monte Carlo and Importance Sampling simulation techniques. We then proceed to analyze the “gain” using Importance Sampling simulations over standard Monte Carlo simulations by comparing the number of simulation trials required by each method to obtain a fixed estimator accuracy. We will consider a simplified single user binary communication system employing an ideal avalanche photodiode just to introduce the development of Importance Sampling techniques for the analysis of direct detection optical systems. The ideas developed will be formulated to analyze more complicated systems for which exact estimates of system performance

Paper approved by Keith Townsend, the Editor for Computer-Aided Design of Communication Systems of the IEEE Communications Society. Manuscript received: January 2, 1992; revised November 20, 1992; March 9, 1993. This work was supported by the Advanced Technology Program of the Texas Higher Education Coordinating Board under Grant 003564-018. This paper was presented in part at the 25th Annual Conference on Information Sciences and Systems, March 20-22, 1991. Johns Hopkins University, Baltimore, MD.

N.B. Mandayam was with the Department of Electrical & Computer Engineering, Rice University, Houston, TX 77251. He is now with WINLAB, Rutgers University, New Brunswick, NJ 08901.

B. Aazhang is with the Department of Electrical & Computer Engineering, Rice University, Houston, TX 77251-1892.

IEEE Log Number 9410555.
are hard to obtain. In particular, we will analyze a single user system employing a random gain avalanche photodiode. For the multi user case, we will consider a FO-CDMA network and develop an Importance Sampling technique of Monte Carlo simulations for the network, taking into consideration the Poisson statistics of the system. We will use the above technique to compare the performance of systems employing OOC’s with those using Prime Sequences. However, for all the systems considered, we assume that the multiple access interference is dominant and the effects of inter symbol interference are negligible. We will also develop lower and upper bounds for the “gain” of Importance Sampling simulations over Monte Carlo simulations, as it may not be possible to calculate the exact “gain” and only simulated values of the “gain” are available.

II. SYSTEMS WITH IDEAL PHOTODIODES

A received optical signal can be modeled as a photon arrival process. In a simple setting, this process has a Poisson density with arrival rate proportional to the received optical power plus a dark current component dependent on the photodetector physical structure. The photodetector then converts the incident optical power, in terms of the number of photons seen, to an electrical signal that is processed to extract the information bit. In this section, we will consider systems with ideal photodiodes which are endowed with a constant avalanche gain of unity.

A. System Description

Consider a binary communications system where the nth information symbol $b_n \in \{0, 1\}$ in the time interval $[(n-1)T, nT]$ is modulated onto the intensity of the incident light, giving rise to the following hypotheses testing problem at the receiver:

$$H_i : \text{photon intensity of } \lambda_i(t), \; i = 0, 1.$$  \hspace{1cm} (1)

Let $T$ be the sampling period and $[(n-1)T, nT]$ denote the nth sampling interval. If $R_{it}$ denotes the number of photons received at the photodetector during this interval, then the photoelectron count at output of the photodetector can be described statistically as

$$p_{R_{it}|H_i}(r) = \exp\left(-\int_{(n-1)T}^{nT} (\lambda_i(t) + \lambda_d)dt\right)$$

$$\left(\int_{(n-1)T}^{nT} (\lambda_i(t) + \lambda_d)dt\right)^r \frac{1}{r!} ; \; i = 0, 1,$$  \hspace{1cm} (2)

where $r \in \mathcal{Z} \equiv \{0, 1, 2, \cdots\}$ denotes the number of photons incident on the photodetector during the nth interval, and $\lambda_d$ accounts for the dark current which is assumed to be constant in time.

We will assume that the modulation used is on-off keying (OOK), however the following analysis can be easily duplicated for Manchester codes. We consider a homogeneous optical communication system employing $\lambda_0(t) = \lambda_0$ and $\lambda_1(t) = \lambda_1$. Without loss of generality, we consider the demodulation of the first bit which is in the interval $[0, T]$.

The probability mass functions that will govern decision making under each hypothesis are given by

$$p_{R_{it}|H_i}(r | H_i) = \frac{\exp\left(-\left(\lambda_i + \lambda_d\right)T\right)\left(\lambda_i + \lambda_d\right)^r}{r!} ; \; i = 0, 1,$$  \hspace{1cm} (3)

where $R_T$ is the photon count in $[0, T]$. Given equally likely hypotheses, it can be easily shown that the optimum maximum likelihood detector counts the number of photoelectrons $r$ in $[0, T]$ and compares it to the threshold $\gamma$ and decides as follows

$$\text{bit 0 was sent if } r < \gamma,$$

$$\text{bit 1 was sent if } r \geq \gamma,$$  \hspace{1cm} (4)

with

$$\gamma = \frac{(\lambda_1 - \lambda_0)T}{\log((\lambda_1 + \lambda_d)/(\lambda_0 + \lambda_d))}.$$  \hspace{1cm} (5)

where we have assumed, without loss of generality, that $\lambda_0 < \lambda_1$. The probability of error for the optimal detector can thus be computed as

$$P_e = \frac{1}{2} \sum_{\Delta_1} p_{R_{it}|H_0}(r) + \frac{1}{2} \sum_{\Delta_0} p_{R_{it}|H_1}(r).$$  \hspace{1cm} (6)

where $\Delta_1$ is the decision region for $H_1$.

In a Monte Carlo (MC) simulation, the probability of error is estimated as the ratio of the number of errors seen to the number of samples considered, that is

$$P_{MC} = \frac{1}{M_{MC}} \sum_{j=1}^{M_{MC}} I(r_j),$$  \hspace{1cm} (7)

where $I(r_j) = 1$ if the jth bit is in error and $I(r_j) = 0$ if the jth bit is correct, and $M_{MC}$ represents the total sample size. It is easily shown that this estimator is unbiased, i.e., $E[P_{MC}] = P_e$, and has variance given by $\text{var}(P_{MC}) = 

B. Importance Sampling Estimation

In Importance Sampling, a reduced variance estimate of $P_e$ is obtained by generating data under each hypothesis from the biased probability mass functions $p_{R_{it}|H_i}$, in place of $p_{R_{it}|H_i}$, which are chosen so as to generate more errors than expected. Then each error is weighted to obtain an unbiased estimate. We obtain the Importance Sampling estimator by rewriting the error probability in (6) as

$$P_e = \frac{1}{2} \left( \sum_{\Delta_1} p_{R_{it}|H_0}(r)w(r | H_0) + \sum_{\Delta_0} p_{R_{it}|H_1}(r | H_1)w(r | H_1) \right).$$  \hspace{1cm} (8)

The resulting sample estimate of the above expectation is given by

$$P_{e} = \frac{1}{M_{IS}} \left( \sum_{j=1}^{M_{IS}} \bar{w}(r_j)I(r_j) \right).$$  \hspace{1cm} (9)
where \( w(r) \) is the weight associated with the random sample \( r \) that is generated with equal probability from the biased densities \( p_{R \mid H_i} \), \( i = 0, 1 \). In the above estimator the weights under hypothesis \( H_i \) are given as

\[
w(r \mid H_i) = \frac{p_{R \mid H_i}(r)}{p_{R \mid H_i}(r)}, \quad \forall r \in \mathcal{Z},
\]

such that \( p_{R \mid H_i}(r) > 0 \). The variance of the above Importance Sampling estimate can be shown to be \( \text{var}[\hat{P}] = \frac{W - \hat{P}}{W} \), where \( W \) is defined as

\[
W = \frac{1}{2} \left[ \sum_{\Delta_i} \frac{p_{R \mid H_0}(r)}{p_{R \mid H_0}(r)} I_{\Delta_i}(r) + \sum_{\Delta_0} \frac{p_{R \mid H_1}(r)}{p_{R \mid H_1}(r)} I_{\Delta_0}(r) \right],
\]

and \( I_{\Delta_i}(r) \) is equal to 1, \( \forall r \in \Delta_i \), and is equal to 0, elsewhere.

It has been shown that (see for example [8]) the optimum biasing probability mass function that minimizes the variance of the above estimator is degenerate. Therefore, we must restrict ourselves to determining a suboptimal biasing density which yields a reduced variance estimate compared to conventional Monte Carlo techniques. Since variance reduction is not an easily quantifiable parameter, a useful expression is the “gain” and is given by

\[
\Gamma = \frac{M_{MC}}{M_{IS}} = \frac{P_0(1 - P_2)}{W - P_2},
\]

where the variance of both the estimators are set equal. Thus maximizing this ratio \( \Gamma \) would mean saving time through Importance Sampling as compared to the standard Monte Carlo simulations. The commonly used methods of arriving at a suboptimal density are the linear shift, increased variance, and exponential tilt [8], [20], [21]. We will use the exponential tilt [22], [21] method for our problem as it has certain characteristics which will minimize the variance of our Importance Sampling estimator.

**Definition:** An exponential change of probability measure from \( p_R(r) \) to \( p_H(r) \) is defined at \( r \) as

\[
p_H(r) = \frac{\exp(\lambda r)p_R(r)}{M_H(s)},
\]

where \( s \in \mathbb{R} \) is a free parameter and \( M_H(s) \) is the characteristic function of the measure \( p_H \).

Due to a theorem by Cottrell et al. [22], we have that for a Poisson process, the exponential change of measure from \( p_{R \mid H_0} \) (with parameter \( \lambda \)) to \( p_{R \mid H_i} \) (with parameter \( \lambda^* \)) has the asymptotic property that

\[
\sum_{\Delta_i} p_{R \mid H_i} \rightarrow 0, \quad i = 0, 1,
\]

is minimum as \( \lambda \rightarrow \lambda^* \), when \( p_{R \mid H_0} \) is as given in equation (13) with \( M_{R_0}(s) < \infty \) for \( s > 0 \). In this context, the intensity approaches its extreme value \( \lambda^* \) such that \( \sum_{\Delta_i} p_{R \mid H_i}(r) \rightarrow 0 \). Thus, a biasing density \( p_{R \mid H_i}(r | H_i) \) of the form given in equation (13) can asymptotically minimize the variance of the Importance Sampling estimator.

For the sake of simplicity, if we make our sampling interval such that \( T = 1 \), then under hypothesis \( H_i \), the exponential change of measure in (13) results in a biasing density which is again Poisson distributed with parameter \( \lambda_i + \lambda_4 \exp(s) \). Therefore, this exponential change of measure results in a new parametric class of probability measures with the intensity as the parameter. This closely resembles the linear shift of the mean that is a popularly used strategy for additive noise channels [9].

As a non-degenerate solution, we consider finding the best probability measure \( p_{R \mid H_i} \) from the above parametric class, thereby reducing the original infinite-dimensional optimization problem to a decoupled optimization problem in \( \mathcal{R}^2 \) as,

\[
\min_{\lambda_0^*, \lambda_4^*} \left\{ \frac{1}{2} \sum_{\Delta_i} \left[ \frac{p_{R \mid H_0}(r)}{p_{R \mid H_0}(r)} + \frac{p_{R \mid H_1}(r)}{p_{R \mid H_1}(r)} \right] \right\},
\]

where \( \lambda_0^* \) is the intensity parameter of \( p_{R \mid H_0}(r) \) for \( i = 0, 1 \). For each hypothesis \( H_i \), the above minimization can be solved numerically to find the optimum \( \lambda_i^* \). Instead, we will take advantage of the following sufficient conditions for achieving a “gain” of Importance Sampling simulations over Monte Carlo simulations, and arrive at a suboptimal \( \lambda_0^* \), which will be shown to yield a “gain” comparable to that obtained by using the optimal \( \lambda_0^* \). Clearly, from the expression for \( \Gamma \) we see that the following condition would yield a gain

\[
P_2 < \hat{W} \leq P_0,
\]

which for our binary system, via equations (9) and (11), implies that the tilted biased density will reduce the estimator variance if

\[
[\lambda_0^* - (\lambda_0 + \lambda_4)] + r \log[\lambda_0 + \lambda_4] < 0, \quad \forall r \in [\gamma, \infty).
\]

For each value of \( r \in [\gamma, \infty) \), we need to minimize the left side of the above inequality with respect to \( \lambda_0^* \), since this will yield a value of \( \hat{W} \) which is closest to it’s lower bound, \( P_2 \). Therefore the problem at hand is

\[
\min_{\lambda_0^*} \left\{ [\lambda_0^* - (\lambda_0 + \lambda_4)] + r \log[\lambda_0 + \lambda_4] \right\}, \quad \forall r \in [\gamma, \infty).
\]

Since the resulting rate will be a function of \( r \), we will minimize the maximum value of the above quantity as a function of \( r \), thus giving rise to the following minmax problem:

\[
\min_{\lambda_0^*} \max_{r \in \Delta_i} \left\{ [\lambda_0^* - (\lambda_0 + \lambda_4)] + r \log[\lambda_0 + \lambda_4] \right\}
\]

It is easily seen that the solution to the minmax problem is to find a \( \lambda_0^* \) that minimizes the left hand side of (17) for \( r = \gamma \). Therefore, the value of \( \lambda_0^* = \gamma \) is found to satisfy the inequality (17). This means that the “biased” equivalent arrival rate \( \lambda_0^* \) should be set to the arrival rate
corresponding to the threshold number of arrivals. A similar derivation can be performed for the case of hypothesis \( H_1 \), and it is seen that \( \lambda^* = \gamma \). Thus the biasing densities are identical under each hypothesis, and these constitute the suboptimal probability mass functions that are to be used for the Importance Sampling estimator. For small values of the error rate \( P_e \), the analytical evaluation of the gain \( \Gamma \) is intractable, and therefore upper and lower bounds for \( \Gamma \) are obtained in the Appendix. In the following, we also examine the tightness of these bounds.

C. Simulations and Numerical Results

The biasing densities were computed through both the numerical optimization procedure in (15) and the relatively easier suboptimum minmax rule which requires the intensities to be set to the threshold. As a comparison of the two schemes, the number of simulation runs required through both the suboptimum minmax rule and the numerical optimization procedure are shown in Fig. 2. The simulations were performed for different values \( \lambda_i \) (the intensity of modulation under \( H_i \)) keeping the dark current rate for the photodiode \( \lambda_d \) fixed at a value of unity. We observe that, for small values of the error probability, the number of Importance Sampling trials required under the two schemes are identical, and thus demonstrate the effectiveness of the suboptimum procedure. The lower and upper bounds developed for the simulated “gain” \( \Gamma \) are plotted in Fig. 3 along with the actual simulated “gain” for different values of \( \lambda_1 \). The lower bound is seen to closely follow the actual value of \( \Gamma \) and thus the simulated “gain” is a sufficiently accurate estimate of the gain of Importance Sampling over Monte Carlo simulations. The number of trials required for an estimate \( P_e \) within 0.1% of the actual probability of error \( P_e \) is computed for both Monte Carlo and Importance Sampling simulations and the resulting simulated “gain”, \( \Gamma = M_{MC}/M_{IS} \) is shown in Fig. 4 against the probability of bit-error on a logarithmic scale.

III. SYSTEMS WITH AVALANCHE PHOTODIODES

A. System Description

In an Avalanche photodiode/detector (APD), each primary electron-hole pair generates a random number of secondary pairs by a process of statistical multiplication. For avalanche detection systems employing direct detection, an “exact” method has been proposed [23], where the total probability of the photodetector count at the output of the avalanche photodetector in the time interval \( [0,t] \) can be described as

\[
P_{B_i}(r) = \sum_{m=0}^{\infty} p_{N_i}(m) p_{B_i | N_i}(r | m),
\]

where \( r \) is the photodetector count at the avalanche photodetector output, \( m \) is the number of photons incident on the avalanche photodetector at the input end, \( p_{N_i}(m) \) is the Poisson probability mass function of the photon count \( N_i \) with parameter \( \lambda_i(\sigma) \), \( 0 \leq \sigma < t \) and \( p_{B_i | N_i}(r | m) \) is the conditional probability that \( m \) incident photons give rise to \( r \) photoelectrons at the avalanche photodetector output. The complete expression for the above conditional probability for an avalanche photodetector with a mean avalanche gain \( G \) and ionization ratio \( \alpha \), has been found to be [24]

\[
p_{B_i | N_i}(r | m) = C \cdot \left[ \frac{m(1 - \alpha)^{-m} \Gamma(r \alpha)}{1 - \alpha} \right] \left[ \frac{1 + \alpha(G - \alpha)(r - m) - \alpha G}{G} \right]^{-m} \left[ (m + \alpha(r - m))(r - m) \right]^{1 - \alpha} \text{Gam}^{-1}
\]

where \( \Gamma(.) \) is the Gamma function and \( C \) is a constant such that \( \sum_{r=0}^{\infty} p_{B_i | N_i}(r | m) = 1 \).

We will consider a single-user binary hypothesis system where the intensity is modulated based on the information to be transmitted. The system is assumed to be homogeneous in the sense that the intensities under each hypothesis are kept constant, i.e., the intensity under \( H_i \) is assumed to be \( \lambda_i \). There is a constant dark current factor \( \lambda_d \) under each hypothesis, to account for the extraneous generation of photoelectrons in the photodetector. The probability mass function governing the statistics of the system under hypothesis \( H_i \) is given as

\[
p_{R_i | H_i}(r) = \sum_{m=0}^{\infty} \text{exp}(-\lambda_i - \lambda_d) \frac{(\lambda_i + \lambda_d)^m}{m!} p_{B_i | N_i}(r | m); \quad i = 0, 1,
\]

where \( p_{R_i | N_i}(r | m) \) is given in (19) and is assumed to be identical under both hypotheses and \( r \) is the photoelectron count in the time interval \( [0,1] \), where we have assumed, without loss of generality, that \( T = 1 \). The photoelectron count at the output of the avalanche photodetector is compared to a threshold and gives rise to the decision making process. Using the maximum likelihood decision rule, the threshold for detection is given as

\[
\gamma = \left[ \text{arg} \min \{ p_{R_i | H_i}(r) \} \right]_{0,1}.
\]

The problem at hand is the evaluation of the probability of error in equation (6) through the Importance Sampling technique of Monte Carlo simulations.

B. Importance Sampling Estimation

As before, we will formulate the probability of error in terms of the biased sample function density as

\[
P_e = \pi_0 \sum_{\Delta_1} p_{R_i | H_0}(r) w(r | H_0) + \pi_1 \sum_{\Delta_0} p_{R_i | H_1}(r) w(r | H_1),
\]

where the weights are given in (10) and \( p_{R_i | H_i}(r) \) denotes the biased sample function density under \( H_i \). As the optimum probability mass function yields a degenerative solution to the problem, we need to find a suboptimal probability mass function belonging to the parameter family.
$\{p_{\mathcal{R}^{2} | \mathcal{H}_i}(r); i = 0, 1\}$ with the parameter being the photon intensity, i.e., we look for a biasing density of the form

$$
p_{\mathcal{R}^{2} | \mathcal{H}_i}(r) = \sum_{m=0}^{\infty} \exp(-\lambda'_i) \frac{\lambda'_i^m}{m!} p_{\mathcal{R}^{2} | \mathcal{N}_T}(r | m),$$

(23)

where we have left the conditional probabilities unaltered. We will now present a lemma which is similar to the one due to Brady and Verdu [19] which will justify the selection of a biasing density in (23).

**Lemma 1:** Let $N_{i}(\lambda)$ be a Poisson random variable with mean $\lambda$, and let $R_{i}(\lambda) = \sum_{j=1}^{N_{i}} G_{j}$ where $\{G_{j}\}$ is a collection of nonnegative independent and identically distributed random variables that is independent of $N_{i}(\lambda)$. Let $0 < \lambda' < \lambda$. Then $P(R_{i}(\lambda) = r) \leq P(R_{i}(\lambda') = r)$, $\forall r \in \Delta_{0}$.

Considering hypothesis $H_{1}$, if we can find $p_{\mathcal{R}^{2} | \mathcal{H}_1}(r)$ with parameter $\lambda'_1$, such that $\lambda'_1 < \lambda_1$, then we have from the lemma that $P_{\mathcal{R}^{2} | \mathcal{H}_1}(r) < P_{\mathcal{R}^{2} | \mathcal{H}_1}(r)$, $\forall r \in \Delta_{0}$, and this is a sufficient condition for achieving a gain of Importance Sampling over Monte Carlo simulations. Therefore, it is sufficient to look for biasing probability mass functions $\{p_{\mathcal{R}^{2} | \mathcal{H}_1}(r)\}$ with only the photon intensity as the parameter and thus leave the conditional probability mass functions $P_{\mathcal{R}^{2} | \mathcal{N}_T}(r | m)$ unchanged. The analysis reduces to that considered in Section 2 and we can obtain the biased probability mass function as

$$
p_{\mathcal{R}^{2} | \mathcal{H}_1}(r) = \sum_{m=0}^{\infty} \exp(-\gamma) \frac{\gamma^m}{m!} p_{\mathcal{R}^{2} | \mathcal{N}_T}(r | m),$$

(24)

where $\gamma$ is as given in equation (5), i.e., the intensity under hypothesis $H_1$ is set to $\gamma$, the optimum threshold for the ideal photodetection case. A similar analysis for the case of hypothesis $H_0$ shows that the biased probability mass functions under each hypothesis are set identical and given by (24).

**C. Simulations and Numerical Results**

The biasing densities for the Importance Sampling estimator were taken as given in (24), where only the intensities are biased to new values, leaving the conditional probability mass functions unchanged. The simulations were performed for different values of the intensity of modulation $\lambda_1$, with $\lambda_2$ being held constant at unity. The mean avalanche gain, $G$, of the APD was set to a value of 50.0, with the ionization ratio being held at a value of $\alpha = 0.5$. The simulated "gain" $\Gamma$ is plotted for different values of the bit error probability in Fig. 4 where the number of trials for both Monte Carlo and Importance Sampling simulations are that required for an estimate $\hat{P}_e$ within 0.1% of the actual probability of error $P_e$. For both ideal and avalanche photodiode systems, the gain of Importance Sampling over Monte Carlo simulations is seen to increase with the performance of the system, i.e., as the probability of error for the system decreases, the value of the "gain" $\Gamma$ increases almost linearly as $1/P_e$. A comparison of the number of simulation trials required in each system is shown in Fig. 5. In either case, it is seen that less than 100 trials are required to estimate probabilities of the order of $10^{-16}$ in a 95% confidence interval of $[2P_e/5, 8P_e/5]$. Empirical analysis also supports the conjecture that biasing the gain distribution of the avalanche photodetector does not improve the "gain" of Importance Sampling.

**IV. Optical Code Division Multiple Access Systems**

In this section, we will develop an Importance Sampling scheme for a FO-CDMA system where an optical encoder maps each bit of information into a very high rate optical sequence, that is then coupled into a single-mode fiber channel. At the receiver end, the optical pulse sequence is compared to a stored replica of itself (correlation process) and compared to a threshold at the comparator for data recovery. However, the photodetectors used at the receiving end are endowed with Poisson counting statistics and in direct detection schemes the photocurrent count at the output of the photodetector is compared to a threshold for data recovery. We will develop an Importance Sampling technique that can be used for a wide class of codes, and in particular, illustrate the scheme for the case of optical orthogonal codes (OOC's) [16] and prime-sequences [18]. These optical codes conform to auto and crosscorrelation properties such as $\sum_{n=1}^{N} x_n x_{n+l} = J$ for $l = 0$, and $\sum_{n=1}^{N} x_n x_{n+l} = \rho_a$ for $l \neq 0$, and $\sum_{n=1}^{N} x_n y_{n+l} = \rho_c$ for $l \neq 0$, where $\{x_n\}$ and $\{y_n\}$ are periodic sequences with period $N$, and $J$, $\rho_a$, and $\rho_c$ are constants. In general, an $(N, J, \rho_a, \rho_c)$ optical code is a family of $[0, 1]$ sequences of length $N$ and weight $J$ with autocorrelation and crosscorrelation constraints $\rho_a$ and $\rho_c$. For OOC's $\rho_a = 1$, while for prime-sequences $\rho_c$ takes the value of 1 or 2 irrespective of the length of the codes. For prime-sequences derived from $GF[q]$ where $q$ is some prime number, $N = q^2$.

**A. System Description**

We will consider a $K$ user FO-CDMA system where the information bit of each user is modulated onto the intensity of the laser transmitted through a single-mode fiber channel. The intensities are in the form of optical pulse sequences and satisfy the auto and crosscorrelation properties set forth by codes assigned to each user. If user $k$ is sending bit $i$, under hypothesis $H_i$, then the intensity of the modulated light will be $\lambda^{(k)}_i(t)$ where

$$
\lambda^{(k)}_i(t) = \sum_{n=1}^{N} \lambda^{(1)}_n(t) \Pi_{T_{z}}(t - nT_{z}), \quad t \in [0, T]
$$

(25)

and $\Pi_{T_{z}}(t)$ is a unit rectangular pulse of duration $T_{z}$, and $\lambda^{(1)}_n(t) = [\lambda^{(1)}_1(t), \ldots, \lambda^{(1)}_N(t)]$ is a signature sequence of length $N = T/T_{z}$ with each $\lambda^{(1)}_n(t) \in [0, 1]$. At the receiving end, this gives rise to the following two hypotheses at the receiver of the desired user (taken to be user 1) in
the time interval \([0, T]\) as

\[
H_0 : \lambda_0^{(1)}(t) = \lambda_0^{(1)}(t) + \sum_{k=2}^{K} \lambda_k^{(k)}(t),
\]

\[
H_1 : \lambda_1^{(1)}(t) = \lambda_1^{(1)}(t) + \sum_{k=2}^{K} \lambda_k^{(k)}(t),
\]

(26)

where \(\lambda_0^{(1)}(t)\) is the sum of the intensities in the channel under hypothesis \(H_0\), due to the 1st user and \(K - 1\) interferers, and in \(\lambda_k^{(k)}(t)\) the symbol \(b\) denotes the information of the \(k^{th}\) user. The receiver corresponding to user 1 has a replica of the signature sequence assigned to user 1, and the light in the channel due to the user 1 and other \(K - 1\) interferers is correlated with this replicated signature sequence. This correlation process is carried out by means of a fiber-optic tapped-delay line, the details of which can be found in [17]. The correlated intensities are incident on an ideal photodiode and the photodetector count at the output of the photodiode is compared to a threshold for the purpose of data recovery, as shown in Fig. 1. Without loss of generality, we assume that each user is employing on-off keying, and hence at the 1st receiver, the intensities are correlated with \(\lambda_0^{(1)}(t)\), since \(\lambda_0^{(1)} = 0\). The statistical nature of the photodetector count at the output of the photodiode at the 1st receiver can be described, under \(H_1\), by a doubly stochastic Poisson process as

\[
P_{R_1[H_1]}(r) = \frac{\exp\left[-(\nu_0^{(1)} + \nu_4 T)[\nu_1^{(1)} + \nu_4 T]^r\right]}{r!},
\]

(27)

where \(r\) is the photodetector count in the time interval \([0, T]\), and \(\nu_1^{(1)}(t)\) is the random valued intensity resulting from correlating the replicated signature sequence at the 1st receiver with the intensity of the light transmitted through the channel, and \(\nu_4\) is a deterministic constant associated with the dark current of the photodiode. From the analysis presented in [17], it can be shown that \(\nu_1^{(1)} = \int_0^T \hat{\phi}_V(t)dt\), and for chip-synchronous codes (but bit-asynchronous) it can be written as \(\nu_1^{(1)} = (iJ + T^{(1)})T_0\), \(i = 0, 1\). The symbol \(T^{(1)}\) is the random valued interference term due to the presence of other users. Without loss of generality, we will consider user 1 with \(T_0 = 1\) (i.e., \(N = T\)), and present our analysis for the data recovery, at the receiver end of user 1, in the presence of interference from the other \(K - 1\) users. In the expression for the intensity of the correlated received optical signal, \(iJ\) is the desired intensity term and \(T^{(1)} = \sum_{k=2}^{K} T_k^{(1)}\) is the undesired intensity term due to the interferers. Each interference term \(T_k^{(1)}\), due to user \(k\), is a discrete random variable with \(T_k^{(1)} \in \{0, 1, \ldots, \rho_k\}\). For a direct detection optical system, it is essential to determine the counting statistics of the underlying process. The probability of \(r\) photoelectrons occurring, under \(H_1\), in the sampling interval \(T\) is given as

\[
P_{R_1[H_1]}(r) \left(r \left| \nu_1^{(1)}\right) = E[r_{R_1[H_1]}(r \left| \nu_1^{(1)})\right] = \frac{(\nu_0^{(1)} + \nu_4 T)[\nu_1^{(1)} + \nu_4 T]^r}{r!}, \quad i = 0, 1.
\]

(29)

where \(\nu_0^{(1)}(k)\) is the probability that the multiple access interference term \(T^{(1)}\) takes on the value \(k\). Using the maximum likelihood decision rule, the threshold \(\gamma\) for the detection process is given as

\[
\gamma = \left[\arg_{\nu_1^{(1)}} \frac{P_{R_1[H_1]}(r \left| \nu_1^{(1)})}{P_{R_1[H_0]}(r \left| \nu_1^{(1)})}\right) = 1\right]
\]

In general, finding the exact analytical probability of bit-error in (6) requires the computation of the probabilities of the multiple access interference which does not involve closed form expressions and therefore, there is a need to resort to simulations for performance evaluation.

B. Importance Sampling Estimation

In order to apply the existing theory for the exponential tilt class of biasing densities we must impose some restrictions on the statistics of the multiple access interference. To make the problem of determining the biasing density tractable, we will choose not to bias the multiple access interference parameters [8]. We will look for biasing densities of the form

\[
P_{R_1[H_1]}(r) = \sum_{k=0}^{\rho_1(K-1)} p_{T_0}(k)p_{R_1[H_1]}(r \left| \nu_1^{(1)})
\]

(30)

Since \(J\) and \(\nu_4\) are the only free parameters (i.e., not involving multiple access interference) in the probability mass functions governing the statistics of the system, we will look for a biased probability mass function \(p_{R_1[H_1]}\) with parameter \(\nu_1^{(1)}\), rather than \((J + \nu_4)\). Under hypothesis \(H_0\) the probability mass function \(p_{R_1[H_0]}\) is shifted to it’s biased version \(p_{R_1[H_1]}\), belonging to above parametric family with \(\nu_1^{(1)}\) as the new parameter, in place of \(\nu_4\). The new biasing densities are of the form

\[
P_{R_1[H_1]}(r) = \sum_{k=0}^{\rho_1(K-1)} p_{T_0}(k)\exp[-(\nu_1^{(1)} + k)]
\]

(31)

Therefore, under our stipulation that the multiple access interference be left unbiased, the conditions for gain through Importance Sampling over Monte Carlo simulations via equation (16), reduce to

\[
\sum_{k=0}^{\rho_1(K-1)} p_{T_0}(k)\exp[-(\nu_1^{(1)} + k)](\nu_1^{(1)} + \nu_4 + k)^r
\]

(32)
Since the probabilities $p_{T(1)}(k)$ are nonnegative for all $k$, we can reformulate the condition in equation (32) as

$$
p_{T(1)}(k) \frac{\exp{-[(J + \nu_0 N + k)(J + \nu_0 N + k)]}}{\exp{-[\nu_1^* + k]}} < 1,
\forall r \in \Delta_0,
$$

and

$$
p_{T(1)}(k) \frac{\exp{-[(\nu_0 N + k)(\nu_0 N + k)]}}{\exp{-[\nu_0^* + k]}} < 1,
\forall r \in \Delta_1,
$$

for each $k \in [0, \rho_e(K - 1)]$. Then, by our analysis in Section 2, the minimax problem for the case of hypothesis $H_1$ can be set up as

$$\min_{\nu_1^*} \max_{r \in \Delta_0} \left\{ -[(J + \nu_0 N + k) + \nu_1^* + k + r \log\left(\frac{J + \nu_0 N + k + r}{\nu_1^* + k + r}\right) \right\},$$

where $k \in [0, \rho_e(K - 1)]$. This results in the solution $\nu_1^* = \gamma - k$ and the new biasing density in equation (31) is given as

$$P_{T(1)}(k) = \sum_{k=0}^{\rho_e(K-1)} p_{T(1)}(k) \frac{\exp{-[(\gamma/\gamma)^r]}}{r!} = \frac{\exp{-[(\gamma/\gamma)^r]}}{r!},$$

since, $\sum_{k=0}^{\rho_e(K-1)} p_{T(1)}(k) = 1$. In effect, we have showed that the Importance Sampling simulations of the K-user FO-CDMA system reduce to just simulating a Poisson random variable with arrival rate set to the threshold (for the FO-CDMA system) and then weighting the errors appropriately. However, the weights for the Importance Sampling estimator will still involve the original density given in (28). A similar analysis of hypothesis $H_2$ shows the biased probability mass functions are identical under both hypotheses and these constitute the suboptimal probability mass functions that are to be used for Importance Sampling simulations for FO-CDMA systems. The above biasing scheme in (35) can also be used for the case of chip asynchronous codes if the interference distributions are known as well as for systems employing an APD.

C. Simulations and Numerical Results

We will consider both optimal orthogonal codes from a family of $\{N, J, 1, 1\}$ and prime sequences. For this case, the intensity term due to the interferers is given by $I^{(1)} = \sum_{k=2}^{N} I^{(1)}_k$, where each interference term $I^{(1)}_k$ is a random variable given by

$$I^{(1)}_k = \begin{cases} 0 & \text{with probability } a \\ 1 & \text{with probability } b \\ 2 & \text{with probability } c \end{cases}$$

For the case of a family of $\{N, J, 1, 1\}$ optical orthogonal codes, $a$, $b$, and $c$ take values $(1 - J^2/2N)$, $(J^2/2N)$ and 0 respectively. For prime sequences generated from $GF[11], GF[17]$ and $GF[31]$, it has been verified that the multiple access interference terms are as given in (36), with $a$, $b$, and $c$ being 0.57, 0.36 and 0.07, respectively [19]. For the FO-CDMA system employing optical orthogonal codes with $p_e$ as one, $I^{(1)}$ has a binomial distribution, while for a system employing prime sequences, $I^{(1)}$ takes on a trinomial distribution. For the FO-CDMA system, the optical orthogonal codes were designed such that the total code length was $N = 2000$, i.e., a family of $\{2000, J, 1, 1\}$ sequences. The weight $J$ for any $K$ corresponds to the maximum allowable weight for OOC's of length of $N$ and is governed by $K \leq \left\lceil \frac{N}{J^{2000-1}} \right\rceil$, where $\lceil z \rceil$ denotes the integer portion of the real value $z$. The simulated “gain” $\Gamma$ is shown in Fig. 6 and is found to be approximately proportional to $1/P_e$ with a gradual decline as the number of users in the system increased. The number of Importance Sampling trials are plotted for a K-user FO-CDMA system against the bit-error probability for the desired user in Fig. 7, and it is observed that less than 100 simulation trials were required for error probabilities of $10^{-10}$ and below. In Fig. 8, the logarithm of the quantity $C = \log(\Gamma \times P_e)$ is plotted against the number of users $K$ for a family of $\{1000, J^{2000}, 1, 1\}$ OOC’s, where $J^{2000}_e$ corresponds to the maximum allowable $J$ for any given number of users $K$, i.e., the maximum value of $J$ that satisfies the above inequality for $N = 1000$ and given value of $K$. For a given number of users $K$, the relationship between the simulated “gain” $\Gamma$ and the performance of the system tends to become linear, i.e., as the probability of bit-error for the desired user decreases there is an almost linearly related increase in the “gain” $\Gamma$. Due to the added interference from other users, the gain of Importance Sampling over Monte Carlo simulations is seen to decrease as the number of users $K$ increases, in the sense that $C = \Gamma \times P_e$ decreases as the number of users increases. To check the consistency of our simulation scheme, the variation of the probability of bit-error with changing weights $J$ is plotted in Fig. 9, and it is observed that the performance of the system is enhanced as the weights increase for a fixed code length, which is in accordance with earlier studies [17]. We use Importance Sampling simulations to compare the performance of a FO-CDMA system employing OOC’s with that employing prime-sequences for the case of identical code length. It is seen in Fig. 10 that for the case of $N = 961$, the system employing prime-sequences generated from $GF[31]$ performs better for the same number of users. Fig. 11 shows the additional capacity (in terms of the additional number of users that are supported for the same level of performance) that is achieved using prime-sequences as compared to OOC’s of the same length.

V. Conclusions

In this paper, we considered different cases of direct-detection optical communication systems and developed an Importance Sampling technique of Monte Carlo simulations for the performance analysis of such systems. In all cases considered, the “gain” of Importance Sampling over Monte Carlo simulations was found to be linearly increasing as the performance of the system, i.e., $\Gamma \approx 1/P_e$ for small values of $P_e$. It was also shown that Importance
Sampling simulation required as low as 50 – 100 trials for estimating error probabilities of the order of $10^{-10}$ and below. This quick and efficient method was used to analyze the performance of avalanche photodiode based systems as well as FO-CDMA systems, and we were able to estimate the system performance without resorting to any kind of approximations that are inherent in the existing analysis of such systems. For the multi-user case, the biasing density for Importance Sampling was shown to be just the Poisson probability mass function as opposed to the complex probability mixtures that govern the statistics of such systems. The above technique can be used for a wide variety of coding schemes, and was used to compare the relative performance of FO-CDMA systems employing OOC’s and prime-sequences. It was seen that performance of FO-CDMA systems with prime-sequences was superior to that of systems employing OOC’s.

APPENDIX
LOWER AND UPPER BOUNDS ON $\Gamma$

For a single user system employing an ideal photodiode, we have from (12) that the gain is given by $\Gamma \approx \frac{P_r}{W}$, where $P_r$ and $W$ are given in (6) and (11) respectively. For small error rates, we can assume equality and write

$$\Gamma \approx \frac{1}{2} \left[ \sum_{r=0}^{\infty} f_1(r) \right] + 1 \left( f_0(r) \right) \left[ \sum_{r=0}^{\infty} f_0(r) \right] \left[ \sum_{r=0}^{\infty} f_0(r) \right] \right]^{-1},$$

where $\gamma$ is given by (5), and the function $f_i(r)$ is given by $f_i(r) = \frac{\exp(-\gamma)}{r!}$. Since we have $(\lambda_0 + \lambda_d) \leq \gamma \leq (\lambda_1 + \lambda_d)$, the lower and upper bounds for $\Gamma$ are given as

$$\frac{1}{2} \left[ \sum_{r=0}^{\infty} f_1(r) \right] + 1 \left( f_0(r) \right) \left[ \sum_{r=0}^{\infty} f_0(r) \right] \left[ \sum_{r=0}^{\infty} f_0(r) \right] \right]^{-1} \leq \Gamma \leq \frac{1}{2} \left[ \sum_{r=0}^{\infty} f_1(r) \right] + \sum_{r=0}^{\infty} f_0(r),$$

where

$$\mu = \left( \sum_{r=0}^{\infty} f_0(r) \right) \left( \sum_{r=0}^{\infty} f_0(r) \right) \left( \sum_{r=0}^{\infty} f_0(r) \right) \right]^{-1} \left( \sum_{r=0}^{\infty} f_0(r) \right) \left( \sum_{r=0}^{\infty} f_0(r) \right) \right]^{-1} \left( \sum_{r=0}^{\infty} f_0(r) \right) \left( \sum_{r=0}^{\infty} f_0(r) \right) \right]^{-1},$$

The above lower and upper bounds for $\Gamma$ can be put in a concise form as

$$\left\{ \max \left[ \frac{\lambda_0 + \lambda_d}{\gamma} \right] \right\}^{-1} \exp(\gamma - \lambda_1 - \lambda_d),$$

where the function $Q(X^2 | \nu)$ is associated with a Chi-Square distribution and has the following description,

$$Q(X^2 | \nu) = \sqrt{2\pi}Z(X)\left[ 1 + \sum_{i=2}^{\infty} \frac{\gamma^{2i}}{2i!} \right],$$

where $Z(X) = \frac{1}{\sqrt{2\pi}} \exp(-X^2/2)$, and $\nu$ is even.

REFERENCES


Fig. 1. A K-user FO-CDMA system with the optical correlation receiver shown for the 1st receiver.


**Fig. 2.** Comparison of minmax rule and numerical optimization procedure.

The number of trials through Importance Sampling required for an estimate \( P_e \in [2 P_e/6, 8 P_e/5] \) with 95% confidence is shown using both numerical optimization and the minmax rule.

**Fig. 3.** Lower and Upper bounds for \( \Gamma \)

The simulated "Gain" \( \Gamma \) is plotted against the intensity \( \lambda_1 \) for a single user ideal photodiode system along with the lower and upper bounds on the "Gain". The dark current was set to a constant value of \( \lambda_d = 1.0 \).

**Fig. 4.** MC versus IS Simulations for PD and APD Systems

The simulated "Gain" \( \Gamma \) of Importance Sampling for a single user system with APD of mean avalanche gain of 50 and an ideal photodiode of constant avalanche gain of unity. The simulations were run till \( P_e \) was within 0.01% of \( P_e \).
Fig. 5. IS Simulations
The number of trials for Importance Sampling for an estimate $P_e \in [2P_e/5, 8P_e/5]$ with 95% confidence for a single user system with an ideal photodiode and an avalanche photodiode.

Fig. 8. Relative Variation of $\Gamma$ and $P_e$ for FO-CDMA
$log[C = \Gamma \times P_e]$ versus Number of users $K$ for a FO-CDMA system employing OOC's of length $N = 1000$, the number of weights $J$ for each user corresponds to the maximum allowable $J$.

Fig. 6. MC versus IS Simulations for K-user FO-CDMA
The simulated “Gain” $\Gamma$ is plotted against the Probability of bit-error of the desired user for a K-user FO-CDMA system employing OOC’s of length $N = 2000$.

Fig. 9. Performance versus Weights for K-user FO-CDMA
Probability of bit-error for the desired user versus the number of weights $J$ for $K = 2$, $K = 3$ and $K = 4$ FO-CDMA system employing OOC’s of length $N = 2000$.

Fig. 7. IS Simulations for K-user FO-CDMA
The number of trials for Importance Sampling for an estimate $P_e \in [2P_e/5, 8P_e/5]$ with 95% confidence for a K-user FO-CDMA system employing OOCs.

Fig. 10. Prime-Sequences vs. OOC’s
Importance Sampling simulations were conducted for 100 trials for OOC’s of length $N = 961$, i.e., from the family $\{961, J_{max}, 1, 1\}$, and prime-sequences generated from $GF[31]$. 
Number of Users that can be supported in the system for the same level of performance using OOC's of length \(N = 961\), i.e., from the family \(\{961, J_{max}, 1, 1\}\), and prime-sequences generated from \(GF[31]\). Importance Sampling simulations were conducted for 100 trials.

Narayan Mandayam was born in Madras, India, on October 18, 1967. He received the B.Tech (Hons.) degree in electrical engineering in 1989 from the Indian Institute of Technology, Kharagpur and the M.S. and Ph.D. degrees in electrical engineering from Rice University, Houston, TX, in 1991 and 1994, respectively. He is currently at WINLAB, Rutgers University, New Brunswick, NJ.

His research interests are in communication theory, radio and optical multiple access communications, discrete event dynamic systems, and wireless ATM.

Dr. Mandayam was a recipient of the Govt. of India National Talent Scholarship, the Institute Silver Medal, awarded to the best graduating student by the EE Department, Indian Institute of Technology, Kharagpur.

Bohnaam Aazhang was born in Bandar Anzali, Iran, on December 7, 1957. He received his B.S. (with highest honors), M.S., and Ph.D. degrees in Electrical and Computer Engineering from University of Illinois at Urbana-Champaign in 1981, 1983, and 1986, respectively.

From 1981 to 1985, he was a Research Assistant in the Coordinated Science Laboratory, University of Illinois. In August 1985, he joined the faculty of Rice University, Houston, Texas, where he is now an Associate Professor of Electrical and Computer Engineering. He was a Visiting Scientist at IBM Federal Systems Company, Houston, Texas, in the Summer of 1989. In the academic year 1991-92 and in the Summer of 1993, he was with the Institut für Kommunikationstechnik at Swiss Federal Institute of Technology (ETH) Zürich, Switzerland as a Visiting Professor. His research interests are in the areas of communication theory, information theory, and their applications with emphasis on radio and optical multiple access communications; discrete event simulations and perturbation analysis; and data driven nonlinear modeling.

Dr. Aazhang is a recipient of the Alosa Foundation Award 1993, the NSF Engineering Initiation Award 1987-1989, and the IBM Graduate Fellowship 1984-1985, and is a member of Tau Beta Pi andEta Kappa Nu. He is currently serving as the Editor for Spread Spectrum Networks of IEEE Transactions on Communications and as the Secretary of the Information Theory Society. He has also served as the Publications Chairman of the 1993 IEEE International Symposium on Information Theory, San Antonio, Texas.