Conditional Monte Carlo with Intermediate Estimations for simulation of Markovian systems

Héctor Cancela
Facultad de Ingeniería
Universidad de la República
Montevideo, Uruguay
cancela@fing.edu.uy

Leslie Murray
FCEIA
Universidad Nacional de Rosario
Rosario. Argentina
leslie@eie.fceia.unr.edu.ar

Gerardo Rubino
INRIA–Rennes
Bretagne Atlantique, Campus de Beaulieu
Rennes, France
Gerardo.Rubino@inria.fr

***Abstract*—For systems that are suitable to be modelled by continuous Markov chains, dependability analysis is not always straightforward. When such systems are large and complex, it is usually impossible to compute their dependability measures exactly. An alternative solution is to estimate them by simulation, typically by Monte Carlo simulation. But for highly reliable systems standard simulation can not reach satisfactory accuracy levels (measured by the variance of the estimator) within reasonable computing times. Conditional Monte Carlo with Intermediate Estimations (CMIE) is a simulation method proposal aimed at making accurate estimations of dependability measures on highly reliable Markovian systems. The basis of CMIE is introduced, the unbiasedness of the corresponding estimator is proven, and its variance is shown to be lower than the variance of the standard estimator. A variant of the basic scheme, that applies to large and highly reliable multicomponent systems, is introduced. Some experimental results are shown.**

***Keywords*—dependability, simulation, rare event, Conditional Monte Carlo.**

I. Introduction

We consider systems which can be modelled by a continuous time homogeneous Markov chain X irreducible on the finite state space S (see [1], [2], [3] or Chapter 6 in [4]). The chain can also be absorbing and the techniques described here still work, but they are easier to present in the irreducible case. In this context, some measures of dependability need the evaluation of the probability $\gamma = \mathbb{P}\{\tau_D < \tau_{\mathbf{u}}\}$, where the times $\tau_{\mathbf{u}}$ and τ_D are defined as follows. The state space of the Markov chain is partitioned as $S = U \cup D$, such that in U the system is *up* and in D the system is *down*. The process X starts at some initial state $\mathbf{u} \in U$. Define $\tau_{\mathbf{u}}$ as the *return time to* $\mathbf{u}$, that is, $\tau_{\mathbf{u}} = \inf\{t > 0 \colon X(t) = \mathbf{u} \text{ and } X(t^-) \neq \mathbf{u}\}$, and τ_D as the *hitting time of* D, that is, $\tau_D = \inf\{t > 0 \colon X(t) \in D\}$.

The simplest and most basic dependability metric is the Mean Time To Failure (MTTF), defined as the expected life–time of the system, that is, the mean time until the system enters the subset D: MTTF = $\mathbb{E}\{\tau_D\}$. This metric admits the well–known representation $\mathrm{MTTF} = \mathbb{E}\{\min(\tau_D, \tau_{\mathbf{u}})\}/\gamma$.

Since we focus on the estimation of γ, we can just collapse all D into a single state $\mathbf{d}$, made absorbing. As before, event $\{\tau_{\mathbf{d}} < \tau_{\mathbf{u}}\}$ means that X gets absorbed at $\mathbf{d}$ before coming back to $\mathbf{u}$. For systems with a large (or infinite) number of states, the exact computation of γ is not feasible, and the standard *Monte Carlo* simulation will work, unless $\gamma \ll 1$, in which case we are facing a *rare event* problem, a context in which acceptable values of the estimator's variance can only be achieved at the expense of a very high number of replications. *Monte Carlo* methods must therefore be improved and adapted to address efficiently this *rare event* problem. Research has resulted in a large number of solutions in this regard, most of which derive from two well known families of techniques named, respectively, *Splitting* [5], [6], [7], [8], [9] and *Importance Sampling* [1], [10], [11].

Some applications of *Splitting* in the context of highly reliable systems can be found in [12] and [13], where the *reliability* and availability estimations of repairable systems are analysed using a variant called RESTART. Recently, some results in the context of static systems have been published in [14] and [15]. Some methods derived from *Importance Sampling*, like Zero-Variance [16], [17], [18] and Cross-Entropy [19], [20] have been successfully applied in the simulation of systems affected by *rare events*.

Conditional Monte Carlo [15], [21] is a classic variance reduction technique that has not given rise to many methods in the field of rare events applied to reliability estimation. However, some applications can be found in [22], [23], [24], [25], but most of them are aimed at the *rare events* probability estimation in models that deal with heavy–tailed distributions. This article addresses the problem of reliability estimation in the model so far defined and introduces a *Conditional Monte Carlo* simulation scheme, suitable for the estimation of γ.

The rest of this paper is organized as follows. Section II shows a basic application of Conditional Monte Carlo on Markovian systems. Section III, the core of this paper, introduces modifications to the basic Conditional Monte Carlo algorithm, in order to make it usable and efficient. Sections IV, V and VI discuss some properties and features of the proposed method. Section VII shows how to apply it to the particular case of Markovian multicomponent systems. Some experimental results are included in Sections VI and VII. A comparison with *Splitting* is shown in Section VIII. Conclusions and future directions can be found in Section IX.

II. Conditional Monte Carlo algorithm

There are different simulation methods to estimate value of γ. In the crude or standard simulation, N_1 replications start at state $\mathbf{u}$ and they are simulated until they either come back to

$\mathbf{u}$, in time $\tau_{\mathbf{u}}$, or hit state $\mathbf{d}$, in time $\tau_{\mathbf{d}}$. Let I be the indicator random variable of the event $\{\tau_{\mathbf{d}} < \tau_{\mathbf{u}}\}$:

$$I = \begin{cases} 1 & \text{w.p.} \quad \gamma, \\ 0 & \text{w.p.} \quad 1-\gamma. \end{cases} \tag{1}$$

Then, $\gamma = \mathbb{E}\{I\}$. Its standard estimator, $\widehat{\gamma}_s$, is:

$$\widehat{\gamma}_s = \frac{1}{N_1} \sum_{j=1}^{N_1} I^{(j)}, \tag{2}$$

where $I^{(j)}, j = 1, 2, \ldots, N_1$ are N_1 independent values sampled from distribution (1).

Let $C = \{\mathbf{d}, k, \mathbf{u}\}$, where k is any state in S, other than $\mathbf{d}$ or $\mathbf{u}$, and let X_C be a random variable defined as *the first state in C hit by a replication started at* $\mathbf{u}$:

$$X_C = \begin{cases} \mathbf{d} & \text{w.p.} \quad p_{\mathbf{d}}, \\ k & \text{w.p.} \quad p_k, \\ \mathbf{u} & \text{w.p.} \quad p_{\mathbf{u}}. \end{cases}$$

See that $p_{\mathbf{d}} \leq \gamma$, because γ is the probability that any replication that starts at $\mathbf{u}$ reaches $\mathbf{d}$ before coming back to $\mathbf{u}$, whereas $p_{\mathbf{d}}$ is the same probability, provided that "the path does not contain k". Similarly, $p_{\mathbf{u}} \leq 1 - \gamma$.

The expectation of I, conditioned on the values of X_C, is given by the following expressions: $\mathbb{E}\{I \mid X_C = \mathbf{d}\} = 1$, $\mathbb{E}\{I \mid X_C = k\} = \gamma_k$ and $\mathbb{E}\{I \mid X_C = \mathbf{u}\} = 0$ (γ_k is the probability that a replication that starts at state k, hits state $\mathbf{d}$ before it hits state $\mathbf{u}$). Thus, $\mathbb{E}\{I \mid X_C\}$ is a random variable with the following probability distribution:

$$\mathbb{E}\{I \mid X_C\} = \begin{cases} 1 & \text{w.p.} \quad p_{\mathbf{d}}, \\ \gamma_k & \text{w.p.} \quad p_k, \\ 0 & \text{w.p.} \quad p_{\mathbf{u}}, \end{cases} \tag{3}$$

and the following expectation:

$$\mathbb{E}\{\mathbb{E}\{I \mid X_C\}\} = \mathbb{E}\{I\} = 1 \times p_{\mathbf{d}} + \gamma_k \times p_k + 0 \times p_{\mathbf{u}} = \gamma.$$

The expected value of both random variables, I and $\mathbb{E}\{I \mid X_C\}$, is γ. As a consequence, another estimator of γ —namely, a *Conditional Monte Carlo* estimator— is:

$$\widehat{\gamma}_c = \frac{1}{N_1} \sum_{j=1}^{N_1} \mathbb{E}\{I \mid X_C^{(j)}\}, \tag{4}$$

where $\mathbb{E}\{I \mid X_C^{(j)}\}, j = 1, 2, \ldots$ are N_1 independent random variables sharing distribution (3). The samples $\mathbb{E}\{I \mid X_C^{(j)}\}$ are obtained in two steps: first, $X_C^{(j)}$ is sampled and then, the corresponding value $\mathbb{E}\{I \mid X_C^{(j)}\}$ is *computed*. In this introductory example the only three possible values of $X_C^{(j)}$ to be sampled are $\{\mathbf{u}, k, \mathbf{d}\}$, whereas the exact values of $\mathbb{E}\{I \mid X_C^{(j)}\}$ associated with them are, respectively, $\{1, \gamma_k, 0\}$.

If the set C includes more intermediate states besides k, the method applies as well. If, for example, $C = \{\mathbf{d}, 1, 2, \ldots, n, \mathbf{u}\}$, the distribution of $\mathbb{E}\{I \mid X_C\}$ becomes:

$$\mathbb{E}\{I \mid X_C\} = \begin{cases} 1 & \text{w.p.} \quad p_{\mathbf{d}}, \\ \gamma_1 & \text{w.p.} \quad p_1, \\ \gamma_2 & \text{w.p.} \quad p_2, \\ \vdots & \\ \gamma_n & \text{w.p.} \quad p_n, \\ 0 & \text{w.p.} \quad p_{\mathbf{u}}, \end{cases} \tag{5}$$

where γ_i is the probability that a replication that starts at state i hits state $\mathbf{d}$ before it hits state $\mathbf{u}$. Now:

$$\begin{aligned} \mathbb{E}\{\mathbb{E}\{I \mid X_C\}\} = \mathbb{E}\{I\} &= 1 \times p_{\mathbf{d}} + \sum_{i=1}^{n} \gamma_i \, p_i + 0 \times p_{\mathbf{u}} \\ &= \sum_{i=0}^{n} \gamma_i \, p_i = \gamma, \end{aligned}$$

where the notation $\gamma_0 = 1$ and $p_0 = p_{\mathbf{d}}$ is included for simplicity. The estimator given in Expression (4) remains valid, with the only difference of sampling from the distribution (5) instead of (3).

Figure 1 depicts the set of probabilities so far defined and shows the nomenclature used to refer to them in the rest of this article (as $\gamma_{\mathbf{u}} = 0$, the term $p_{\mathbf{u}} \times \gamma_{\mathbf{u}}$ equals 0, reason why it is shown in Figure 1 but does not appear in any further expression).

For any given set $C = \{\mathbf{d}, 1, 2, \ldots, n, \mathbf{u}\}$, call $\widetilde{C} = C \setminus \{\mathbf{d}, \mathbf{u}\}$, i.e. the subset formed only by the intermediate states, that is, $\widetilde{C} = \{1, 2, \ldots, n\}$.

The variance of the *Conditional Monte Carlo* estimator in (4) is:

$$\begin{aligned} \mathbb{V}\{\widehat{\gamma}_c\} &= \frac{1}{N_1} \left(\mathbb{E}\{\mathbb{E}\{I \mid X_C\}^2\} - \mathbb{E}\{\mathbb{E}\{I \mid X_C\}\}^2 \right) \\ &= \frac{1}{N_1} \left(\sum_{i=0}^{n} p_i \gamma_i^2 - \gamma^2 \right). \end{aligned} \tag{6}$$

On the other hand, the variance of the standard estimator given in (2) is known to be:

$$\mathbb{V}\{\widehat{\gamma}_s\} = \frac{1}{N_1} \left(\gamma - \gamma^2 \right) = \frac{1}{N_1} \left(\sum_{i=0}^{n} p_i \gamma_i - \gamma^2 \right). \tag{7}$$

Comparing expressions (6) and (7) and considering that $\gamma_i \leq 1$, $i = 0, \ldots, n$, because all these values are probabilities, it is clear that:

$$\sum_{i=0}^{n} p_i \gamma_i^2 \leq \sum_{i=0}^{n} p_i \gamma_i,$$

what means that the variance of the *Conditional Monte Carlo* estimator given in (6), is never larger than the *Standard Monte Carlo* estimator variance given in (7). This is, of course, a general fact on *Conditional Monte Carlo* methods, but it is worth making it explicit in our context.

III. Conditional Monte Carlo with Intermediate Estimations

The main problem in the use of *Conditional Monte Carlo*, as it was introduced so far, is the fact that the values γ_1, γ_2, …, γ_n are unknown, and that may be even as hard to evaluate as the exact value of γ itself. To work around this problem, these values will be now replaced by estimators.

It will be shown that after such replacement, the method is still unbiased. This is the core of the proposal introduced in this article and the basis of the so–called *Conditional Monte Carlo with Intermediate Estimations* (*CMIE*) method. The method

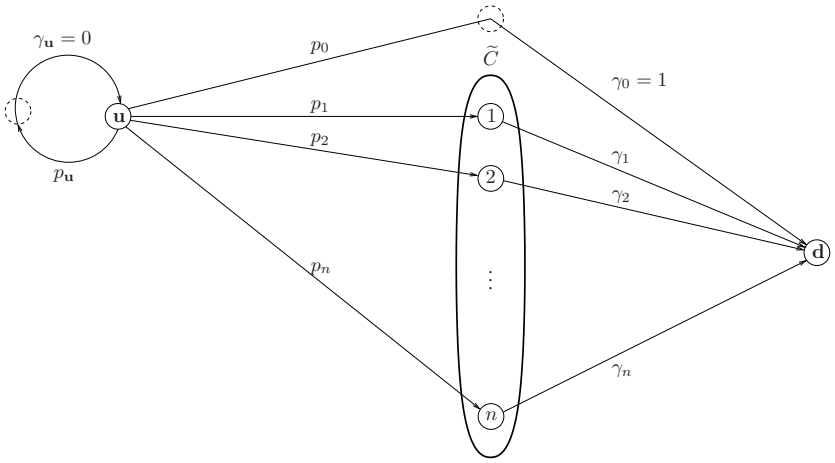

Fig. 1: The set of probabilities used in all calculations.

will now be described and, at the end of this section, the variance of the corresponding estimator will be determined.

To address the following calculation, it is better to express $\widehat{\gamma}$ in terms of the random vector $\bar{I} = (I_0, I_1, \ldots, I_{n+1})$, whose components are dependent binary random variables such that one and only one has value 1:

$$\bar{I} = \begin{cases} (1,0,0,\ldots,0,0) & \text{w.p.} \quad p_{\mathbf{d}}, \\ (0,1,0,\ldots,0,0) & \text{w.p.} \quad p_1, \\ (0,0,1,\ldots,0,0) & \text{w.p.} \quad p_2, \\ \qquad\vdots & \\ (0,0,0,\ldots,1,0) & \text{w.p.} \quad p_n, \\ (0,0,0,\ldots,0,1) & \text{w.p.} \quad p_{\mathbf{u}}. \end{cases} \tag{8}$$

Then, being $\gamma_0 = 1$, the standard estimator, $\widehat{\gamma}_s$, is:

$$\begin{aligned} \widehat{\gamma}_s &= \frac{1}{N_1}\sum_{j=1}^{N_1} I_0^{(j)} \times \gamma_0 + I_1^{(j)} \times \gamma_1 + I_2^{(j)} \times \gamma_2 + \ldots \\ &\quad + I_n^{(j)} \times \gamma_n + I_{n+1}^{(j)} \times 0 \\ &= \frac{1}{N_1}\sum_{j=1}^{N_1} \sum_{k=0}^{n} I_k^{(j)} \times \gamma_k \end{aligned} \tag{9}$$

In (9), the samples $I_0^{(j)}$, $I_1^{(j)}$, …, $I_n^{(j)}$ are obtained by the simulation, whereas the values γ_1, γ_2, …, γ_n must be calculated. However, if such calculation is too hard, or simply impossible, these values can be replaced by standard estimators. In order to do this, every time the simulation reaches a state $i \in \widetilde{C}$, N_2 independent replications must be started at i and simulated until they either reach $\mathbf{d}$ (and accumulate 1) or $\mathbf{u}$ (and accumulate 0). Once these N_2 replications started at i are completed, a standard estimator $\widehat{\gamma}_i$ can be evaluated and used in place of γ_i. To compute these estimations, define the set of Bernoulli random variables $\{J_i\}_{i=1}^n$, with the following probability distribution:

$$J_i = \begin{cases} 1 & \text{w.p.} \quad \gamma_i, \\ 0 & \text{w.p.} \quad 1-\gamma_i, \end{cases} \qquad i = 1, 2, \ldots, n. \tag{10}$$

The samples of J_i are obtained from the actual simulation of the Markov chain, which is the same as sampling them from distribution (10) ($J_0 = 1$ w.p. 1). Then, if γ_k is replaced by the estimator $\widehat{\gamma}_k$ in (9), the standard estimator is transformed into the *CMIE* estimator $\widehat{\gamma}_{cie}$:

$$\begin{aligned} \widehat{\gamma}_{cie} &= \frac{1}{N_1}\sum_{j=1}^{N_1} \left(\sum_{k=0}^{n} I_k^{(j)} \times \frac{1}{N_2}\sum_{i=1}^{N_2} J_k^{(i)} \right) \\ &= \frac{1}{N_1 N_2}\sum_{j=1}^{N_1}\sum_{k=0}^{n}\sum_{i=1}^{N_2} I_k^{(j)} J_k^{(i)}. \end{aligned}$$

It is simple to show that $\widehat{\gamma}_{cie}$ is unbiased:

$$\begin{aligned} \mathbb{E}\{\widehat{\gamma}_{cie}\} &= \frac{1}{N_1 N_2}\,\mathbb{E}\left\{ \sum_{j=1}^{N_1}\sum_{k=0}^{n}\sum_{i=1}^{N_2} I_k^{(j)} J_k^{(i)} \right\} \\ &= \frac{1}{N_1 N_2}\sum_{j=1}^{N_1}\sum_{k=0}^{n}\sum_{i=1}^{N_2} p_k\gamma_k \\ &= \sum_{k=0}^{n} p_k\gamma_k \;=\; \gamma. \end{aligned}$$

To determine the the variance of $\widehat{\gamma}_{cie}$, let $\bar{I}^{(x)}$ be any possible replication of $\bar{I}$, what means that $I_0^{(x)}$, $I_1^{(x)}$, …, $I_n^{(x)}$ are the components of this replication. Using the variance decomposition formula, the variance of the estimator can be written as:

$$\mathbb{V}\{\widehat{\gamma}_{cie}\} = \underbrace{\mathbb{V}\{\mathbb{E}\{\widehat{\gamma}_{cie} \mid \bar{I}^{(x)}\}\}}_{A} + \underbrace{\mathbb{E}\{\mathbb{V}\{\widehat{\gamma}_{cie} \mid \bar{I}^{(x)}\}\}}_{B}.$$

Terms A and B are analysed separately and after some algebra (see [26]), we obtain:

$$\begin{aligned} \mathbb{V}\{\widehat{\gamma}_{cie}\} = A + B = \frac{1}{N_1}\left(\sum_{k=0}^{n} p_k\gamma_k^2 - \gamma^2 \right) + \\ \frac{1}{N_1 N_2}\left(\gamma - \sum_{k=0}^{n} p_k\gamma_k^2 \right). \end{aligned} \tag{11}$$

Term A is the value of the variance of the *Conditional Monte Carlo* estimator when the values γ_1, γ_2, …, γ_n are known exactly (see (6)). Term B is the variance increase due to the fact that the values γ_1, γ_2, …, γ_n are replaced by estimators.

IV. Multiple Sets of Intermediate States

The key to the application of our *Conditional Monte Carlo* to Markov chains (as described in Section II) —call it *pure Conditional Monte Carlo*— is the knowledge of the probabilities γ_1, γ_2, ..., γ_n. The lack of these values makes it necessary to use estimations instead (as described in Section III). This technique is the heart of the *CMIE* method proposed in this article. As shown, the estimators $\widehat{\gamma}_1$, $\widehat{\gamma}_2$, ..., $\widehat{\gamma}_n$ can be obtained by standard simulation started every time one of the intermediate states 1, 2, ..., n is reached. But these values can be estimated more accurately, applying the same *Conditional Monte Carlo* method recursively, in the following way.

Suppose that two sets of intermediate states, $\widetilde{C}_1$ and $\widetilde{C}_2$, are defined, instead of one, as shown Figure 2. Assume that $\widetilde{C}_1 \cap \widetilde{C}_2 = \emptyset$ and $\mathbf{u}, \mathbf{d} \notin \widetilde{C}_1, \widetilde{C}_2$. Then, once a state $i \in \widetilde{C}_1$ is reached, N_2 replications must be started at state i, and they must be simulated until they either hit a state in $\widetilde{C}_2$, go back to $\mathbf{u}$, or get absorbed at $\mathbf{d}$. This can be considered the second recursive level of the simulation. It is intended to obtain the values γ'_1, γ'_2, ..., γ'_{n_1}, which indicate the probability that each of these N_2 replications get absorbed at $\mathbf{d}$. These values are not estimated by means of standard simulation, they are estimated more accurately by this recursive level of *Conditional Monte Carlo* simulation that makes use of $\widetilde{C}_2$ as the set of intermediate states. It is simple to extend this mechanism to more recursive levels (with more sets of intermediate states).

The variance analysis can be extended to the case of two sets of intermediate states, $\widetilde{C}_1$ and $\widetilde{C}_2$, in a straightforward manner. The probabilities involved are shown in Figure 2. The variance, computed in [26], is the following:

$$\begin{aligned}\mathbb{V}\{\widehat{\gamma}_{cie}\} =& \frac{1}{N_1}\left(\sum_{l=0}^{n_1} p_l\,{\gamma'_l}^2 - \gamma^2\right) + \\ & \frac{1}{N_1}\sum_{l=0}^{n_1} p_l\left(1/N_2\left(\sum_{k=0}^{n_2} p_{lk}\gamma_k^2 - {\gamma'_l}^2\right) + \right. \\ & \left. \frac{1}{N_2 N}\left(\gamma'_l - \sum_{k=0}^{n_2} p_{lk}\,\gamma_k^2\right)\right.\end{aligned}$$

Given this expression, it follows that the variance obtained in a model with two sets of intermediate states, $\widetilde{C}_1$ and $\widetilde{C}_2$, is lower than or equal to the variance obtained in a model with the single intermediate set of states $\widetilde{C}_1$.

V. Comparative Analysis of Variances

The variance of the *CMIE* estimator for the case of only one set of intermediate states, $\mathbb{V}\{\widehat{\gamma}_{cie}\}$, was derived in Section III. In this section, this variance is compared to the variance of other estimators, namely, the variance of the standard estimator, $\mathbb{V}\{\widehat{\gamma}_s\}$, shown in (7), and the variance of the *pure Conditional Monte Carlo* estimator, $\mathbb{V}\{\widehat{\gamma}_c\}$, derived in (6).

As $N_2 \to \infty$, $\mathbb{V}\{\widehat{\gamma}_{cie}\} \to \mathbb{V}\{\widehat{\gamma}_c\}$. Clearly, if the number of replications used in the estimation of the probabilities $\gamma_i,\ i = 0, \ldots, n$, is infinite, the estimators converge to the corresponding exact values, and the method becomes the *pure Conditional Monte Carlo*.

At the end of Section II it has been shown that $\mathbb{V}\{\widehat{\gamma}_c\} \leq \mathbb{V}\{\widehat{\gamma}_s\}$, meaning that the accuracy of *pure Conditional Monte Carlo* is never less than the accuracy of *Standard Monte Carlo*. It is clear that $\mathbb{V}\{\widehat{\gamma}_c\} \leq \mathbb{V}\{\widehat{\gamma}_{cie}\}$. We now prove that $\mathbb{V}\{\widehat{\gamma}_{cie}\} \leq \mathbb{V}\{\widehat{\gamma}_s\}$, meaning that the proposed estimator is never less accurate than *Standard Monte Carlo* estimator.

$$\begin{aligned}\mathbb{V}\{\widehat{\gamma}_{cie}\} &= \frac{\sum_{k=0}^{n} p_k\gamma_k^2 - \gamma^2}{N_1} + \frac{\gamma - \sum_{k=0}^{n} p_k\gamma_k^2}{N_1 N_2} \\ &\leq \frac{\sum_{k=0}^{n} p_k\gamma_k^2 - \gamma^2}{N_1} + \frac{\gamma - \sum_{k=0}^{n} p_k\gamma_k^2}{N_1} = \frac{\gamma - \gamma^2}{N_1} \\ &\leq \mathbb{V}\{\widehat{\gamma}_s\}.\end{aligned}$$

The inequality holds, no matter the values of N_1 and N_2. This means that the proposed estimator, $\widehat{\gamma}_{cie}$, is never less accurate than *Standard Monte Carlo* estimator, $\widehat{\gamma}_s$, even for a low number of replications N_1 and N_2.

Finally, the three variances involved are related as follows: $\mathbb{V}\{\widehat{\gamma}_c\} \leq \mathbb{V}\{\widehat{\gamma}_{cie}\} \leq \mathbb{V}\{\widehat{\gamma}_s\}$, which means that *CMIE* is always more accurate than crude or *Standard Monte Carlo*, but never as accurate as *pure Conditional Monte Carlo*, in which the exact values $\gamma_i,\ i = 0, \ldots, n$, are used.

VI. Intermediate States Analysis

The variance reduction capacity of *CMIE* depends on the choice of the set of intermediate states. In this section two properties of the sets of intermediate states are considered. Their proofs can be seen in [26]. The first one states that after adding a new state to an existing set, the variance of the estimator never increases and, therefore, a variance reduction may be expected. The second one says that if we compare the variance reduction obtained by two disjoint sets, the highest variance reduction comes from the cut that is somehow "closer" to the initial state, $\mathbf{u}$.

These two properties are consistent because, if the addition of one state to an existing set of intermediate states yields a variance that is lower than, or at worst equal to, the variance before the addition, the set that yields the least variance is the one composed of all the states: $C = S$. But, from the implementation point of view, the set $C = S$ is equivalent to the set formed by the adjacent states to $\mathbf{u}$, because, if $C = S$, for any replication started at the initial state $\mathbf{u}$, the only reachable states are the adjacent ones.

In the case of two or more sets of intermediate states, the choice of the second, and the consecutive ones, must be somehow similar to the choice of the first one with respect to the initial state $\mathbf{u}$. Whenever possible, the second set (between the existing one and state $\mathbf{d}$) must be formed by the adjacent states to the existing set. However, this is not straightforward and must be analyzed for every particular model.

These two properties are now tested on a continuous time Markov chain proposed and used by Juneja and Shahabuddin in [3] and shown in Figure 3. The system has 2 components of class A and 4 components of class B. The components can only be *operational* or *failed*. The state is the pair (N_A,N_B), where N_i indicates the number of *failed* components in class i. Failure rates of classes A and B are, respectively, $\epsilon/2$ and

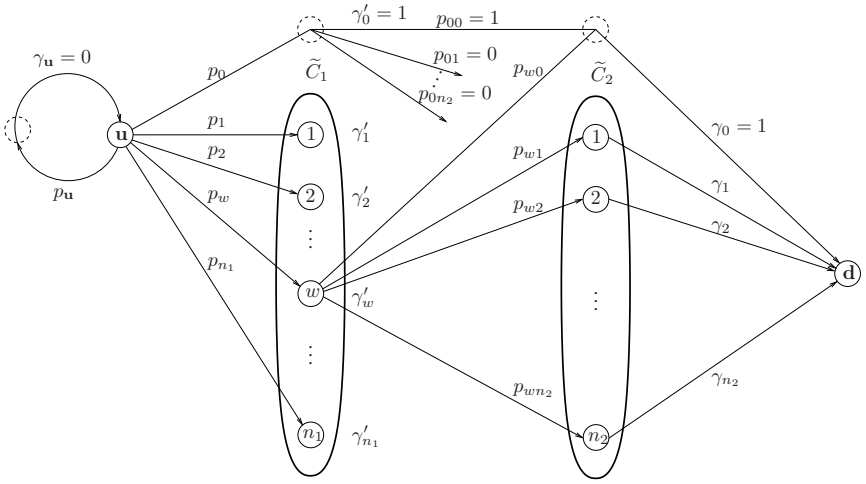

Fig. 2: The case of two sets of Intermediate States, $\widetilde{C}_1$ and $\widetilde{C}_2$

ϵ. The system fails if all components of all classes fail. Group repair (all *failed* components of a class are repaired simultaneously) begins if two components of the same class fail. Group repair rates for both classes are equal to 1. There is one repair-person in the system, and class A gets preemptive priority over class B.

Tables I, II and III show the results obtained by *CMIE* simulations over this model. The sets $\widetilde{C}_1$, $\widetilde{C}_2$ and $\widetilde{C}_3$ are all cuts between $\mathbf{u}$ and $\mathbf{d}$, and they are referred to, making use of the numbers placed above each state in Figure 3. Table III shows the ratio $\widehat{\mathbb{V}}\{\widehat{\gamma}_s\}/\widehat{\mathbb{V}}\{\widehat{\gamma}_{cie}\}$ when *CMIE* and standard simulation run the same execution time.

The *CMIE* method was programmed in the `C` language, using the `gcc` compiler. The estimator $\widehat{\gamma}_{cie}$ and an unbiased estimator of its variance were calculated as follows:

$$\widehat{\gamma}_{cie} = \frac{1}{N_1}\sum_{j=1}^{N_1}\gamma^{(j)} \quad \text{and}$$

$$\widehat{\mathbb{V}}\{\widehat{\gamma}_{cie}\} = \frac{1}{N_1 - 1}\left(\frac{1}{N_1}\left(\sum_{j=1}^{N_1}{\gamma^{(j)}}^2\right) - \widehat{\gamma}^2_{cie}\right). \qquad (12)$$

The results in Table I show that the cut that attains the lowest variance is the one formed by the adjacent states to $\mathbf{u}$. The variances of the estimators whose associated cut is close to state $\mathbf{u}$ are also low and quite similar. But, when the cuts are "far" from state $\mathbf{u}$, the variance greatly increases. In these experiments the number of replications started at state $\mathbf{u}$ was 10,000 and the number of replications launched from the intermediate states was also 10,000.

The experiment whose results are in Table II show that as the number of cuts increases, the variance of the estimator $\widehat{\gamma}_{cie}$ decreases. In these experiments the number of replications started at state $\mathbf{u}$ was 10,000 and the number of replications launched from the intermediate states was 100 for all cases.

The experiments in Table III are included to briefly show the variance reduction capacity of the *CMIE* method. The number of replications launched from intermediate states was 1,000 for all cases; the number of replications starting at state $\mathbf{u}$ was adjusted so that the total execution time of each of the four simulations was $t = 500$ sec. This time was fixed in advance and equal for all methods in order to have a fair comparison of the accuracy that was obtained by the different experiments.

TABLE I: Model in Figure 3, $\epsilon = 0.01$

$\widetilde{C}_1$	$\widehat{\gamma}_{cie}$	$\widehat{\mathbb{V}}\{\widehat{\gamma}_{cie}\}$
1–5	6.18E−06	6.28E−14
2–6–10	6.23E−06	6.88E−14
3–7–11	6.36E−06	6.90E−14
4–8–12	6.34E−06	1.53E−13
9–13	6.93E−06	6.56E−12

TABLE II: Model in Figure 3, $\epsilon = 0.01$

$\widetilde{C}_1$	$\widetilde{C}_2$	$\widetilde{C}_3$	$\widehat{\gamma}_{cie}$	$\widehat{\mathbb{V}}\{\widehat{\gamma}_{cie}\}$
1–5	—	—	4.00E−06	4.00E−12
1–5	2–6–10	—	6.19E−06	6.31E−14
1–5	2–6–10	3–7–11	6.13E−06	2.27E−15

VII. Application to Large Systems

Sometimes the state space S of the Markov chain is extremely large and, therefore, the choice of intermediate states is hard to be done explicitly. The idea of *CMIE* fits better to these models if it is adapted in the following way. In every replication, the computed values are samples of the probability

TABLE III: Model in Figure 3, $\epsilon = 0.001$

$\widetilde{C}_1$	$\widetilde{C}_2$	$\widetilde{C}_3$	$\widehat{\gamma}_{cie}$	$\widehat{\mathbb{V}}\{\widehat{\gamma}_{cie}\}$	$\widehat{\mathbb{V}}\{\widehat{\gamma}_s\}/\widehat{\mathbb{V}}\{\widehat{\gamma}_{cie}\}$
4–8–12	—	—	6.53E−09	4.60E−20	87
3–7–11	9–13	—	6.38E−09	7.75E−21	516
3–7–11	4–8–12	9–13	6.44E−09	1.28E−21	3125

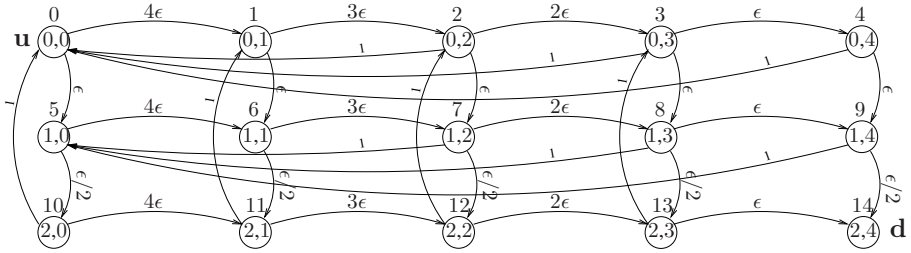

Fig. 3: Continuous time Markov chain used in the experimental variance analysis

of interest, γ, conditioned to the values of the random variable X_C (or to the intermediate states that are reached). In the end, the values of X_C are a function of the path $\pi_{\mathbf{u}}$ followed by the trajectory started at state $\mathbf{u}$. But the value of γ can be conditioned to events related to different functions that take the path $\pi_{\mathbf{u}}$ as their argument. In the following subsections three possible alternatives for these functions are introduced.

A. Forward Steps

For some systems it is possible to detect whether, at every jump, the replications move towards the final state $\mathbf{d}$, or not. In these cases it is possible to make the recursive calls only after moving $D \geq 1$ steps closer to the target state, every time. Proceeding this way, wherever the simulation starts, it must keep moving forwards and backwards until it either comes back to $\mathbf{u}$, or moves D steps forward to $\mathbf{d}$. If $\mathbf{u}$ is reached, the replication terminates; if the simulation moves D steps closer to $\mathbf{d}$, a new replication (recursive call) is launched.

B. Consecutive Failures

In multicomponent systems subject to fails and repairs, every failure produces a forward step, that is a step towards the final state $\mathbf{d}$. For a replication that starts at some state i, there are many ways (paths) to get D steps closer to the target state $\mathbf{d}$. One of them corresponds to the case in which D consecutive failures occur. If the system is composed of more that D components, there will be many different ways in which D consecutive failures may occur, all of them rarer than the case in which the D steps are completed after a zigzag of forward and backward steps. Thus, the indicator random variable I can be conditioned on such a sequence of D consecutive failures.

C. Measure of Rarity

Let $\pi_{i,j}$ be a path that starts at state i and ends at state j, without hitting state $\mathbf{u}$. If this path is composed of the sequence of states i, k, ..., l, j, the probability that the simulation goes through it, is: $\mathbb{P}\{\pi_{i,j}\} = p_{ik} \times \ldots \times p_{lj}$, where p_{xy} is the probability of going from state x to the neighbour state y, no matter if this jump is a fail or a repair.

In order to apply *CMIE*, the indicator variable I can be conditioned on the event $\mathbb{P}\{\pi_{i,j}\} \leq B$, where B is a fixed bound. But, as in highly reliable systems most of the probabilities p_{xy} are likely to be low, the values $\mathbb{P}\{\pi_{i,j}\}$ are likely to be extremely low. Therefore, it may be better to apply logarithm as follows: $-\log(\mathbb{P}\{\pi_{i,j}\}) = -\log(p_{ik}) - \ldots - \log(p_{lj})$ and to condition on the event $-\log(\mathbb{P}\{\pi_{i,j}\}) \geq W$ (where $W = -\log(B)$).

TABLE IV: *Tandem computer*, 1st version in [1]

Method	$\widehat{\gamma}_{cie}$	$\widehat{\mathbb{V}}\{\widehat{\gamma}_{cie}\} \times t$
FB	1.33E−06	3.37E−14
SFB	1.27E−06	2.06E−15
SFBP	1.27E−06	2.20E−15
Forward Steps	1.21E−06	4.20E−13
C. Failures	1.19E−06	3.93E−13
M. of Rarity	1.20E−06	5.38E−13

D. Experimental Comparison

The three implementations proposed are now subject to an experimental comparison. All the values —obtained by using the formulas indicated in (12)— are compared to results obtained from published papers.

The model used in the first set of experiments was used by Cancela et. al. in [1]. It is a computer that is composed of a multiprocessor, a dual disk controller, two RAID disk drives, two fans, two power supplies, and one dual interprocessor bus. When a component in a dual fails, the subsystem is reconfigured into a simplex. This tandem computer system requires all subsystems, one fan, and one power supply for it to be *operational*. The failure rates are 5ϵ, 2ϵ, 4ϵ, 0.1ϵ and 3ϵ for the processors, the disk controller, the disks, the fans, the power supplies and the bus respectively, with $\epsilon = 10^{-5}$ failures/hour. There is only one repairman and the repair rates are 30 repairs/hour for all the components, except for the bus, which has repair rate equal to 15 repairs/hour. In the experiments shown in Table IV, the multiprocessor and the disks have two units each, and only one is needed for the system to be working. FB, SFB and SFBP are all *Importance Sampling* methods used in [1]. Table V shows the results obtained for the same system, but with with a four-unit multiprocessor (only one of the four processors is required to have an *operational* system), and with each RAID being composed of 5 drives, only 3 of which are required for the system to be *operational*.

The third system used, also taken from [1], consists of a replicated database in which there are four sites, and each site has a whole copy of the database, on a RAID disk cluster. All clusters are identical, with the same redundancies (7-out-of-9), and with failure rate (for each disk) equal to $\epsilon = 10^{-2}$. There is one repairman per class, and the repair rate is 1. The system is considered *up* if there is at least one copy of the database working. Results are shown in Table VI.

Measure of Rarity is efficient only if failure and repair rates are considerably different. When this is not the case, the

TABLE V: *Tandem computer*, 2nd version in [1]

Method	$\widehat{\gamma}_{cie}$	$\widehat{\mathbb{V}}\{\widehat{\gamma}_{cie}\} \times t$
FB	1.24E−07	1.88E−15
SFB	—	1.57E−16
SFBP	1.25E−07	9.05E−17
Forward Steps	1.19E−07	5.55E−14
C. Failures	1.30E−07	6.17E−14
M. of Rarity	1.24E−07	1.11E−14

TABLE VI: *Replicated database* in [1]

Method	$\widehat{\gamma}_{cie}$	$\widehat{\mathbb{V}}\{\widehat{\gamma}_{cie}\} \times t$
FB	8.54E−13	8.65E−25
SFB	1.16E−12	3.93E−23
SFBP	8.87E−13	2.37E−25
Forward Steps	8.04E−13	4.41E−26
Consecutive Failures	8.10E−13	4.18E−23
Measure of Rarity	—	—

measure of rarity increases significantly at both, failures and repairs and, as a consequence, an increase of such measure is not an indication that the systems is moving towards the target event. In the case of the replicated database, failure and repair rates are, respectively, 10^{-2} and 1. Compared to the rates of the other systems analyzed, these rates are considerably close. This is the reason why *Measure of Rarity* is not computed in Table VI.

In the second set of experiments the models are the ones used by L'Ecuyer et. al. in [18]. In the first case (Example 5 in [18]), the system is composed of two sets of processors with two processors per set, two sets of disk controllers with two controllers per set, and six clusters of disks with four disks per cluster. The failure rates for processors, controllers, and disks are 5×10^{-5}, 2×10^{-5} and 2×10^{-5}, respectively. The repair rate is 1 for each type of component. In each disk cluster, data is replicated, which means that the failure of a single disk does not provoke a system's failure. The system is *operational* if all data is accessible from both processor types, meaning that at least one processor of each type, one controller of each set, and three disks of each cluster are *operational*. Results are shown in Table VII. BFB, SBLR, ZVA(v_0), ZVA(v_1), ZVA(v_2), and ZVA(v_3) are all *Importance Sampling* methods used in [18].

The last example is the one referred to as Example 6 in [18]. The system is composed of 20 types of components numbered from 0 to 19, with 4 components of each type. All repair rates are assumed to be 1, but component's failure rates differ: type–i components have failure rate $\lambda_i = (1+i/10)\epsilon$ for $0 \leq 1 \leq 9$ and $\lambda_i = i\epsilon^2/10$ for $10 \leq 1 \leq 19$, where $\epsilon = 10^{-3}$. The system is *failed* whenever a total of 7 components are *failed*. Results are shown in Table VIII.

All the *CMIE* estimations can be considered in the same order of precision and efficiency of the other methods to which the comparisons has been made.

VIII. CMIE vs. Splitting

If the sets of intermediate states are cuts (between $\mathbf{u}$ and $\mathbf{d}$) in the graph that models the Markov chain, there is a formal equivalence between *CMIE* and *Splitting* [5], [6], [7], [9], [27].

TABLE VII: Example 5 in [18] ($\gamma = 5.60\text{E}{-}05$)

Method	$\widehat{\gamma}_{cie}$	$\widehat{\mathbb{V}}\{\widehat{\gamma}_{cie}\} \times t$
BFB	—	4.93E−07
SBLR	—	1.17E−03
ZVA(v_0)	—	6.21E−11
ZVA(v_1)	—	3.90E−11
ZVA(v_2)	—	4.80E−11
Forward Steps	5.59E−05	3.08E−11
C. Failures	5.51E−05	2.83E−11
M. of Rarity	5.28E−05	9.88E−11

TABLE VIII: Example 6 in [18]

Method	$\widehat{\gamma}_{cie}$	$\widehat{\mathbb{V}}\{\widehat{\gamma}_{cie}\} \times t$
BFB	3.10E−11	9.35E−17
SBLR	—	—
ZVA(v_3)	3.00E−11	1.26E−22
Forward Steps	3.03E−11	1.74E−23
C. Failures	2.93E−11	4.28E−22
M. of Rarity	2.38E−11	8.70E−21

When the set $\widetilde{C}$ is a cut, the *CMIE* estimator takes the form:

$$\widehat{\gamma}_{cie} = \frac{1}{N_1}\sum_{j=1}^{N_1}\left(\sum_{k=1}^{n} I_k^{(j)} \times \frac{1}{N_2}\sum_{i=1}^{N_2} J_k^{(i)}\right) = \frac{1}{N_1 N_2}\sum_{j=1}^{N_1}\sum_{k=1}^{n}\sum_{i=1}^{N_2} I_k^{(j)} J_k^{(i)}. \quad (13)$$

A different analysis on the same model shows that, any path starting at state $\mathbf{u}$ has a probability, say P_1, to reach —any state of— the set $\widetilde{C}$ before coming back to $\mathbf{u}$. In the same way, a path starting from any state in the set $\widetilde{C}$ has a probability, say P_2, to reach state $\mathbf{d}$ before coming back to $\mathbf{u}$. The set $\widetilde{C}$ can be seen as a *bound* or *threshold* in the paths going from $\mathbf{u}$ to $\mathbf{d}$ and, therefore, *Splitting* can be applied in the estimation of γ. This *Splitting* estimation takes the form: $\widehat{\gamma} = \widehat{P}_1 \times \widehat{P}_2$, where $\widehat{P}_1$ and $\widehat{P}_2$ are, respectively, standard estimators of P_1 and P_2, as in any ordinary *Splitting* application. Figure 4 shows part of a set of replications, some of which start at $\mathbf{u}$ and goes forward to $\widetilde{C}$, and some others that start at $\widetilde{C}$ and goes forward to $\mathbf{d}$. According to this approach, the estimators of P_1 and P_2 are:

$$\widehat{P}_1 = \frac{1}{N_1}\sum_{j=1}^{N_1}\sum_{k=1}^{n} I_k^{(j)} \quad \text{and} \quad \widehat{P}_2 = \frac{\sum_{j=1}^{N_1}\sum_{k=1}^{n} I_k^{(j)} \sum_{i=1}^{N_2} J_k^{(i)}}{N_2 \sum_{j=1}^{N_1}\sum_{k=1}^{n} I_k^{(j)}},$$

and the *Splitting* estimator is:

$$\widehat{\gamma}_{spl} = \widehat{P}_1 \times \widehat{P}_2 = \frac{1}{N_1 N_2}\sum_{j=1}^{N_1}\sum_{k=1}^{n} I_k^{(j)} \sum_{i=1}^{N_2} J_k^{(i)} = \widehat{\gamma}_{cie}. \quad (14)$$

This leads to the conclusion that, if the set $\widetilde{C} = \{1, 2, \ldots, n\}$ is a cut in the graph of the Markov chain, *CMIE* and *Splitting* (based on a single level set) produce the same estimation. In other words, *Splitting* with a single level set $\widetilde{C}$ is the particular

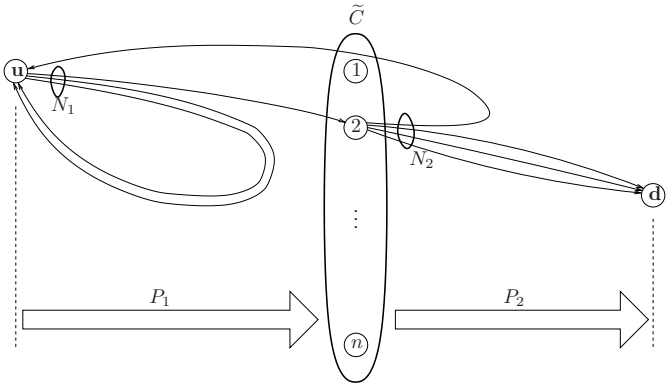

Fig. 4: Some trajectories in a *CMIE* vs. *Splitting* comparison

case of *CMIE* in which the set $\widetilde{C}$ is a cut in the graph of the Markov chain.

In a basic *Splitting* model there are *bounds* or *thresholds* between the initial and the final state, just like the set $\widetilde{C}$ in Figure 4. The consecutive probabilities P_1, P_2, ..., need to be estimated somehow. One of them is the probability of reaching the final state from the *threshold* that is immediately before. In systems like the ones used in Section VII, there are usually more than one final state scattered through all the Markov chain, some of which may be located between *thresholds*. This requires a particular effort to design a *Splitting* function of importance, while the application of *CMIE* is straightforward.

Another feature that may cause complications in a basic *Splitting* model is failure propagation. Sometimes a particular failure may cause the simultaneous occurrence of a set of other failures, with a given probability. In a basic *Splitting* model this translates into crossing more than one *threshold* simultaneously, what makes necessary to modify the basic approach according to system under analysis. *CMIE* is not affected by failure propagation.

IX. Conclusions and Open Research Lines

This article proposes a *Monte Carlo* method, referred to as Conditional Monte Carlo with Intermediate Estimations (*CMIE*), designed to reduce the variance of the estimator in a context of large and highly reliable Markovian systems.

CMIE was conceived to estimate the probability of visiting the failure state before coming back to the initial state (accepted as the state in which the system is *up*). The application of ordinary *Conditional Monte Carlo* to this type of model requires the knowledge of the exact value of some probabilities in the model. To overcome the fact that this probabilities are unknown, the proposal of *CMIE* is to estimate them, for which it is necessary to launch the method, recursively, from some selected states called intermediate states.

Splitting can be seen as the particular case of *CMIE* in which the events are implicitly defined by means of *thresholds* (cuts) in the state space of the Markov chain. However, the way in which the target probability is recursively computed in *CMIE* is simpler than the *Splitting* algorithm, which needs to determine the probabilities of crossing each *threshold* conditioned to the previous cross, keeping track of the number of times each *threshold* is crossed. Another advantage of *CMIE* over *Splitting*, comes up in systems in which there are more than one target state and/or fault propagation. The presence of more than one target state is a drawback in the determination of *thresholds* (cuts in the graph). Due to the presence of fault propagation, multiple *thresholds* crosses may occur. A particular effort then is required to adapt *Splitting* to these particular settings, whereas the *CMIE* implementations are straightforward and do not differ with respect to ones in which there is only one target state and there is no fault propagation.

CMIE was empirically tested against some other methods taken from the literature. In all cases, the results show that *CMIE* is in the same range of efficiency as the methods to which it was compared, not only in the variance, but also in the precision gain comparison. Some properties of *CMIE* were demonstrated and, besides, its variance was given a closed form. *CMIE* can be easily extended to other types of rare event problems like, for instance, network reliability estimation, either under a static (classic) or a dynamic approach [8].

One possible line of future work is to refine the asymptotic analysis of the behaviour of *CMIE* and to see, for instance, how close or how far it is to have Bounded Relative Error and/or Bounded Normal Approximation. Besides, two important issues to work on are: methods and criteria for the selection of the intermediate states sets and the trade–off between accuracy and execution time.

References

[1] H. Cancela, G. Rubino, and B. Tuffin, "MTTF estimation using Importance Sampling on Markov models." *Monte Carlo Methods and Applications*, vol. 8, no. 4, pp. 321–341, 2002.

[2] A. Goyal, P. Shahabuddin, P. Heidelberger, V. F. Nicola, and P. W. Glynn, "A unified framework for simulating markovian models of highly dependable systems," *IEEE Transactions on Computers*, vol. 41, no. 1, pp. 36–51, 1992.

[3] S. Juneja and P. Shahabuddin, "Fast Simulation of Markov Chains with Small Transition Probabilities," *Management Science*, vol. 47, no. 4, pp. 547–562, Apr. 2001. [Online]. Available: http://dx.doi.org/10.1287/mnsc.47.4.547.9827

[4] G. Rubino and B. Tuffin, *Rare Event Simulation Using Monte Carlo Methods*. Wiley, 2009. [Online]. Available: http://books.google.co.uk/books?id=KoRdR-pSfKsC

[5] M. J. J. Garvels, "The splitting method in rare event simulation," Ph.D. dissertation, Faculty of mathematical Science, University of Twente, The Netherlands, 2000.

[6] P. Glasserman, P. Heidelberger, P. Shahabuddin, and T. Zajic, "Splitting for rare event simulation: Analysis of simple cases," in *Proceedings of the 1996 Winter Simulation Conference*. San Diego, California: IEEE Computer Society Press, 1996, pp. 302–308.

[7] P. L'Ecuyer, V. Demers, and B. Tuffin, "Rare events, splitting, and quasi-Monte Carlo," *ACM Trans. Model. Comput. Simul.*, vol. 17, no. 2, Apr. 2007. [Online]. Available: http://doi.acm.org/10.1145/1225275.1225280

[8] L. Murray, H. Cancela, and G. Rubino, "A Splitting algorithm for network reliability estimation," *IIE Transactions*, vol. 45, no. 2, pp. 177–189, 2013. [Online]. Available: http://hal.inria.fr/hal-00773319

[9] M. Villén-Altamirano and J. Villén-Altamirano, "RESTART: A method for accelerating rare events simulations," in *Proceedings of the 13th International Teletraffic Congress*. North–Holland, 1991, pp. 71–76.

[10] Z. I. Botev, P. L'Ecuyer, and B. Tuffin, "Markov chain importance sampling with applications to rare event probability estimation," *Statistics and Computing*, vol. 23, no. 2, pp. 271–285, 2013. [Online]. Available: http://dx.doi.org/10.1007/s11222-011-9308-2

[11] P. W. Glynn and D. L. Iglehart, "Importance sampling for stochastic simulations," *Management Science*, vol. 35, no. 11, pp. 1367–1392, November 1989.

[12] J. Villén-Altamirano, "Importance functions for RESTART simulation of highly–dependable systems," *Simulation*, vol. 83, pp. 821–828, December 2007. [Online]. Available: http://portal.acm.org/citation.cfm?id=1363142.1363148

[13] ——, "RESTART simulation of non–Markov consecutive–k–out–of–n: F repairable systems," *Reliability Engineering & System Safety*, vol. 95, no. 3, pp. 247–254, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/B6V4T-4XHVH09-2/2/976438347aa6e9fd178b753fe1646bc7

[14] Z. I. Botev and D. P. Kroese, "Efficient Monte Carlo simulation via the generalized splitting method," *Statistics and Computing*, pp. 1–16, 2010, dOI: 10.1007/s11222-010-9201-4. [Online]. Available: http://dx.doi.org/10.1007/s11222-010-9201-4

[15] D. P. Kroese, T. Taimre, and Z. I. Botev, *Handbook of Monte Carlo Methods*, ser. Wiley Series in Probability and Statistics. Wiley, 2013. [Online]. Available: http://books.google.fr/books?id=Trj9HQ7G8TUC

[16] P. L'Ecuyer, G. Rubino, S. Saggadi, and B. Tuffin, "Approximate zero-variance importance sampling for static network reliability estimation," *IEEE Transactions on Reliability*, vol. 8, no. 4, pp. 590–604, 2011.

[17] P. L'Ecuyer and B. Tuffin, "Effective approximation of zero-variance simulation in a reliability setting," in *Proceedings of the 2007 European Simulation and Modeling Conference*. Ghent, Belgium: EUROSIS, 2007, pp. 48–54.

[18] ——, "Approximating Zero-Variance Importance Sampling in a Reliability Setting," *Annals of Operations Research*, vol. 189, no. 1, pp. 277–297, 2011.

[19] R. Y. Rubinstein and D. P. Kroese, *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte Carlo Simulation and Machine Learning*, ser. Information Science and Statistics. Springer, 2004. [Online]. Available: http://books.google.fr/books?id=9KZCZIogU4YC

[20] A. Ridder, "Importance sampling simulations of Markovian reliability systems using cross-entropy," *Annals of Operations Research*, vol. 134, pp. 119–136, 2005.

[21] S. M. Ross, *Simulation, Fourth Edition*. Orlando, FL, USA: Academic Press, Inc., 2006.

[22] J. Zuk and L. Rosenberg, "Rare-event simulation for radar threshold estimation in heavy-tailed sea clutter," in *IEEE Workshop on Statistical Signal Processing, SSP 2014, Gold Coast, Australia, June 29 - July 2, 2014*, 2014, pp. 436–439. [Online]. Available: http://dx.doi.org/10.1109/SSP.2014.6884669

[23] S. Asmussen and D. Kortschak, "Error rates and improved algorithms for rare event simulation with heavy Weibull tails," *Methodology and Computing in Applied Probability*, pp. 1–21, Jan. 2013.

[24] J. Blanchet, J. Li, and M. K. Nakayama, "A conditional monte carlo method for estimating the failure probability of a distribution network with random demands," in *Proceedings of the Winter Simulation Conference*, ser. WSC '11. Winter Simulation Conference, 2011, pp. 3837–3848. [Online]. Available: http://dl.acm.org/citation.cfm?id=2431518.2431974

[25] J. Chan and D. Kroese, "Rare-event probability estimation with conditional monte carlo," *Annals of Operations Research*, vol. 189, no. 1, pp. 43–61, 2011. [Online]. Available: http://dx.doi.org/10.1007/s10479-009-0539-y

[26] L. Murray, "New variance reduction methods in Monte Carlo rare event simulation," Ph.D. dissertation, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay, 2014.

[27] P. L'Ecuyer, F. Le Gland, P. Lezaud, and B. Tuffin, "Splitting techniques," in *Rare Event Simulation using Monte Carlo Methods*, G. Rubino and B. Tuffin, Eds. John Wiley & Sons, 2009, ch. 3, pp. 39–61.