Transactions Papers

A Simple Sequential Stopping Rule for Monte Carlo Simulation

Luis Mendo and José M. Hernando, *Member, IEEE*

***Abstract*—In this paper, a sequential stopping rule for the estimation of a probability p by means of Monte Carlo simulation is analyzed. It is shown that the proposed estimator is almost unbiased, and guarantees a given relative precision irrespective of p. Under very mild conditions, the method also guarantees a certain confidence level for a given relative estimation error, provided that p does not exceed a maximum value. An extension to importance sampling is discussed.**

***Index Terms*—Importance sampling (IS), Monte Carlo (MC) methods, sequential stopping rule, simulation.**

I. Introduction

Monte Carlo (MC) simulation is a powerful tool which enables the investigation of complex systems. MC techniques essentially aim at estimating an unknown parameter. In this paper we restrict to the case when the desired parameter is a probability, p.

The reliability of the estimation is usually described by its mean-square error (MSE) or by the confidence level associated with a given error interval. In conventional MC, no matter what quality measure is used, the required sample size for a given quality depends on the unknown p, with sample size increasing as p decreases. Importance sampling (IS) or other variance-reduction techniques can reduce the required sample size [1]; however, in general, its dependence on p cannot be avoided.

In practice, two methods can be used to determine sample size in MC simulations. In the first approach, which we will term *fixed-size MC* (FSMC), sample size is fixed beforehand, based on certain *a priori* knowledge. The second approach is to use a *sequential stopping* procedure, in which sample size is (randomly) determined by the outcome of the simulation, using knowledge gained from the simulation itself. A common rule is to let the simulation run until the empirical (i.e., estimated) relative precision (standard deviation divided by mean) reaches a prescribed value [2]. In [3], sufficient conditions are given for the asymptotic validity of confidence-based sequential stopping rules, as the desired confidence volume tends to zero. However, the nonasymptotic behavior of this and other classes of stopping rules is difficult to analyze.

In this paper, we study a simple sequential stopping rule for MC simulation, based on simulating until a given number of "important" events are observed. Although this kind of stopping rule is common practice in simulation [4], no rigorous analysis of its statistical properties is available, to the authors' knowledge. We will see that, for all practical values of p, the behavior of this method is similar (in terms of bias, MSE, and confidence level) to that exhibited by FSMC when the required sample size is assumed to be known in advance. The advantage of the proposed method is that, without assuming such knowledge, it can achieve a prescribed estimation quality.

The rest of the paper is organized as follows. Section II states the estimation method and analyzes its bias. Section III characterizes its MSE and relative precision, and Section IV analyzes confidence level as a function of error interval. Section V proposes an extension to IS and discusses its properties. Finally, Section VI gives conclusions and suggests future work.

II. Motivation and Statement of the Estimator. Bias

Let S be a system, and let H be an event defined in terms of the system output variables. We focus on the problem of estimating the probability p of the event H. To this end, we assume that a number of statistically *independent* simulations of the system are carried out. The independence assumption is standard in analyses of MC and IS techniques [5], [6], and is partly justified by analytical simplicity. Note that this assumption does not prevent application to systems with memory, as will be discussed in Section VI. In the following, we will refer to each independent system simulation as a *realization*. The number of realizations is the *sample size* of the simulation experiment. The outcome of the ith realization is a Bernoulli random variable[1] $\boldsymbol{h}_i$ which equals one if H occurs, and zero, otherwise. Each realization for which H occurs will be called a *successful* realization, or a *hit*.

In FSMC with sample size n, the number of successful realizations is a binomial random variable $\boldsymbol{N}$ with parameters n and

[1]Random variables are denoted in boldface throughout the paper, and vectors are underlined.

Paper approved by F. Santucci, the Editor for Transmission Systems of the IEEE Communications Society. Manuscript received November 18, 2003; revised February 11, 2005 and June 3, 2005.

The authors are with the Signals, Systems and Radiocommunications Department, Polytechnic University, 28040 Madrid, Spain (e-mail: lmendo@grc.ssr.upm.es; hernando@grc.ssr.upm.es).

Digital Object Identifier 10.1109/TCOMM.2005.863780

0090-6778/$20.00 © 2006 IEEE

p, and p is estimated as $\hat{\boldsymbol{p}} = \sum_{i=1}^{n} \boldsymbol{h}_i/n = \boldsymbol{N}/n$. This estimator is unbiased and has an MSE [1]

$$\mathrm{E}[(\hat{\boldsymbol{p}} - p)^2] = \frac{pq}{n} \tag{1}$$

where $q = 1 - p$. A standard quality measure for an estimator $\hat{\boldsymbol{p}}$ is its relative precision, defined as the square root of the MSE divided by p.[2] From (1), in order to obtain a given relative precision in FSMC, the sample size n should be inversely proportional to $p/q \approx p$.

A sequential stopping method is proposed in which realizations are run until a given number of hits N is reached. The sample size is thus a negative binomial random variable $\boldsymbol{n}$ [7, p. 96]:[3]

$$\Pr[\boldsymbol{n} = k] = \binom{k-1}{N-1} p^N q^{k-N}, \quad k \geq N \tag{2}$$

with $\mathrm{E}[\boldsymbol{n}] = N/p$ and $\mathrm{Var}[\boldsymbol{n}] = Nq/p^2$. It is shown in Appendix A that for $0 < p < 1$

$$\mathrm{E}\left[\frac{1}{\boldsymbol{n}}\right] = \sum_{i=1}^{N-1} \frac{(-1)^{i+1}}{N-i} \left(\frac{p}{q}\right)^i + (-1)^N \left(\frac{p}{q}\right)^N \ln p. \tag{3}$$

For low values of p, (3) implies that $\mathrm{E}[1/\boldsymbol{n}] \approx p/(N-1)$. This suggests the following estimator:

$$\hat{\boldsymbol{p}} = \frac{N-1}{\boldsymbol{n}}. \tag{4}$$

We will refer to this estimation method as *negative-binomial MC* (NBMC).

The normalized bias $b_{N,p}$ for the estimator (4) is obtained from (3) as

$$b_{N,p} = \frac{\mathrm{E}[\hat{\boldsymbol{p}}] - p}{p} = \sum_{i=1}^{N-1} (-1)^{i+1} \frac{N-1}{N-i} \frac{p^{i-1}}{q^i} + (-1)^N (N-1) \frac{p^{N-1}}{q^N} \ln p - 1. \tag{5}$$

It is also shown in Appendix A that for $N \geq 3$, $b_{N,p}$ is a negative, decreasing function of p, with

$$\max\left\{-\frac{1}{N}, -\frac{p}{N-2}\right\} \leq b_{N,p} < -\frac{p}{N-1+p} \tag{6}$$

and that $b_{N,p}$ increases with N. As a consequence of (6), $b_{N,p} \to 0$ for $p \to 0$ and $b_{N,p} \to -1/N$ for $p \to 1$. The normalized bias is plotted in Fig. 1 for several values of N. It can be seen that $|b_{N,p}|$ is very small; indeed, it is observed from (6) that, for $N \geq 3$, $|b_{N,p}| \leq p$. Therefore, $\hat{\boldsymbol{p}}$ can be considered essentially unbiased for practical purposes.

III. Mean-Square Error and Relative Precision

In NBMC with $N \geq 3$, it is shown in Appendix B that

$$\frac{\mathrm{E}[(\hat{\boldsymbol{p}} - p)^2]}{p^2} < \frac{1}{N-2} \tag{7}$$

$$\frac{\mathrm{E}[(\hat{\boldsymbol{p}} - p)^2]}{p^2} < \frac{1-p}{N-2}, \quad \text{for } p \leq \frac{1}{3}. \tag{8}$$

[2]For an unbiased estimator, this coincides with the coefficient of variation, $\sqrt{\mathrm{Var}[\hat{\boldsymbol{p}}]}/p$.

[3]The negative binomial distribution is sometimes defined as the number of unsuccessful realizations $\boldsymbol{n} - N$.

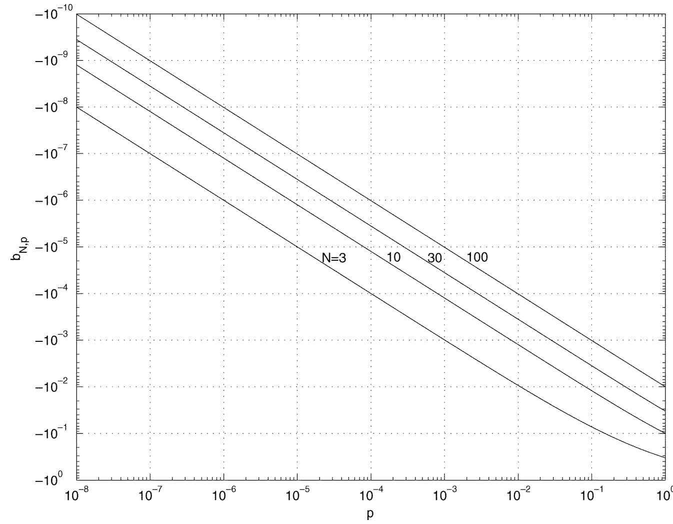


Fig. 1. Normalized bias in NBMC.

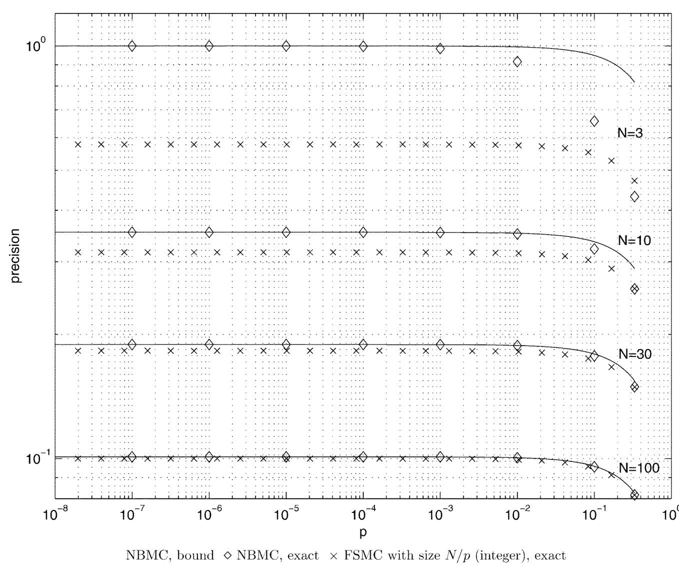


Fig. 2. Relative precision in NBMC and in FSMC.

According to (7), the relative precision $\sqrt{\mathrm{E}[(\hat{\boldsymbol{p}} - p)^2]}/p$ is better than $1/\sqrt{N-2}$ for all values of p. Assuming that the average sample size N/p is an integer, we can make a straightforward comparison with FSMC with that sample size. From (1) and (8), it is seen that for a given average sample size, and assuming $p \leq 1/3$, the relative precision in NBMC is degraded no more than a factor $\sqrt{N/(N-2)}$ with respect to that in FSMC. In general, for N/p not necessarily integer, considering a sample size n' in FSMC

$$\frac{\mathrm{E}[(\hat{\boldsymbol{p}} - p)^2]|_{\mathrm{NBMC}}}{\mathrm{E}[(\hat{\boldsymbol{p}} - p)^2]|_{\mathrm{FSMC}}} = \frac{N}{N-2} \frac{n'p}{N} \tag{9}$$

for $N \geq 3$ and $p \leq 1/3$. Thus, the degradation factor for the relative precision is $\sqrt{N/(N-2)}$ divided by the square root of the ratio of average sample sizes.

Fig. 2 shows the bound in (8), exact values obtained by numerical computation, and exact values in FSMC with the same (integer) average size. It is seen that the relative precision in NBMC is very similar to that in FSMC. It can also be observed that the bound (8) is a very good approximation for the exact MSE in NBMC, except when p is high and N is very small.

In view of (7), NBMC can be used to assure a relative precision better than a prescribed value, despite p being unknown. In fact, it is seen that relative precision is nearly independent of p.

IV. Confidence Level for an Error Interval

The confidence c for an interval $[p/\mu_2, p\mu_1]$; $\mu_2, \mu_1 > 1$, is defined as the probability that the estimated value lies in that interval, and can be computed in NBMC as $c = P_2(p) - P_1(p)$, with $P_2(p) = \Pr[\hat{\boldsymbol{p}} \geq p/\mu_2] = \Pr[\boldsymbol{n} \leq n_2]$, $P_1(p) = \Pr[\hat{\boldsymbol{p}} > p\mu_1] = \Pr[\boldsymbol{n} \leq n_1 - 1]$, and

$$n_2 = \left\lfloor \frac{(N-1)\mu_2}{p} \right\rfloor \quad (10)$$

$$n_1 = \left\lceil \frac{N-1}{p\mu_1} \right\rceil \quad (11)$$

where $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$, respectively, denote rounding to the nearest integer toward $-\infty$ and toward ∞.

Let $\bar{c} = \gamma(N, (N-1)\mu_2) - \gamma(N, (N-1)/\mu_1)$, where $\gamma(r, x)$ is the incomplete gamma function, $\gamma(r, x) = (1/\Gamma(r)) \int_0^x e^{-t} t^{r-1} dt$. The following results, given by (12)–(15), hold for $N \geq 3$, as shown in Appendix C:

$$\lim_{p \to 0} c = \bar{c} \quad (12)$$

$$|c - \bar{c}| < \frac{p}{2}\sqrt{\frac{\left(\frac{N-1}{N-1-p}\right)^2}{1 - \frac{(N-1)p}{N-1-p}} + \frac{2}{N-1}} + \frac{p}{2}\sqrt{\frac{\left(\frac{N}{N-1}\right)^2}{1 - \frac{Np}{N-1}} + \frac{2}{Np}} \quad \text{for } p < \frac{N-1}{N\mu_1} \quad (13)$$

$$|c - \gamma(N, (N-1)\mu_2)| < \frac{p}{2}\sqrt{\frac{\left(\frac{N-1}{N-1-p}\right)^2}{1 - \frac{(N-1)p}{N-1-p}} + \frac{2}{N-1}} \quad \text{for } p \geq \frac{N-1}{N\mu_1} \quad (14)$$

$$c > \bar{c} \quad \text{for } \mu_2 \geq \frac{N+\sqrt{N}}{N-1}, \mu_1 \geq \frac{N-1}{N - \sqrt{\frac{3}{2}N}}, \quad p < \frac{N-1}{\left\lceil \frac{7}{2}N - 1 \right\rceil \mu_1}. \quad (15)$$

Expressions (12) and (13) characterize the convergence of c to an asymptotic value $\bar{c}$ as $p \to 0$. Equation (15) establishes that, if the confidence interval is wide enough and p does not exceed a given value, the confidence is guaranteed to be greater than its asymptotic value. It should be remarked that the conditions expressed in (13)–(15) are not necessary. In fact, numerical computation shows that $c > \bar{c}$ for some combinations of parameters outside the scope of the conditions in (15).

We now particularize to intervals of the form $[p/(1+m), p(1+m)]$, where $m > 0$ is a relative error margin, and compare the confidence obtained in NBMC with that in FSMC. In FSMC with sample size n', the confidence level for a given m is

$$c = \sum_{i=\lceil n'p/(1+m) \rceil}^{\lfloor n'p(1+m) \rfloor} \binom{n'}{i} p^i q^{n'-i}. \quad (16)$$

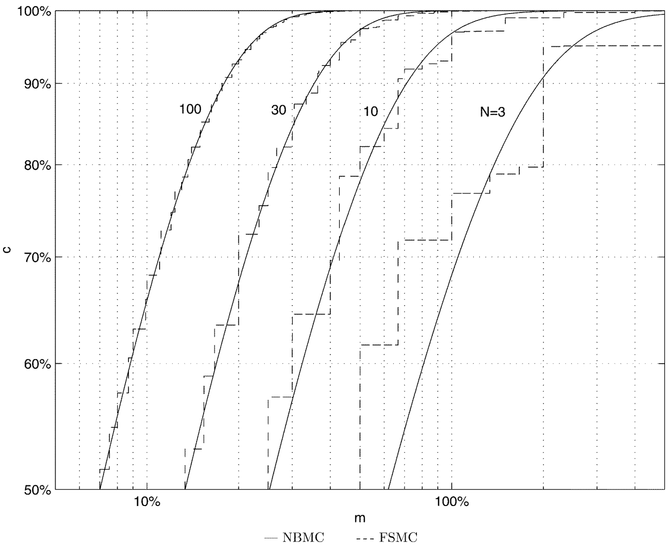


Fig. 3. Confidence as a function of error margin in NBMC and in FSMC, for $p = 10^{-3}$.

To obtain a meaningful comparison, we choose n' equal to the average sample size in NBMC, N/p, for N/p integer. Fig. 3 depicts c as a function of m in NBMC and in FSMC, for several values of N and for $p = 10^{-3}$ (note that in all cases shown N/p is an integer). Results for other values of p are very similar. The curves have jumps, caused by the discrete character of $\hat{\boldsymbol{p}}$. These jumps are much more evident in FSMC, due to the smaller number of possible values of $\hat{\boldsymbol{p}}$ with that method. It can be seen that the confidence level in NBMC is similar to that in FSMC with the same error margin and average (or deterministic, in FSMC) number of realizations.

Fig. 4(a) shows, with dashed lines, the guaranteed confidence $\bar{c}$ as a function of m, with N as a parameter. This confidence level is guaranteed only for p not exceeding the value represented in Fig. 4(b). The dashed curves in Fig. 4(a) are the limit of those in Fig. 3 as $p \to 0$, for the range of m satisfying the conditions in (15) with $\mu_1 = \mu_2 = 1 + m$. Fig. 4(a) also shows, with a solid line, the minimum confidence level $\bar{c}_{\min}$ that can be guaranteed (by proper selection of N) as a function of m. This is given as $\min_{N \geq 3} \bar{c}$ subject to the conditions expressed in (15). The resulting curve delimits the range of confidence levels that can be guaranteed. For a given error margin m, any confidence value above this curve can be assured using adequate N, for p below the maximum value corresponding to that N. The figure shows that the sufficient conditions (15) cover all cases of practical interest. Namely, any confidence greater than 85% can be assured for any margin lower than 120%; and any confidence greater than 80% can be assured for any margin lower than 67%.

The selection of N as a function of the desired confidence-error performance is depicted in Fig. 5. Fig. 5(a) shows the minimum value of N that assures a given confidence level for a given relative error m, provided that p does not exceed the value in Fig. 5(b). These can be used as "design curves" in NBMC simulations. As an example, if we wish to guarantee that $\hat{\boldsymbol{p}}$ lies in the interval $(0.5p, 2p)$ ($m = 100\%$) with 95% confidence, $N = 9$ should be used, with a maximum p of 0.129 (this perfor-

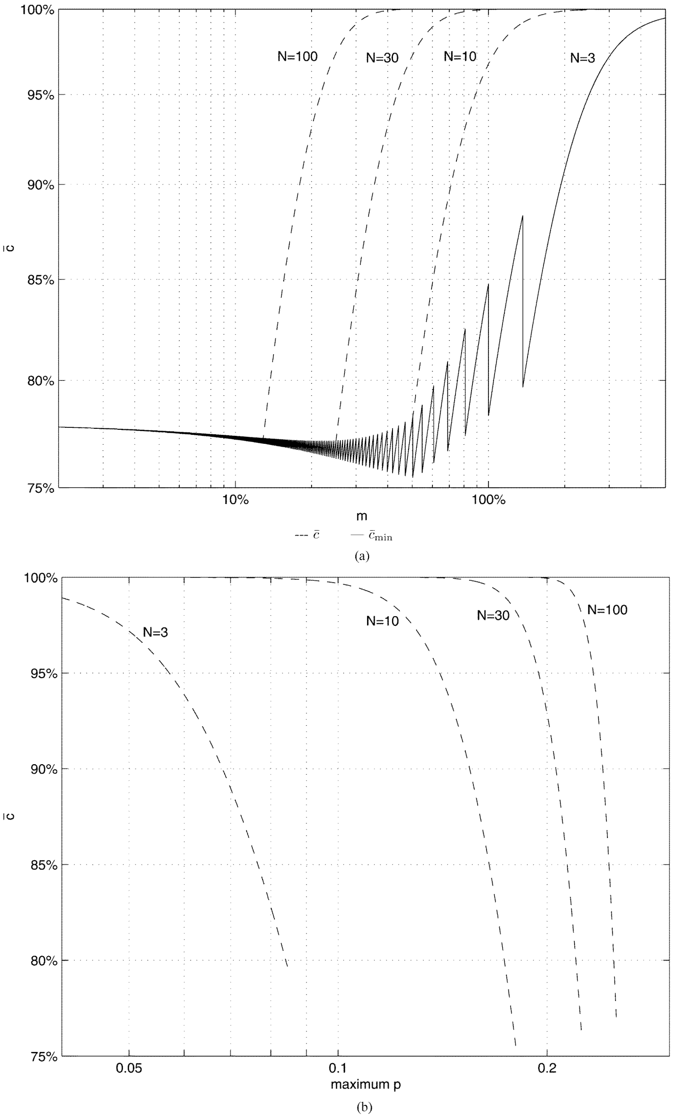


Fig. 4. Guaranteed confidence in NBMC. (a) Guaranteed confidence $\bar{c}$ and minimum confidence that can be guaranteed $\bar{c}_{\min}$. (b) Maximum p for guaranteed confidence $\bar{c}$.

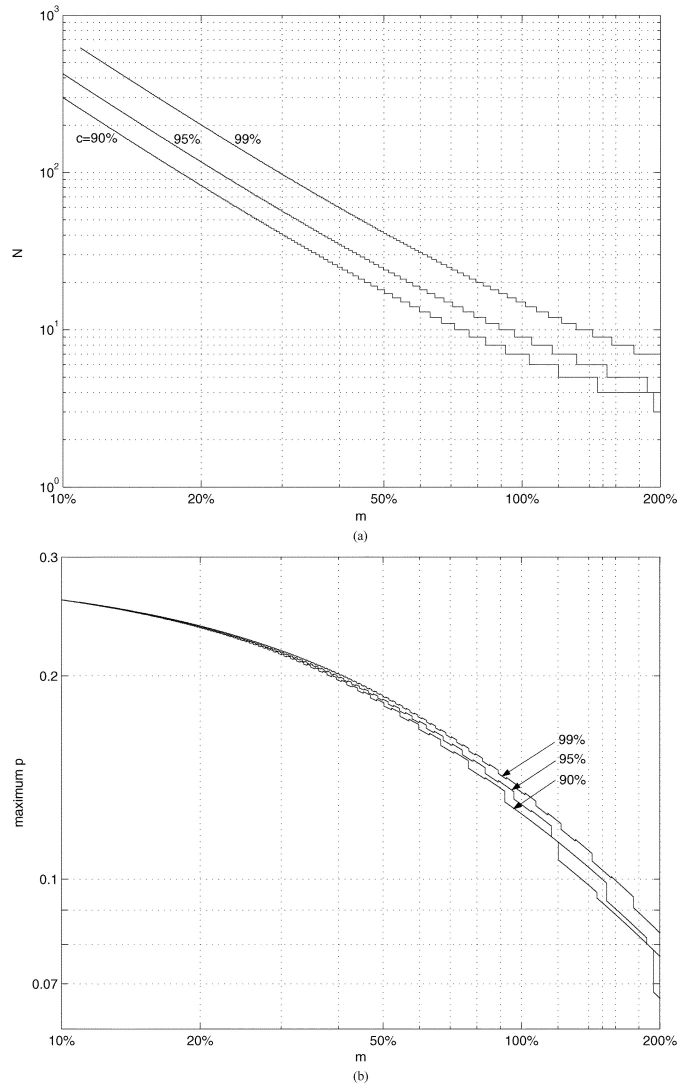


Fig. 5. N and maximum p for guaranteed confidence level in NBMC. (a) Minimum N that guarantees confidence level. (b) Maximum p for guaranteed confidence level.

mance is quite similar to the "rule of thumb" for FSMC stated in [5, p. 157]).

V. An Extension to Importance Sampling

In this section, we discuss a straightforward generalization of NBMC to IS that preserves its bias properties. Unfortunately, this generalization lacks the feature of guaranteed relative precision, due to the variability introduced by the weighting in IS. However, we will see that this degradation is approximately the same as that in the fixed-size case.

Let $\underline{\boldsymbol{x}}_i$ be a vector containing the random variables that constitute the input of the system at the ith realization, with joint probability density function (pdf) $f_{\underline{\boldsymbol{x}}}(\underline{x})$ (independent of i). With IS, the pdf of $\underline{\boldsymbol{x}}$ is changed, or "biased," to $f^*_{\underline{\boldsymbol{x}}}(\underline{x})$. In *fixed-size IS* (FSIS) with sample size n, p is estimated as $\hat{\boldsymbol{p}} = \sum_{i\in\boldsymbol{\mathcal{H}}} \boldsymbol{w}_i/n$, where $\boldsymbol{\mathcal{H}}$ denotes the set of successful realizations, and $\boldsymbol{w}_i = f_{\underline{\boldsymbol{x}}}(\underline{x}_i)/f^*_{\underline{\boldsymbol{x}}}(\underline{x}_i)$ is the weight that must be applied to each hit so as to unbias the estimation [8].

The stopping rule for MC introduced in Section II can directly be applied to IS. $\boldsymbol{n}$ realizations are carried out to obtain N hits, and p is estimated as

$$\hat{\boldsymbol{p}} = \frac{N-1}{\boldsymbol{n}} \frac{\sum_{i\in\boldsymbol{\mathcal{H}}} \boldsymbol{w}_i}{N}. \tag{17}$$

We will term this method *negative-binomial IS* (NBIS).

The two factors in (17) are statistically independent, because the weight associated with a successful realization depends only on the input variables of that realization. Due to the IS biasing, the simulation actually generates hits with a probability p^* different (usually larger) than p. The first factor in (17) can be seen as an NBMC estimator of p^*. Defining $\mathcal{R}$ as the region of input values that produce a hit, the mean of the second factor can be computed as

$$\mathrm{E}\left[\frac{\sum_{i\in\boldsymbol{\mathcal{H}}} \boldsymbol{w}_i}{N}\right] = \mathrm{E}[\boldsymbol{w}_i | i \in \boldsymbol{\mathcal{H}}] = \frac{\int_{\mathcal{R}} \frac{f_{\underline{\boldsymbol{x}}}(\underline{x})}{f^*_{\underline{\boldsymbol{x}}}(\underline{x})} f^*_{\underline{\boldsymbol{x}}}(\underline{x}) d\underline{x}}{\int_{\mathcal{R}} f^*_{\underline{\boldsymbol{x}}}(\underline{x}) d\underline{x}} = \frac{p}{p^*}. \tag{18}$$

This implies that the normalized bias in NBIS coincides with that of an NBMC estimator of p^*. The relative precision, however, is degraded with respect to NBMC, because of the variability introduced by the second factor.

The squared relative precision in NBIS is computed as follows. Let ε and μ denote, respectively, the MSE and the mean-square value of the associated NBMC estimator of p^*. Let σ_{w}^2 denote the conditional variance of the weight in a successful realization, and $c_{\mathrm{w}} = \sigma_{\mathrm{w}} p^*/p$ denote the coefficient of variation of the conditional weight. Taking into account that

$$\mathrm{E}\left[\left(\frac{\sum_{i \in \mathcal{H}} \boldsymbol{w}_i}{N}\right)^2\right] = \frac{p^2}{p^{*2}} + \frac{\sigma_{\mathrm{w}}^2}{N} \tag{19}$$

the following expression is obtained:

$$\begin{aligned}\frac{\mathrm{E}[(\hat{\boldsymbol{p}} - p)^2]}{p^2} &= \mu\left(\frac{1}{p^{*2}} + \frac{\sigma_{\mathrm{w}}^2}{Np^2}\right) - \frac{2\mathrm{E}[\hat{\boldsymbol{p}}]}{p^*} + 1 \\ &= \frac{\varepsilon}{p^{*2}} + \frac{\mu\sigma_{\mathrm{w}}^2}{Np^2} = \frac{\varepsilon}{p^{*2}}\left(1 + \frac{\mu c_{\mathrm{w}}^2}{N\varepsilon}\right).\end{aligned} \tag{20}$$

According to (20), the relative precision in NBIS is expressed as that in the underlying NBMC process, $\sqrt{\varepsilon}/p^*$, augmented by a factor that depends on the conditional weight distribution only through c_{w}. The referred factor indicates, for a given p and p^*, how efficiently we are sampling the region $\mathcal{R}$. In view of this, we can define the *efficiency factor* ν of an IS technique (whether NBIS or FSIS) as the relative precision in the underlying MC process, divided by the actual relative precision with IS. In the following, we compare the efficiency factor in NBIS and in FSIS.

As seen in Section III, $\varepsilon/p^{*2} < 1/(N-2)$, with $\varepsilon/p^{*2} \approx 1/(N-2)$ for all p of practical interest, and thus $p^{*2}/(N\varepsilon) \approx (N-2)/N$. With this approximation, (20) is transformed into

$$\frac{\mathrm{E}[(\hat{\boldsymbol{p}} - p)^2]}{p^2} \approx \frac{\varepsilon}{p^{*2}}\left(1 + \frac{N-2}{N}\frac{\mu c_{\mathrm{w}}^2}{p^{*2}}\right). \tag{21}$$

Since NBMC is approximately unbiased, $\mu/p^{*2} \approx (\varepsilon + p^{*2})/p^{*2} \approx 1 + 1/(N-2)$, and, therefore

$$\frac{\mathrm{E}[(\hat{\boldsymbol{p}} - p)^2]}{p^2} \approx \frac{\varepsilon}{p^{*2}}\left(1 + c_{\mathrm{w}}^2\right). \tag{22}$$

On the other hand, in FSIS with a given sample size n'

$$\begin{aligned}\frac{\mathrm{E}[(\hat{\boldsymbol{p}} - p)^2]}{p^2} &= \frac{\mathrm{Var}[\boldsymbol{w}_i \boldsymbol{h}_i]}{n'p^2} = \frac{\left(\sigma_{\mathrm{w}}^2 + \left(\frac{p}{p^*}\right)^2\right)p^* - p^2}{n'p^2} \\ &= \frac{\varepsilon'}{p^{*2}}\left(1 + \frac{c_{\mathrm{w}}^2}{q^*}\right) \approx \frac{\varepsilon'}{p^{*2}}\left(1 + c_{\mathrm{w}}^2\right)\end{aligned} \tag{23}$$

where $q^* = 1 - p^*$ and $\varepsilon' = p^* q^*/n'$ is the MSE in FSMC with probability p^* [see (1)]. It is observed that the efficiency factor in FSIS and in NBIS is approximately the same, namely $\nu \approx 1/\sqrt{1 + c_{\mathrm{w}}^2}$. In both estimation methods, the best input distributions, in the sense of the efficiency factor, are proportional to the original distribution across $\mathcal{R}$, so that $c_{\mathrm{w}} = 0$; any such input distribution produces a relative precision equal to that in the underlying MC process.

Clearly, the efficiency factor alone does not characterize the relative precision of the IS estimation method. In FSIS, for ν and p given, relative precision can always be improved by increasing p^* (i.e., concentrating the input distribution on $\mathcal{R}$), as stems from (23) and (1). In NBIS, for ν and p given, increasing p^* is also beneficial, since it reduces simulation time while maintaining a given relative precision. Consequently, the techniques that have been developed for the design of good biasing (i.e., for achieving small c_{w} and large p^*) in FSIS [1] are fully applicable to NBIS.

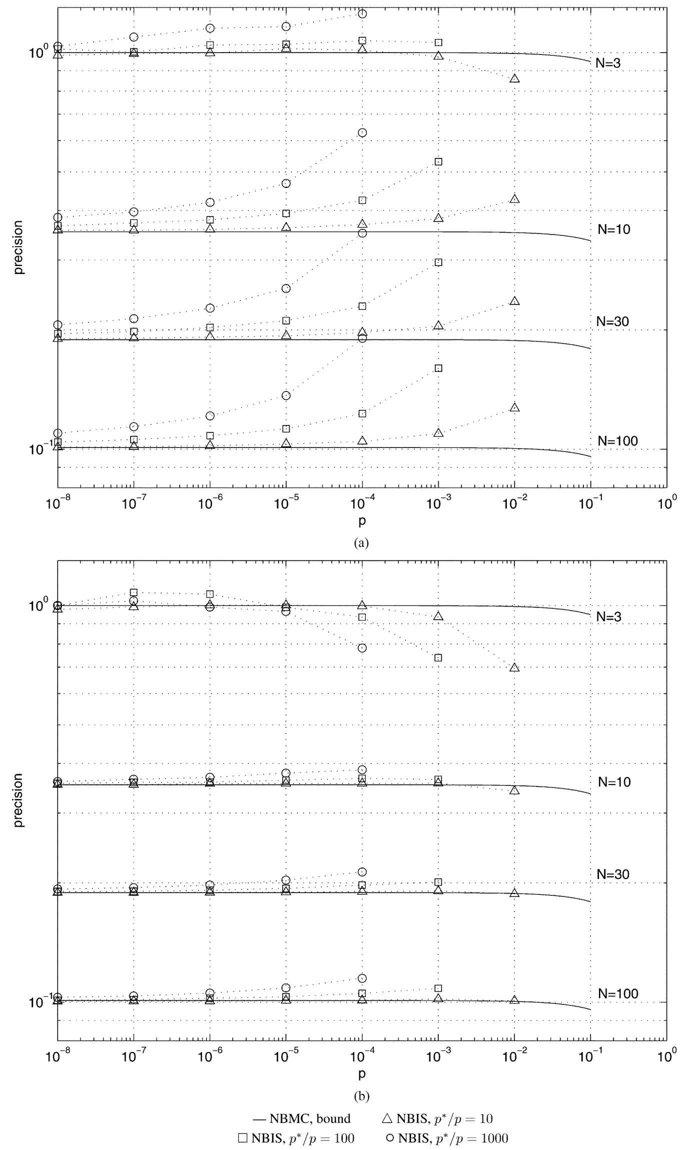


Fig. 6. Relative precision in NBIS example. (a) Variance scaling. (b) Mean translation.

In order to assess the degradation in NBIS, we have simulated two simplified estimation experiments, in which H is the event that a normalized Gaussian random variable exceeds a given threshold, and the input distribution is biased using variance scaling or mean translation, respectively. In either case, the simulation consists of a number M of trials of the NBIS experiment (each with a number of realizations determined by the stopping rule). The jth trial yields a value $\hat{\boldsymbol{p}}_j$ of the estimator, and the MSE is *estimated* as $\sum_{j=1}^{M}(\hat{\boldsymbol{p}}_j - p)^2/M$. M is chosen as 10^6.

Fig. 6 shows the resulting relative precision for several values of N with p^*/p as a parameter, as well as the bound (8) for NBMC. Note that in our examples, p determines the threshold, and p^*/p is related to the scaled variance or the translated mean, with $p^*/p = 1$ corresponding to no biasing, i.e., NBMC. Table I

TABLE I
c_{w} IN NBIS EXAMPLES

	Variance scaling			Mean traslation		
	$p^*/p=10$	$p^*/p=100$	$p^*/p=1000$	$p^*/p=10$	$p^*/p=100$	$p^*/p=1000$
$p=10^{-2}$	0.839	—	—	0.339	—	—
$p=10^{-3}$	0.445	1.273	—	0.202	0.519	—
$p=10^{-4}$	0.308	0.713	1.605	0.145	0.332	0.644
$p=10^{-5}$	0.236	0.513	0.919	0.113	0.248	0.429
$p=10^{-6}$	0.192	0.404	0.673	0.092	0.198	0.328
$p=10^{-7}$	0.161	0.335	0.539	0.078	0.166	0.267
$p=10^{-8}$	0.139	0.286	0.452	0.068	0.142	0.226

shows the computed c_{w}. The relative precision is similar to that of the associated NBMC when mean translation is used, because c_{w} is low, and is larger for variance scaling, with higher values of c_{w}. In both cases, for N not too small ($N \geq 10$), the relative precision deviates more from NBMC as p^*/p is increased. This is because a greater p^*/p implies, in our examples, a larger c_{w}. For N small ($N = 3$) this behavior changes, because the approximations leading to (22) are less accurate. Compared with FSMC with sample size N/p (see Fig. 2), NBIS shows some degradation, which is larger for variance scaling. The reduced MSE in mean translation compared with variance scaling is in accordance with general behavior in FSIS [2].

VI. Conclusions and Further Work

A stopping rule has been proposed for MC simulation which estimates a probability p without any *a priori* information, in particular without knowledge of the required sample size. The rule consists in simulating as many realizations as necessary to obtain N hits. The resulting method, NBMC, is essentially unbiased, and its relative precision is bounded irrespective of p. Under very mild conditions, the estimator also guarantees a certain confidence level for a given error margin. The quality of the estimation can thus be fixed *a priori* by properly selecting N. NBMC achieves similar performance (either in terms of relative precision or confidence for an error margin) to that obtained in FSMC with perfect knowledge of the necessary sample size. The sample size in NBMC is random, with mean N/p and coefficient of variation bounded by $1/\sqrt{N}$, for a single simulation; in a series of simulations, this variability is reduced by the law of large numbers.

The estimation method is applicable for simulations composed of independent realizations. An example is static ("snapshot") simulations, as used in system-level analyses of mobile cellular networks. In particular, NBMC has been successfully used for estimating the outage probability in code-division multiple-access systems [9]. When statistical dependence (system memory) needs to be modeled, the technique of event simulation, introduced in [10], can be exploited to attain independent realizations.

A generalization to IS has been proposed. NBIS retains good bias properties, but its relative precision is not guaranteed. The degradation of relative precision, compared with NBMC is approximately the same as in the fixed-size case, as are the design criteria for the IS biasing.

Further work is required to relax the sufficient conditions stated in (8) and in (13)–(15). An interesting line of research is also the analysis of other generalizations of NBMC to IS. An IS method that guarantees precision or confidence level would be highly valuable. Another topic for future research is the possibility to incorporate adaptive IS techniques [11] into this framework.

Appendix A
Computation of $\mathrm{E}[1/\boldsymbol{n}]$ and Bias in NBMC

Let $f_N[k]$ denote $\Pr[\boldsymbol{n} = k]$ for a given N, let $\Phi_N(z)$ be its z-transform

$$\Phi_N(z) = \sum_{k=N}^{\infty} f_N[k] z^{-k} = \frac{p^N}{(z-q)^N}, \quad |z| > q \tag{24}$$

and let η_N denote $\mathrm{E}[1/\boldsymbol{n}] = \sum_{k=N}^{\infty} f_N[k]/k$. Dividing (24) by z and integrating from 1 to ∞ yields

$$\eta_N = \sum_{k=N}^{\infty} \frac{1}{k} f_N[k] = \int_1^{\infty} \frac{p^N}{z(z-q)^N} dz. \tag{25}$$

This integral is computed carrying out a partial fraction expansion of the integrand, and (3) results.

We point out that for p close to 1, (3) contains large terms that partially cancel, making its numerical evaluation difficult. In this case, a more convenient expression can be derived, noting that $\ln p = -\ln(1 + q/p)$ and using a Taylor expansion of $\ln(1 + 1/t)$, $t \geq 1$ with $t = p/q$ in (3).

The lower bound in (6) will hold if and only if $b_{N,p} \geq -1/N$ and $b_{N,p} \geq -p/(N-2)$. To prove the first inequality, we note that for any positive random variable $\boldsymbol{x}$

$$\mathrm{E}\left[\frac{1}{\boldsymbol{x}}\right] \geq \frac{1}{\mathrm{E}[\boldsymbol{x}]}. \tag{26}$$

Applying (26) to the variable $\boldsymbol{n}$ and taking into account that $\mathrm{E}[\boldsymbol{n}] = N/p$, the desired result is obtained. To establish the second inequality, we introduce the following notation: $k^{(i)} = k(k-1)\cdots(k-i+1)$, $k^{(0)} = 1$. Since $\sum_{k=N}^{\infty} \Pr[\boldsymbol{n} = k] = 1$, from (2), it stems that

$$\sum_{k=N}^{\infty} (k-1)^{(N-1)} q^{k-N} = \frac{(N-1)!}{p^N}. \tag{27}$$

Using (2) and (27) we express η_N for $N \geq 3$ as

$$\begin{aligned}\eta_N &= \frac{p^N}{(N-1)!}\sum_{k=N}^{\infty}\left(1-\frac{1}{k}\right)(k-2)^{(N-2)}q^{k-N} \\ &\geq \frac{p^N}{(N-1)!}\left(\sum_{k=N}^{\infty}(k-2)^{(N-2)}q^{k-N}\right. \\ &\quad \left. - \sum_{k=N}^{\infty}(k-3)^{(N-3)}q^{k-N}\right) \\ &= \frac{p}{N-1}\left(1-\frac{p}{N-2}\right). \end{aligned} \tag{28}$$

The upper bound in (6) is equivalent to $\eta_N < p/(N-q)$, and is established as follows. We observe that for $t > 0$ and $0 < p < 1$, the inequality $1 + tp > (1+t)^p$ holds. With the substitution $t = (z-1)/p$, the foregoing inequality is transformed into $z > (z-q)^p/p^p$ for $z > 1$. Using this to bound the integral in (25), the desired result is obtained as

$$\eta_N = \int_1^{\infty}\frac{p^N}{z(z-q)^N}dz < \int_1^{\infty}\frac{p^{N+p}}{(z-q)^{N+p}}dz = \frac{p}{N-q}. \tag{29}$$

Since $b_{N,p}$ is a derivable function of p, it will be monotonically decreasing with p if and only if $d(\eta_N/p)/dp < 0$. The integrand in (25) is derivable, and thus, from the Leibniz rule

$$\begin{aligned}\frac{d\left(\frac{\eta_N}{p}\right)}{dp} &= \int_1^{\infty}\frac{(N-1)p^{N-2}(z-q) - Np^{N-1}}{z(z-q)^{N+1}}dz \\ &= \frac{(N-1)\eta_N - N\eta_{N+1}}{p^2}. \end{aligned} \tag{30}$$

From (3), the following relationship is obtained for $N \geq 2$:

$$\eta_{N+1} = \frac{p}{q}\left(\frac{1}{N} - \eta_N\right). \tag{31}$$

Using (31), (30) is transformed into

$$\frac{d\left(\frac{\eta_N}{p}\right)}{dp} = -\frac{1}{pq} + \frac{N-q}{p^2 q}\eta_N \tag{32}$$

which is negative because of (29).

Using (31) and (29), we can write

$$\begin{aligned}\frac{b_{N+1,p}+1}{b_{N,p}+1} &= \frac{N\eta_{N+1}}{(N-1)\eta_N} = \frac{p}{(N-1)q\eta_N} - \frac{Np}{(N-1)q} \\ &> \frac{N-q}{(N-1)q} - \frac{Np}{(N-1)q} = 1 \end{aligned} \tag{33}$$

which establishes the increasing character of $b_{N,p}$ with N.

Appendix B
Mean-Square Error in NBMC

The MSE in NBMC can be expressed as

$$\begin{aligned}\mathrm{E}[(\hat{\boldsymbol{p}}-p)^2] &= \mathrm{E}[\hat{\boldsymbol{p}}^2] - 2p\mathrm{E}[\hat{\boldsymbol{p}}] + p^2 \\ &= (N-1)^2\mathrm{E}\left[\frac{1}{\boldsymbol{n}^2}\right] - 2p^2 b_{N,p} - p^2. \end{aligned} \tag{34}$$

Using the $k^{(i)}$ notation introduced in Appendix A, the term $\mathrm{E}[1/\boldsymbol{n}^2]$ is given by

$$\mathrm{E}\left[\frac{1}{\boldsymbol{n}^2}\right] = \frac{p^N}{(N-1)!}\sum_{k=N}^{\infty}\frac{(k-1)^{(N-1)}q^{k-N}}{k^2}. \tag{35}$$

Considering that $k^2 > (k-1)(k+1)$, for $N \geq 3$ the following bound is obtained from (35):

$$\begin{aligned}\mathrm{E}\left[\frac{1}{\boldsymbol{n}^2}\right] &< \frac{p^N}{(N-1)!}\sum_{k=N}^{\infty}\frac{(k-2)^{(N-2)}q^{k-N}}{k+1} \\ &= \frac{p^N}{(N-1)!}\sum_{k=N}^{\infty}(k-3)^{(N-3)}q^{k-N} - \frac{3p^N}{(N-1)!} \\ &\quad \cdot \sum_{k=N}^{\infty}\frac{(k-3)^{(N-3)}q^{k-N}}{k+1} = \frac{p^2}{(N-1)(N-2)} \\ &\quad - \frac{3p^2}{(N-1)(N-2)}\sum_{k=N-2}^{\infty}\frac{f_{N-2}[k]}{k+3}. \end{aligned} \tag{36}$$

According to (26)

$$\sum_{k=N-2}^{\infty}\frac{f_{N-2}[k]}{k+3} > \frac{p}{N-2+3p}. \tag{37}$$

From (34), (36), (37), and the lower bound in (6), it stems that

$$\frac{\mathrm{E}[(\hat{\boldsymbol{p}}-p)^2]}{p^2} < \frac{\min\{\alpha_1(p), \alpha_2(p)\}}{N-2} \tag{38}$$

with $\alpha_1(p) = 1 + 2p - 3p(N-1)/(N-2+3p)$ and $\alpha_2(p) = 1 + 2(N-2)/N - 3p(N-1)/(N-2+3p)$.

We will show that $\alpha_1(p) < 1$ for $p < 2/3$, whereas $\alpha_2(p) < 1$ for $p \geq 2/3$. Computing

$$\frac{d\alpha_1(p)}{dp} = 2 - \frac{3(N-1)(N-2)}{(N-2+3p)^2} \tag{39}$$

and solving for $d\alpha_1(p)/dp = 0$, it is seen that only one solution exists in $0 \leq p \leq 1$, given by

$$p_1 = -\frac{N-2}{3} + \sqrt{\frac{(N-1)(N-2)}{6}} \tag{40}$$

with $d\alpha_1(p)/dp < 0$ for $0 \leq p < p_1$ and $d\alpha_1(p)/dp > 0$ for $p_1 < p \leq 1$. $\alpha_1(0) = 1$ and $\alpha_1(2/3) = 1/3 + 2/N \leq 1$ if $N \geq 3$, which implies that $\alpha_1(p) < 1$ for $0 \leq p < 2/3$, $N \geq 3$. $\alpha_2(p)$ is a decreasing function with $\alpha_2(2/3) = 1 - 2/N < 1$, and thus $\alpha_2(p) < 1$ for $2/3 \leq p \leq 1$. Therefore, (38) assures that (7) holds for $N \geq 3$.

The behavior of $d\alpha_1(p)/dp$ implies that $\alpha_1(p)$ is concave, and solving for $\alpha_1(p) = 1 - p$ yields 0 and 1/3. Thus $\alpha_1(p) \leq$

$1-p$ for $0 \le p \le 1/3$, and then (38) implies (8) for this range of values.

Appendix C
Confidence Level in NBMC

In the following, we denote, for convenience, $a = (N-1)\mu_2$ and $b = (N-1)/\mu_1$.

Proof of (12)

Let $\beta(n,p;i)$ denote the binomial probability function with parameters n and p evaluated at i. From the relationship between binomial and negative binomial distributions [7, p. 96]

$$P_2(p) = 1 - \sum_{i=0}^{N-1} \beta(n_2,p;i) \tag{41}$$

$$P_2(p) = \sum_{i=N}^{n_2} \beta(n_2,p;i) \tag{42}$$

$$P_1(p) = 1 - \sum_{i=0}^{N-1} \beta(n_1-1,p;i) \tag{43}$$

$$P_1(p) = \sum_{i=N}^{n_1-1} \beta(n_1-1,p;i). \tag{44}$$

Let $\pi(\lambda;i)$ denote the Poisson probability function with parameter λ evaluated at i. Since $n_2 p \to a$ and $n_1 p \to b$ as $p \to 0$, the Poisson theorem [7, p. 113] implies that $P_2(p) \to \bar{P}_2$ and $P_1(p) \to \bar{P}_1$, with $\bar{P}_2 = 1 - \sum_{i=0}^{N-1} \pi(a;i) = \gamma(N,a)$, $\bar{P}_1 = 1 - \sum_{i=0}^{N-1} \pi(b;i) = \gamma(N,b)$. This establishes (12).

Proof of (13) and (14)

We first consider the term $P_2(p)$. For a given $n_2 \ge a$, which corresponds to p in the interval $(a/(n_2+1), a/n_2]$, it is seen from (41) that $P_2(p)$ is a continuous, differentiable function of p. In addition, $P_2(p)$ is monotonically increasing in the referred interval, as we now show. It is easily seen that the term $\beta(n_2,p;i)$ is monotonically increasing with p for $p < i/n_2$, and decreasing for $p > i/n_2$. If $a \ge N$, we have that $a \ge (N-1)a/(a-1) \ge (N-1)(n_2+1)/n_2$, which substituted into the inequality $p > a/(n_2+1)$ implies that $p \ge (N-1)/n_2$. Therefore, all terms $\beta(n_2,p;i)$ in (41) are decreasing, and $P_2(p)$ is increasing. If $a < N$, $p \le a/n_2$ implies that $p < N/n_2$, therefore all terms $\beta(n_2,p;i)$ in (42) are increasing, and so is $P_2(p)$.

Let β_χ be the binomial distribution with parameters n_2, $p_\chi = a/n_2$, and π the Poisson distribution with parameter a. The information divergence $D(\beta_\chi \| \pi)$ between these distributions satisfies [12, Th. 12]

$$D(\beta_\chi \| \pi) \le \left(\frac{1}{2(1-p_\chi)} + \frac{1}{4n_2} - \frac{1}{4} \right) p_\chi^2 \le \frac{p_\chi^2}{2(1-p_\chi)}. \tag{45}$$

For a binomial distribution β_ϕ with parameters n_2, $p_\phi = a/(n_2+1)$, $D(\beta_\phi \| \pi)$ can be computed in a similar manner to that used in [12] for $D(\beta_\chi \| \pi)$. Let $a' = n_2 p_\phi$, and let π' denote the Poisson distribution with parameter a'. Then

$$\begin{aligned} D(\beta_\phi \| \pi) &= \sum_{i=0}^{n_2} \beta_\phi[i] \left(\ln \frac{n_2^{(i)} (1-p_\phi)^{n_2-i}}{(n_2+1)^i} + a' + p_\phi \right) \\ &= D(\beta_\phi \| \pi') - n_2 p_\phi \ln \frac{n_2+1}{n_2} + p_\phi. \end{aligned} \tag{46}$$

Using (46), (45), and the inequality $\ln(1+1/n_2) > 1/(n_2+1)$, $D(\beta_\phi \| \pi)$ can be bounded as

$$D(\beta_\phi \| \pi) < D(\beta_\phi \| \pi') + \frac{p_\phi}{n_2+1} \le \left(\frac{1}{2(1-p_\phi)} + \frac{1}{a} \right) p_\phi^2. \tag{47}$$

From (45), using Pinsker's inequality [12, Sec. 2] and taking into account that $\sum_{i=0}^{\infty} (\beta_\chi[i] - \pi[i]) = 0$, we obtain

$$\begin{aligned} |P_2(p_\chi) - \bar{P}_2| &\le \sum_{i=0}^{N-1} |\beta_\chi[i] - \pi[i]| \\ &\le \frac{1}{2} \sum_{i=0}^{\infty} |\beta_\chi[i] - \pi[i]| < \frac{p_\chi}{2} \sqrt{\frac{1}{1-p_\chi}}. \end{aligned} \tag{48}$$

Likewise

$$\lim_{p \to p_\phi +} |P_2(p) - \bar{P}_2| = |P_2(p_\phi) - \bar{P}_2| < \frac{p_\phi}{2} \sqrt{\frac{1}{1-p_\phi} + \frac{2}{a}}. \tag{49}$$

Since $P_2(p)$ is a continuous, monotonic increasing function in $(p_\phi, p_\chi]$, $|P_2(p) - \bar{P}_2|$ is upper bounded in that interval by $\max\{|P_2(p_\chi) - \bar{P}_2|, |P_2(p_\phi) - \bar{P}_2|\}$. Taking into account that $p_\chi = a/\lfloor a/p \rfloor < a/(a/p - 1) = ap/(a-p)$ and $p_\phi = a/(\lfloor a/p \rfloor + 1) < p$, we obtain from (48) and (49)

$$|P_2(p_\chi) - \bar{P}_2| < \frac{ap}{2(a-p)} \sqrt{\frac{1}{1 - \frac{ap}{a-p}}} \tag{50}$$

$$|P_2(p_\phi) - \bar{P}_2| < \frac{p}{2} \sqrt{\frac{1}{1-p} + \frac{2}{a}}. \tag{51}$$

Since $a > N-1$, we can write

$$|P_2(p_\chi) - \bar{P}_2| < \frac{p}{2} \frac{N-1}{N-1-p} \sqrt{\frac{1}{1 - \frac{(N-1)p}{N-1-p}}} \tag{52}$$

$$|P_2(p_\phi) - \bar{P}_2| < \frac{p}{2} \sqrt{\frac{1}{1-p} + \frac{2}{N-1}}. \tag{53}$$

Therefore, $|P_2(p) - \bar{P}_2|$ can be bounded as

$$\begin{aligned} |P_2(p) - \bar{P}_2| &\le \max\left\{ |P_2(p_\chi) - \bar{P}_2|, |P_2(p_\phi) - \bar{P}_2| \right\} \\ &< \frac{p}{2} \sqrt{\frac{\left(\frac{N-1}{N-1-p}\right)^2}{1 - \frac{(N-1)p}{N-1-p}} + \frac{2}{N-1}}. \end{aligned} \tag{54}$$

A similar approach can be used for $P_1(p)$. If $b \le Np$, (11) implies that $n_1 < N+1$, and therefore, $P_1(p) = 0$. Let us assume that $b > Np$. The interval of p corresponding to a given $n_1 \ge b+1$ is $[b/n_1, b/(n_1-1))$. Applying analogous arguments as before, with b replacing a, n_1-1 replacing n_2, $p_\phi = b/\lceil b/p\rceil \le p$, and $p_\chi = b/(\lceil b/p\rceil - 1) < bp/(b-p)$, we obtain

$$|P_1(p_\chi) - \bar{P}_1| < \frac{bp}{2(b-p)}\sqrt{\frac{1}{1-\frac{bp}{b-p}}} \tag{55}$$

$$|P_1(p_\phi) - \bar{P}_1| < \frac{p}{2}\sqrt{\frac{1}{1-p}+\frac{2}{b}} \tag{56}$$

and

$$|P_1(p) - \bar{P}_1| < \frac{p}{2}\sqrt{\frac{\left(\frac{N}{N-1}\right)^2}{1-\frac{Np}{N-1}}+\frac{2}{Np}}. \tag{57}$$

From inequalities (54) and (57), we obtain (13) for $b > Np$. If $b \le Np$, we have $P_1(p) = 0$, so that $c - \bar{c} = P_2(p) - \bar{P}_2 + \bar{P}_1$, and (14) results.

Proof of (15)

We consider a set of parameters N, p, a, and b, with $N \ge 3$. First, we will show that $P_2(p) > \bar{P}_2$ for $\mu_2 \ge (N+\sqrt{N})/(N-1)$, i.e., for $a \ge N+\sqrt{N}$. Then we will show that $P_1(p) < \bar{P}_1$ for $\mu_1 \ge (N-1)/(N-\sqrt{3N/2})$, i.e., for $b \le N - \sqrt{3N/2}$, and p as in (15). Combining these results, we will obtain the inequality $c > \bar{c}$.

In order to prove that $P_2(p) > \bar{P}_2$, let us consider an arbitrary $n_2 \ge a$, and the corresponding interval for p, $(a/(n_2+1), a/n_2]$. As noted in the proof of (13) and (14), since $p > a/(n_2+1) > (N-1)/n_2$, each of the terms $\beta(n_2, p; i)$ in (41) is a continuous, monotonically decreasing function of p in that interval. Therefore, $P_2(p)$ is monotonically increasing in that interval, and thus it suffices to consider $p = a/(n_2+1)$. We will show that $\beta(n_2, a/(n_2+1); i)$ increases with n_2 for $i \le N-1$, i.e., that $\beta(n_2, a/(n_2+1); i) \ge \beta(n_2-1, a/n_2; i)$ for $i \le N-1$, $n_2 - 1 \ge a$. In view of (41), this will establish that $P_2(a/(n_2+1))$ decreases with n_2, and thus $P_2(p) \ge P_2(a/(n_2+1)) > \bar{P}_2$.

Let us define $U(i) = \ln\beta(n_2, a/(n_2+1); i) - \ln\beta(n_2-1, a/n_2; i)$ with $n_2 \ge a+1$. We need to show that $U(i) \ge 0$ for $i \le N-1$. Computing $dU(i)/di$ as if i were a continuous variable and taking into account that $i \le N-1 < a-1$, it is seen that $dU(i)/di \le 0$. Therefore, it suffices to show that $U(N-1) \ge 0$. In the following, we denote $U = U(N-1)$.

U is an infinitely derivable function of a for $n_2 > a$, with a minimum at $a = N$. Let us denote

$$\begin{aligned} U_{\min} &= U|_{a=N} \\ &= -n_2 \ln\frac{n_2+1}{n_2} + (n_2-N)\ln\frac{n_2-N+1}{n_2-N}. \end{aligned} \tag{58}$$

Using Taylor's theorem, we can express U for $a > N$ as

$$\begin{aligned} U = U_{\min} &+ \frac{(a-N)^2}{2(n_2-N)(n_2-N+1)} \\ &+ \left(\frac{n_2-N}{(n_2-\theta)^3} - \frac{n_2-N+1}{(n_2-\theta+1)^3}\right)\frac{(a-N)^3}{3} \end{aligned} \tag{59}$$

with $N < \theta < a$. It is easily verified that the last term in (59) is positive. Therefore, it suffices to prove that the sum of the other two terms is nonnegative, i.e., that

$$\begin{aligned} (a-N)^2 \ge 2(n_2-N)(n_2-N+1) &\left(n_2 \ln\frac{n_2+1}{n_2}\right. \\ &\left.-(n_2-N)\ln\frac{n_2-N+1}{n_2-N}\right). \end{aligned} \tag{60}$$

Applying Taylor's theorem to $\ln(1+t)$ and substituting $t = 1/n_2$ we can write

$$n_2 \ln\frac{n_2+1}{n_2} < 1 - \frac{1}{2n_2} + \frac{1}{3n_2^2}. \tag{61}$$

Similarly, with $t = 1/(n_2-N)$ we have

$$(n_2-N)\ln\frac{n_2-N+1}{n_2-N} > 1 - \frac{1}{2(n_2-N)} + \frac{n_2-N}{3(n_2-N+1)^3}. \tag{62}$$

In view of (61) and (62), a sufficient condition for inequality (60) to hold is

$$\begin{aligned} (a-N)^2 \ge 2(n_2-N)(n_2-N+1) &\left[\frac{1}{2(n_2-N)} - \frac{1}{2n_2}\right. \\ &\left.+\frac{1}{3n_2^2} - \frac{n_2-N}{3(n_2-N+1)^3}\right]. \end{aligned} \tag{63}$$

We now show that, for $N \ge 3$, the sum of the last two terms within the brackets, which will be denoted as x, is negative. Expressing

$$\begin{aligned} x &= \frac{1}{3n_2^2} - \frac{n_2-N}{3(n_2-N+1)^3} \\ &= \frac{(-2N+3)n_2^2 + 3(N-1)^2 n_2 - (N-1)^3}{3n_2^2(n_2-N+1)^3} \end{aligned} \tag{64}$$

it is easily seen that the roots of numerator and denominator are lower than $N+1+\sqrt{N}$, and $\lim_{n_2\to\infty} n_2^3 x = -2N/3+1 < 0$. Thus, for $n_2 \ge a+1 \ge N+1+\sqrt{N}$, x is negative. Then (63) can be replaced by the sufficient condition

$$\begin{aligned} (a-N)^2 &\ge 2(n_2-N)(n_2-N+1)\left(\frac{1}{2(n_2-N)} - \frac{1}{2n_2}\right) \\ &= \frac{N(n_2-N+1)}{n_2} \end{aligned} \tag{65}$$

which holds for $(a-N)^2 \ge N$, i.e., for $a \ge N+\sqrt{N}$.

In order to prove that $P_1(p) < \bar{P}_1$, let us consider an arbitrary $n_1 \ge b+1$ and the corresponding interval for p, $[b/n_1, b/(n_1-1))$. Since $p < b/(n_1-1) < (N-1)/(n_1-1)$, each of the terms $\beta(n_1-1, p; i)$ in (44) is a continuous, monotonically increasing function of p in that interval. Therefore, $P_1(p)$ is also monotonically increasing in that interval, and thus, it suffices to consider $p = b/(n_1-1)$. From (44), we can write

$$\begin{aligned} P_1\left(\frac{b}{n_1-1}\right) &= \sum_{i=N}^{n_1-1}\beta\left(n_1-1, \frac{b}{n_1-1}; i\right) \\ &= \sum_{i=0}^{n_1-N-1}\beta\left(n_1-1, 1-\frac{b}{n_1-1}; i\right). \end{aligned} \tag{66}$$

We will show that $\beta(n_1, 1-b/n_1; i+1) \geq \beta(n_1-1, 1-b/(n_1-1); i)$ for $i \leq n_1 - N - 1$, $n_1 - 1 \geq b+1$. In view of (66), this will establish that $P_1(b/(n_1-1))$ increases with n_1, and thus $P_1(p) \leq P_1(b/(n_1-1)) < \bar{P}_1$.

Let us define in analogy with $U(i)$

$$V(i) = \ln\beta\left(n_1, 1-\frac{b}{n_1}; i+1\right) - \ln\beta\left(n_1-1, 1-\frac{b}{n_1-1}; i\right)$$

with $n_1 \geq b+2$. We need to show that $V(i) \geq 0$ for $i \leq n_1 - N - 1$. $V(i)$ is minimum for $i = n_1 - N - 1$, V defined as $V(n_1 - N - 1)$ is minimum for $b = N$, and for $b < N$

$$V = V_{\min} + \frac{(b-N)^2}{2(n_1-N)(n_1-N-1)} + \left(-\frac{n_1-N}{3(n_1-\theta)^3} + \frac{n_1-N-1}{3(n_1-\theta-1)^3}\right)(b-N)^3 \quad (67)$$

$$V_{\min} = V|_{b=N} = (n_1-N-1)\ln\frac{n_1-N}{n_1-N-1} - (n_1-1)\ln\frac{n_1}{n_1-1} \quad (68)$$

with $b < \theta < N$. Now the last term in (67) is not positive [as it was in (59)]; however, it is a decreasing function of θ, and therefore, we can bound

$$V \geq V_{\min} + \frac{(b-N)^2}{2(n_1-N)(n_1-N-1)} + \left(-\frac{1}{3(n_1-N)^2} + \frac{1}{3(n_1-N-1)^2}\right)(b-N)^3. \quad (69)$$

Let $y = (n_1-N)^2 V_{\min}$. We now show that $y > -N/2$ for $n_1 \geq 7N/2$. Considering y as a function of n_1, from (68), we compute

$$\frac{1}{n_1-N}\frac{dy}{dn_1} = (3(n_1-N)-2)\ln\frac{n_1-N}{n_1-N-1} + (3n_1-N-2)\ln\frac{n_1-1}{n_1} - \frac{N}{n_1} \quad (70)$$

$$\frac{d}{dn_1}\left(\frac{1}{n_1-N}\frac{dy}{dn_1}\right) = 3\ln\frac{n_1-N}{n_1-N-1} - \frac{2}{n_1-N} + \frac{N}{n_1^2} - \frac{1}{n_1-N-1} + 3\ln\frac{n_1-1}{n_1} - \frac{N-1}{n_1-1} + \frac{N+2}{n_1} \quad (71)$$

$$\frac{d^2}{dn_1^2}\left(\frac{1}{n_1-N}\frac{dy}{dn_1}\right) = N\frac{Q(n_1)}{(n_1-N)^2(n_1-N-1)^2 n_1^3 (n_1-1)^2} \quad (72)$$

with

$$Q(n_1) = (-6N+4)n_1^4 + (14N^2+8N-6)n_1^3 - (11N^3+20N^2+5N-2)n_1^2 + (3N^4+12N^3+11N^2+2N)n_1 - (2N^4+4N^3+2N^2). \quad (73)$$

Expressing $Q(n_1) = (n_1 - 7N/2)S(n_1) + R$, it can be seen that all the coefficients of $S(n_1)$, as well as the term R, are negative for $N \geq 3$. Thus $Q(n_1) < 0$ for $n_1 \geq 7N/2$. Therefore, (72) is also negative, and (71) is a decreasing function of n_1, so that

$$\frac{d}{dn_1}\left(\frac{1}{n_1-N}\frac{dy}{dn_1}\right) > \lim_{n_1\to\infty}\frac{d}{dn_1}\left(\frac{1}{n_1-N}\frac{dy}{dn_1}\right) = 0. \quad (74)$$

Then (70) is an increasing function, with

$$\frac{1}{n_1-N}\frac{dy}{dn_1} < \lim_{n_1\to\infty}\frac{1}{n_1-N}\frac{dy}{dn_1} = 0 \quad (75)$$

and thus, $y > \lim_{n_1\to\infty} y$. Using Taylor expansions of the logarithms in (68), we can express

$$V_{\min} = -\frac{N}{2(n_1-N-1)(n_1-1)} + v(n_1) \quad (76)$$

with $\lim_{n_1\to\infty} n_1^2 v(n_1) = 0$, and therefore, $y > \lim_{n_1\to\infty}(n_1-N)^2 V_{\min} = -N/2$.

From (69), as $(n_1-N)^2 V_{\min} > -N/2$

$$V > -\frac{N}{2(n_1-N)^2} + \frac{(b-N)^2}{2(n_1-N)(n_1-N-1)} + \frac{2(n_1-N)-1}{3(n_1-N)^2(n_1-N-1)^2}(b-N)^3. \quad (77)$$

It is easily seen that, for $n_1 \geq 7N/2$, the right-hand side of (77) decreases with b. Thus, for μ_1 as in (15), i.e., for $b < N - \sqrt{3N/2}$, (77) implies that

$$V > -\frac{N}{2(n_1-N)^2} + \frac{3N}{4(n_1-N)(n_1-N-1)} - \frac{N}{(n_1-N)(n_1-N-1)^2}\sqrt{\frac{3N}{2}}. \quad (78)$$

Multiplying (78) by $(n_1-N)^2(n_1-N-1)^2$, it stems that the inequality $V \geq 0$ will be satisfied if

$$-\frac{1}{2}(n_1-N-1)^2 + \frac{3}{4}(n_1-N)(n_1-N-1) - \sqrt{\frac{3N}{2}}(n_1-N) \geq 0 \quad (79)$$

which is easily seen to hold for $n_1 \geq 7N/2$, i.e., for p as given in (15).

ACKNOWLEDGMENT

The authors are grateful to the anonymous reviewers and to the Editor for Wireless Systems Performance, F. Santucci, for their valuable comments. L. Mendo also thanks P. W. Glynn and P. Harremoës for providing some of the references, and J. Grajal for helpful discussions.

REFERENCES

[1] P. J. Smith, M. Shafi, and H. Gao, "Quick simulation: A review of importance sampling techniques in communications systems," *IEEE J. Sel. Commun.*, vol. 15, no. 5, pp. 597–613, May 1997.

[2] J.-C. Chen, D. Lu, J. S. Sadowsky, and K. Yao, "On importance sampling in digital communications—Part I: Fundamentals," *IEEE J. Sel. Commun.*, vol. 11, no. 4, pp. 289–299, Apr. 1993.

[3] P. W. Glynn and W. Whitt, "The asymptotic validity of sequential stopping rules for stochastic simulations," *Ann. Appl. Probab.*, vol. 2, no. 1, pp. 180–198, 1992.

[4] J.-C. Chen and J. S. Sadowsky, "On importance sampling in digital communications—Part II: Trellis-coded modulation," *IEEE J. Sel. Commun.*, vol. 11, no. 4, pp. 300–308, Apr. 1993.

[5] M. C. Jeruchim, "Techniques for estimating the bit error rate in the simulation of digital communication systems," *IEEE J. Sel. Commun.*, vol. SAC-2, no. 1, pp. 153–170, Jan. 1984.

[6] P. M. Hahn and M. Jeruchim, "Developments in the theory and application of importance sampling," *IEEE Trans. Commun.*, vol. COM-35, no. 7, pp. 706–714, Jul. 1987.

[7] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, 4th ed. New York: McGraw-Hill, 2002.

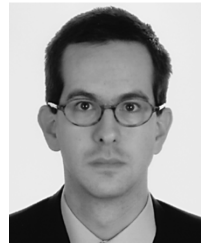
[8] K. S. Shanmugan and P. Balaban, "A modified Monte-Carlo simulation technique for the evaluation of error rate in digital communication systems," *IEEE Trans. Commun.*, vol. COM-28, no. 11, pp. 1916–1924, Nov. 1980.

[9] L. Mendo and J. M. Hernando, "Uplink and downlink traffic capacity performance in WCDMA systems," in *Proc. Wireless Des. Conf.*, vol. 1, May 2002, pp. 117–120.

[10] J. S. Sadowsky, "A new method for Viterbi decoder simulation using importance sampling," *IEEE Trans. Commun.*, vol. 38, no. 9, pp. 1341–1351, Sep. 1990.

[11] J. S. Stadler and S. Roy, "Adaptive importance sampling," *IEEE J. Sel. Commun.*, vol. 11, no. 4, pp. 309–316, Apr. 1993.

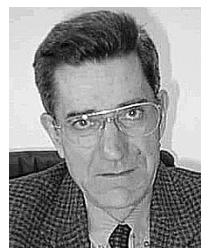
[12] P. Harremoës, "Convergence to Poisson distribution in information divergence," Math. Dept., Univ. of Copenhagen, Copenhagen, Denmark, Tech. Rep., 2003.



Luis Mendo was born in Madrid, Spain, on June 20, 1973. He received the M.Sc. and the Ph.D. degrees in telecommunication engineering from the Polytechnic University of Madrid, Madrid, Spain, in 1997 and 2001, respectively.

He has worked in radio network planning, and is currently with the Polytechnic University of Madrid. He is a coauthor of two textbooks on CDMA cellular systems, and has published several papers in the fields of CDMA system capacity and power control. His current research interests include CDMA networks, teletraffic, and Monte Carlo methods.

Dr. Mendo is the recipient of a national prize from the Telecommunication Engineering Professional Organization for his Ph.D. dissertation, and of two national awards for his M.Sc. thesis.



José M. Hernando (M'94) received the M.S. degree in communications engineering and the Ph.D. degree from Madrid Polytechnic University, Madrid, Spain, in 1967 and 1970, respectively.

Between 1967–1970, he was with ITT Laboratories of Spain doing research work in teletraffic theory. From 1970 until 1977, he was a Senior Engineer in the Communications Department of Iberia Airlines, Spain, working on planning, implementation, and maintenance of mobile radio networks, ground and air-to-ground in HF and VHF/UHF for both voice and data radiocommunications. In 1977, he joined the Signals Systems and Radiocommunications Department, Madrid Polytechnic University. Since then, he has been a Senior Professor devoted to educational and research work in radiocommunications. He has served as the Spanish Delegate in conferences and working groups in the Radiocommunications Sector of the International Telecommunications Union.

Dr. Hernando is member of the Spanish Bar of Telecommunications Engineers.