RARE EVENT SIMULATION

Carmelita Görg, Eugen Lamers, Oliver Fuß*
Communication Networks, University of Bremen, Germany

Poul Heegaard†
Telenor, Norway

Abstract

Discrete event simulation as a method for performance evaluation has become an indispensable tool in many fields, e.g., teletraffic engineering. New communication networks and services pose extreme requirements regarding the quality of service, e.g., cell loss probabilities in ATM[1] networks in the order of 10^{-9}. Straightforward simulation for this type of rare event leads to simulation run times in the order of months and years, thus requiring new methods to be investigated and employed. In the following the state of the art in the field of Rare Event Simulation is described and examples are given for the RESTART[2]/LRE[3] method to demonstrate the type of results that can be achieved by using these methods. In recent years these methods have been applied to relevant quality of service parameters in ATM networks. Future applications will concern the investigation of the emerging quality of service concepts for Next Generation Internets.

RESTART/LRE is a multi-step simulation approach which reduces the simulation run time by several orders of magnitude from, e.g., years to minutes, thus making simulation studies for rare events feasible. The approach can be applied to single nodes and networks. It is a so-called importance splitting method, where system states that lead to the rare event are saved and used as the starting point for new simulation sub-runs. The results of the sub-runs are multiplied with the corresponding weights which are obtained in the previous step and result in the desired probabilities. The statistical evaluation method LRE used in this combined approach has the additional advantage of evaluating the so-called *local correlation function* which gives insight into the correlation

* {cg,elm,fuss}@comnets.uni-bremen.de, www.comnets.uni-bremen.de
† poul.heegaard@telenor.com

[1] ATM: Asynchronous Transfer Mode
[2] REpetitive Simulation Trials After Reaching Threshold
[3] Limited Relative Error

structure of the investigated random variable and is part of an error measure for controlling all phases of a simulation. Examples are given for the loss probability of ATM cells in G/G/1/N[4] type systems as well as certain tandem networks that represent a model for ATM reference connections. Additional results are presented where the rare event details of the delay time distribution are the object of the investigation. The RESTART/LRE simulation offers as a result the complementary distribution function of the delay time distribution so that the probability of delays longer than a given maximum can be deduced directly from this result. This is one of the quality of service parameters defined for ATM. Furthermore, new results are available where the method has been used to investigate rare event details of the loss probabilities for handover in wireless ATM networks. The underlying model for an ATM reference connection consists of a network with a sequence of */D/1/N[5] queues and an extra delay (propagation and switching) between these queues. Additional traffic is offered to the outgoing connections to model the interference of other ATM connections. Simulation results for the complementary distribution function of the delay time are presented.

Keywords: rare event simulation, RES, output analysis, variance reduction techniques, importance sampling, IS, importance splitting, optimization techniques, ATM networks, loss probability, occupancy distribution, delay time distribution, RESTART

1 INTRODUCTION

Performance of computer and communication systems is commonly characterized by the occurrence of rare events. For example, system unavailability requirements are typically less than 10^{-5} and cell loss probability in asynchronous transfer mode (ATM) switches less than 10^{-9}. The performance of such systems is frequently studied through simulation. However, estimation of rare event probabilities with the naive Monte Carlo techniques requires a prohibitively large number of trials in most interesting cases. Figure 1 shows the basic methods of simulation speed-up by using parallel or distributed computing power or stochastic methods. For rare event simulation (RES) two stochastic methods, called importance splitting/RESTART and importance sampling (IS), have been extensively investigated by the research simulation community in the last decade.

The basic idea of splitting [28] is to partition the state-space of the system into a series of nested subsets and to consider the rare event as the intersection of a nested sequence of events. When a given subset is entered by a sample

[4]G/G/1/N: Kendall classification for queuing systems, General Input, General Output, 1 Server, N Buffer Spaces, usually FIFO (First In First Out)

[5]*/D/1/N: queuing system with input from previous stages and deterministic service time

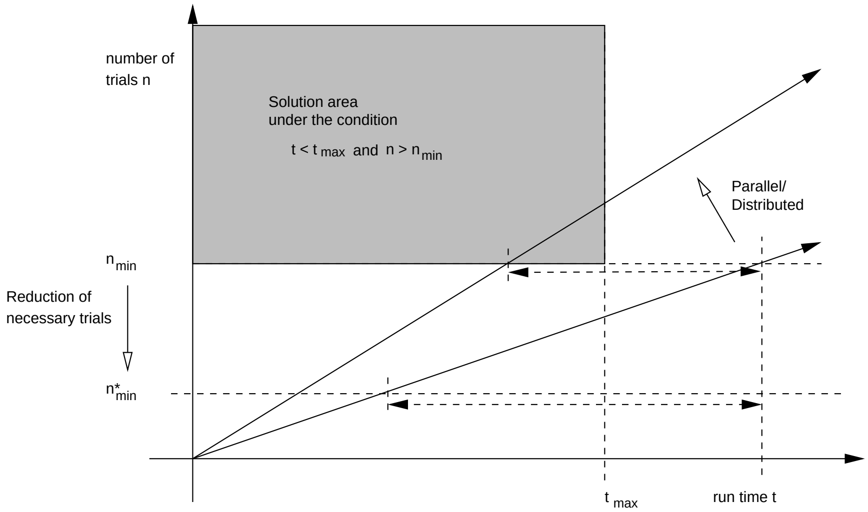


Figure 1: Simulation Speed-Up

trajectory during the simulation, numerous random retrials are generated with the initial state for each retrial being the state of the system at the entry point. Thus, by doing so, the system trajectory has been split into a number of new sub-trajectories, hence the name splitting. A similar idea has been developed in [52] [53] into a refined simulation technique under the name RESTART which has been extended by different authors to the multiple threshold case. A comparison of the different splitting/RESTART methods and recent substantial extensions are described in [12].

Importance sampling (IS) is also based on the idea to make the occurrence of rare events more frequent, or in other words, to *speed up* the simulation. Technically, IS aims to select a probability distribution (change of measure) that minimizes the variance of the IS estimate. The efficiency of IS (strongly) depends on obtaining a good change of measure. Large deviation results have proven their applicability for small systems with medium complexity, see e.g. [40]. This type of analysis is feasible only for relatively simple models, see [3] and [24] for surveys. To cope with large and realistic models, different techniques such as dynamic biasing [9] [14] [22] have been developed.

Most successful applications of rare event simulation techniques consider the evaluation of small ATM systems with a single switch or tandem queue. Studies include: consecutive cell loss in a single buffer, e.g., [8] [37], transfer delay and jitter [17], Connection Admission Control [35], Usage Parameter Control [21], and rate-controlled ATM switch (leaky bucket algorithm) [1].

Some speed-up is reported in studies of large networks that involve a multidimensional state space. In optimal design and optimal network routing one needs to solve stochastic optimization problems closely related to the classic traveling salesman problem (TSP), the shortest/longest path (SLP) and some other NP-hard problems. The work in [44] is an important contribution to this field. Another network issue is estimation of the call blocking probabilities. These have been obtained by Monte Carlo simulation with importance sampling on loss networks [35] [33]. In [23], well-engineered loss networks have been studied by simulation of Markov chain models with importance sampling. This is a flexible approach that is suitable for more general network problems, which has been recently extended in [7]. Other work on multidimensional problems has focused on a single network bottleneck that reduces the dimensionality of the problem [10]. Rare event simulation has also been applied to examples of multidimensional models with highly reliable components for dependability evaluation using importance sampling with failure biasing strategies [14] [25] [36], and the RESTART method [31] [50].

Several approaches have been used for modeling network traffic sources, including fluid-flow sources [30] [42], on-off sources [41], MPEG coded video sources [2], Composite Source Models [26], feedback and throttled sources [21], and self-similar sources [4] [6] [19].

RES approaches have proven to be efficient for quite a number of both trivial and non-trivial system examples. However, to become a success and to be applied by simulation users, the RES techniques have to be incorporated into a simulation tool and the setting of the critical parameters (both for RESTART and IS) have to be automated, or at least a "wizard" is required to assist in setting these. Examples of simulation tools where RES support is included are ASTRO with RESTART [54], UltraSAN with an IS module [38] [39], and MuSICS [11] [32] with RESTART/LRE.

In [34] a new adaptive IS algorithm is presented for rare event simulation based on the cross-entropy (CE) method, which works very efficiently for complex static systems, like the stochastic PERT networks. Its applicability to complex dynamic systems such as ATM networks is under investigation.

The next two sections describe the CE method and IS in more detail. The remaining sections are devoted to the RESTART/LRE method, first describing the method in general (section 4), then extending it to a queuing network (section 5). In section 6 the basic formulas of the splitting technique for the discrete arrival occupancy and the continuous delay time are given. The ATM reference connection model is described in section 7.1. Parameters and results are presented in section 7.2.

2 THE CROSS ENTROPY METHOD

An extremely useful feature of the CE method is that it can be readily modified for finding (estimating) the optimal solution in an NP-hard combinatorial problem [43] [45]. The CE method for combinatorial optimization problems (COP) presents an adaptive randomized algorithm, equipped with a random mechanism (an auxiliary discrete distribution or an auxiliary stochastic process, like a Markov chain). The aim of the random mechanism is to transform the original (deterministic) network into a stochastic one, called the associated stochastic network (ASN). As soon as the ASN is obtained, one casts the COP into the rare event framework in the sense that a very small probability associated with the optimal solution of COP is defined for the ASN. It can be shown numerically [44] [45] that the randomized algorithm converges with a very high probability to a very small subset of the original (huge) set of feasible solutions. Moreover, the optimal solution of COP belongs to this rare set with very high probability. Each iteration of the CE randomized algorithm contains the following two phases:

1. Generating random trajectories (walks) on the network graph using the auxiliary process and simultaneous calculation of the objective function, such as the trajectory length.

2. Updating the parameters of the auxiliary process on the basis of the data collected in the first phase.

It is shown in [44] [45] that after some finite number of iterations the auxiliary process converges to a degenerated one, which uniquely defines the optimal trajectory, such as the shortest tour in the TSP. Depending on the particular problem, the randomness in ASN is introduced by making either the nodes or the edges of the network random. More specifically, one has to distinguish between the so-called:

- Stochastic node networks (SNN), where the trajectories are generated using an n-dimensional discrete distribution, like the n-dimensional Bernoulli distribution with independent components, such that each of its components is uniquely associated with a fixed node in the network. Examples of SNN are the maximal cut (max-cut) and the bipartition problems.

- Stochastic edge networks (SEN), where the trajectories are generated using a Markov chain (MC) with a given probability transition matrix, such that the transition from state i to state j in the MC uniquely defines the edge *(ij)* in the network. Examples of SEN are the SLP, TSP and the assignment problem.

It is crucial to note that the above CE method is suitable for both deterministic and stochastic (noisy) combinatorial optimization problems, where, for example, one may assume that deterministic and stochastic combinatorial optimization differ with respect to whether the length of their edges (arcs) in the network are deterministic quantities or random variables, respectively. Note in addition that optimization of stochastic (noisy) networks is associated with flow control, optimal routing of data and communication networks and with other simulation-based optimization models. Note finally that most combinatorial optimization problems involving data networks make simplified (deterministic) assumptions, since no algorithms are available for combinatorial optimization in a more realistic (stochastic) environment. CE may offer a major breakthrough to the entire field of stochastic network optimization.

3 IMPORTANCE SAMPLING

Importance sampling changes the dynamics of the simulation model. The dynamics must be changed to increase the number of rare events of interests. Let $f(x)$ be the original sampling distribution and X the stochastic variable. If $g(x)$ is the property of interest, the original problem is that the probability of observing $g(x) > 0$ is very small when taking samples from $f(x)$. To solve this problem, importance sampling changes the sampling distribution to a new distribution $f^*(x)$. Taking samples from $f^*(x)$ should increase the probability of observing $g(x) > 0$ significantly. To retain an unbiased estimate the observations must be corrected because the samples x are from $f^*(x)$ and not $f(x)$.

Let the property of interest be $\gamma = E[g(X)]$. This can now be rewritten:

$$\gamma = E_f[g(X)] = \int g(x)f(x)\mathrm{d}x = \int g(x)\frac{f(x)}{f^*(x)}f^*(x)\mathrm{d}x = \\ E_{f^*}[g(X) \cdot L(X)] \tag{1}$$

where $L(x) = f(x)/f^*(x)$ is denoted the *likelihood ratio*. Observe that the expected value of the observations under f is equal to the expected value of the observations under f^* corrected for bias by the likelihood ratio, $E_f[g(X)] = E_{f^*}[g(X) \cdot L(X)]$.

An unbiased estimator for γ, taking samples x from $f^*(x)$, is

$$\hat{\gamma}_{\mathrm{IS}} = \frac{1}{R}\sum_{r=1}^{R} g(x_r)L(x_r) \tag{2}$$

with variance

$$\mathrm{Var}[\hat{\gamma}_{IS}] = \frac{1}{R}\mathrm{Var}_{f^*}[g(X)L(X)] = \frac{1}{R}\mathrm{E}_{f^*}[(g(X)L(X) - \gamma)^2]\,. \qquad (3)$$

To obtain the optimal – or at least a good – *change of measure* from the original sampling distribution $f(x)$ to a new distribution $f^*(x)$, is the main challenge with respect to making importance sampling efficient and robust. From equation (1), the only restriction to the probability density $f^*(x)$ is that $f^*(x) > 0$ for all samples x where $g(x)f(x) \neq 0$. This means that the samples x with a positive probability $f(x) > 0$ and a non-zero contribution $g(x)$ must have a positive probability also in the new distribution, $f^*(x)$. This is a necessary condition which serves as a guideline for choosing $f^*(x)$. The efficiency of the importance sampling estimate in (2) is dependent on an appropriate choice of the new distribution. An unfortunate choice of $f^*(x)$ may cause the variance of $\hat{\gamma}_{\mathrm{IS}}$ to be larger than the variance of the estimates obtained by direct simulation. For some $f^*(x)$, the variance may be infinite, and, hence, it is crucial to find a good $f^*(x)$. The *optimal* $f^*(x)$ is the one that minimizes the variance of $\hat{\gamma}_{\mathrm{IS}}$ in (3). The variance is minimized to $\mathrm{Var}[\hat{\gamma}_{IS}] = 0$ when $g(x)f(x) = \gamma f^*(x)$. However, this requires knowledge of γ, the property of interest. In any case, this observation may serve as a guideline stating that $g(x)f(x) \propto f^*(x)$. This makes efficient application of importance sampling very model specific. Two general guidelines can be given:

- if $g(x)f(x) > 0$ then $f^*(x) > 0$ (from (1)),
- $g(x)f(x) = \gamma f^*(x)$ which implies that the sampling should be in proportion to the original importance and likelihood of the samples, i.e $g(x)f(x) \propto f^*(x)$ (from (3)).

Because the efficiency of importance sampling is observed to be rather sensitive to the change of measure, a great deal of work is being done to obtain optimal – or at least good – simulation parameters. In Section 1, a few comments on this work are given. For an extensive and excellent survey on selecting $f^*(x)$ the reader is referred to [24].

4 IMPORTANCE SPLITTING – RESTART/LRE

The RESTART method (REpetitive Simulation Trials After Reaching Threshold) [53] [54] [51] [29] [20] [13] belongs to the so-called importance splitting methods [48] which allow the evaluation of extremely low probabilities, e.g., 10^{-10} or 10^{-25}, by simulation. This type of rare event simulation was proposed in 1970 [5] [27] and has been refined in the context of ATM performance evaluation studies. The method is based on *restarting* the simulation in valid states of the system that provoke the events of interest more often than in a straightforward simulation, thus leading to a shorter simulation run time. The results of the simulation have to be modified to reflect the correct weights of the events. Applications of this approach and the so-called importance sampling approach are the subject of a special issue on rare event simulation [18].

Combining the RESTART method with the LRE algorithm (Limited Relative Error) [46] [47] for the statistical evaluation of the distribution function yields the stationary complementary distribution function[6] $G(x)$ of the investigated random x-sequence and also the so-called *local correlation function* $\varrho(x)$, which represents relevant correlation evidence to be included in the error measure for controlling the simulation run time. For discrete random sequences like the arrival occupancy this procedure for automatically performing this general, rather complicated evaluation task can be substantially simplified by applying a simplified version of the LRE algorithm (LRE algorithm III [47]). The evaluation task of this algorithm has been restricted to the measurement of those discrete random sequences. The basic idea of this algorithm maps the values that are smaller than a given x-value and the values that are greater than this value to a 2-node Markov chain and defines the first order correlation coefficient of this Markov chain as the local correlation function $\rho(x)$ which depends on the location x.

The random variables investigated in this paper are the arrival occupancy α of a queuing system within a queuing network and the delay time τ_D of a section of a queuing network. For discrete sequences the complementary distribution function is defined as

$$G(x) = G_i = \mathsf{P}\{\alpha \geq i\} \qquad \text{for} \quad i - 1 \leq x < i\,. \tag{4}$$

The complementary distribution function of the arrival occupancy G_i ($i = 0, \ldots, B$) contains the loss probability as the last value G_B, i.e., the probability that an arrival sees the system with the maximum of B jobs or cells. For the delay time τ_D the complementary distribution function is defined as $G_D(x) = \mathsf{P}\{\tau_D \geq x\}$.

Different run time strategies are possible for this method. Figure 2 shows the step-by-step or single-step approach where each step is evaluated one after the other until the requested accuracy is reached. During each step, those states of the system that can be used for *restarting* the simulation in the next step are saved. Within one step, only the next step (threshold) needs to be known. Subsequent thresholds can be determined using the results of the simulation.

In the global-step approach, all steps are simulated simultaneously, see figure 4 and [16]. The main advantage of this approach is that only a maximum

[6]The LRE algorithm can also evaluate the distribution function $F(x)$.

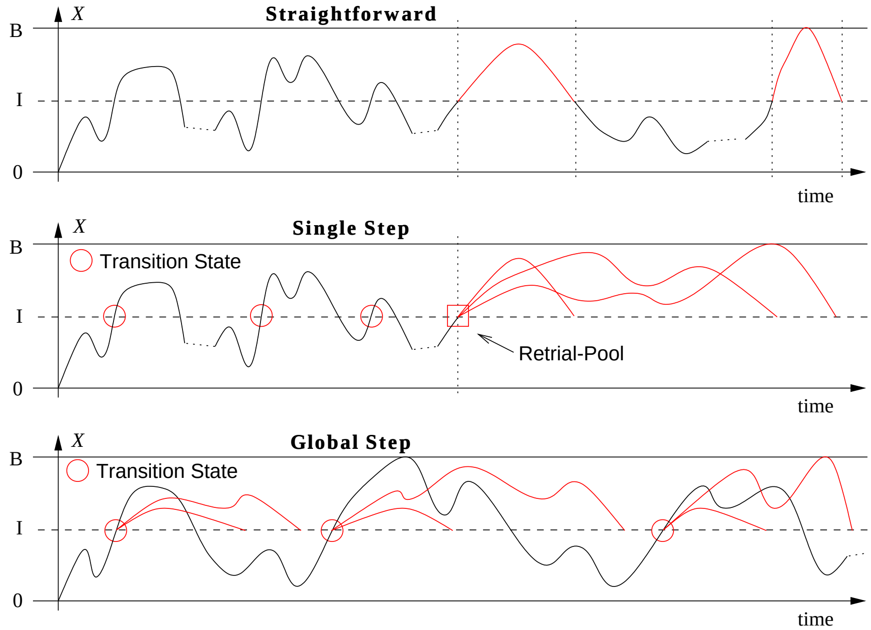


Figure 2: Importance Splitting with RESTART

of m states – one state per threshold – needs to be stored at any one time. The disadvantage is that all the thresholds need to be defined in advance.

Combinations of the two methods lead to a very efficient approach, where the parameters are determined with the step-by-step approach and the desired accuracy is gained with the global-step approach. Figure 5 shows a validation of the method for an M/M/1 system. The method has also been applied very successfully to a slightly modified cell loss simulation during handover within a wireless ATM system. Cell loss has a different meaning in this context, as all cells of the reference connection in the queue at the time of the handover are lost.

5 RESTART/LRE FOR QUEUING NETWORKS

The method can easily be extended to queuing networks as shown in figure 6 by choosing one node to be evaluated and combing the remaining parts of the system as a so-called *rest system*, which also contains the sources of the system. Saving a system state now means saving the state of all sources, all queues, and all servers including the rest service time of the current job being served. The chosen node with index r controls the simulation. To evaluate the complete network one RESTART/LRE simulation for each node has to be performed. This is a feasible procedure given the enormous speed-up of the

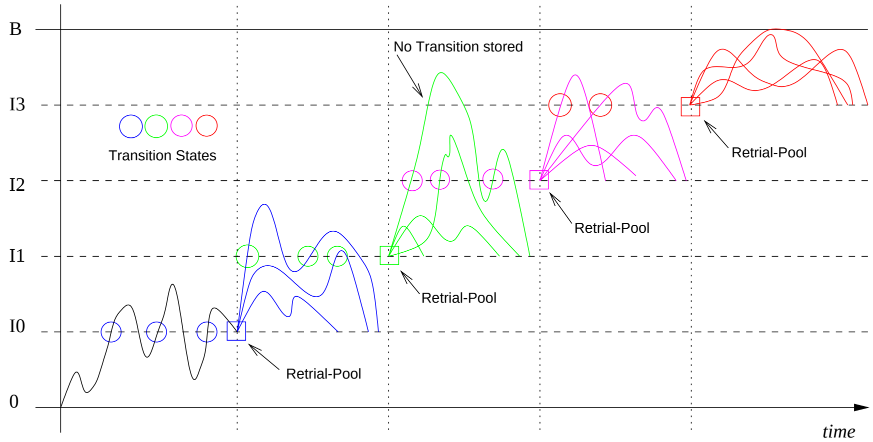
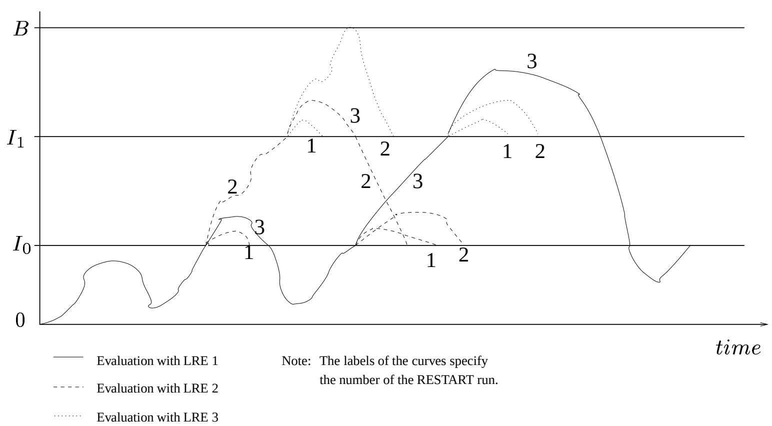


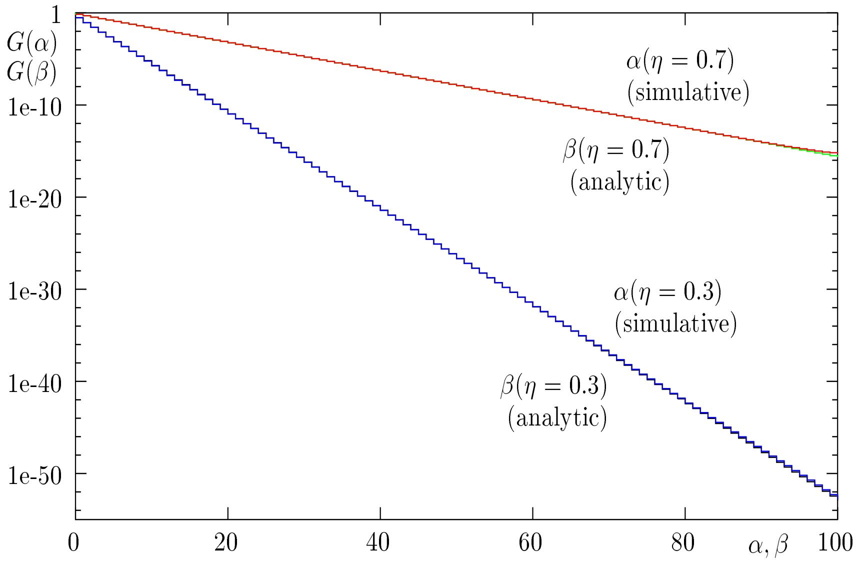
Figure 3: Step-by-Step RESTART



Splitting Factor $R_i = 3, i = 1, 2$

The label of the curve specifies the number of the RESTART run.

Figure 4: Global-Step RESTART



Simple M/M/1 model
α: arrival occupancy, β: overall time weighted occupancy

Figure 5: RESTART/LRE Example: simulated vs. analytic results

method.

6 RESTART/LRE: BASIC FORMULAS

Measuring the complementary distribution function $G_i = \mathsf{P}\{\alpha \geq i\}$ of the arrival occupancy α and the conditional complementary distribution functions $\mathsf{P}\{\alpha \geq I_j | \alpha \geq I_{j-1}\}$ during the simulation allows the calculation of the loss probability from the results of each step by using the following formula, see [16] [20]:

$$\begin{aligned} G_B = \mathsf{P}\{\alpha \geq B\} = \\ \mathsf{P}\{\alpha \geq B | \alpha \geq I_{m-1}\}\mathsf{P}\{\alpha \geq I_{m-1} | \alpha \geq I_{m-2}\} \\ \cdots \mathsf{P}\{\alpha \geq I_1 | \alpha \geq I_0\}\mathsf{P}\{\alpha \geq I_0\}\,. \end{aligned} \tag{5}$$

The RESTART/LRE simulations for the delay measurements use the following approach to calculate the complementary distribution function:

$$\begin{aligned} G_D(x) = \mathsf{P}\{\tau_D \geq x\} = \\ \mathsf{P}\{\tau_D \geq x | \tau_D \geq I_{m-1}\}\mathsf{P}\{\tau_D \geq I_{m-1} | \tau_D \geq I_{m-2}\} \\ \cdots \mathsf{P}\{\tau_D \geq I_1 | \tau_D \geq I_0\}\mathsf{P}\{\tau_D \geq I_0\}\,. \end{aligned} \tag{6}$$

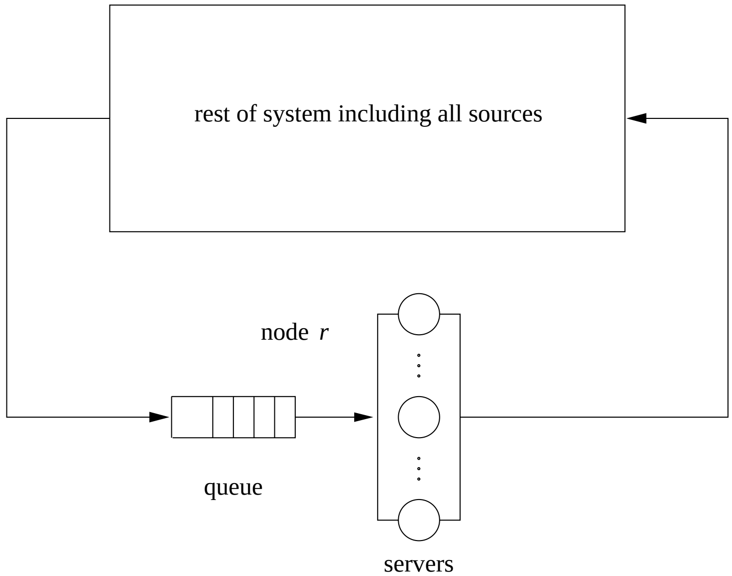


Figure 6: Queueing network

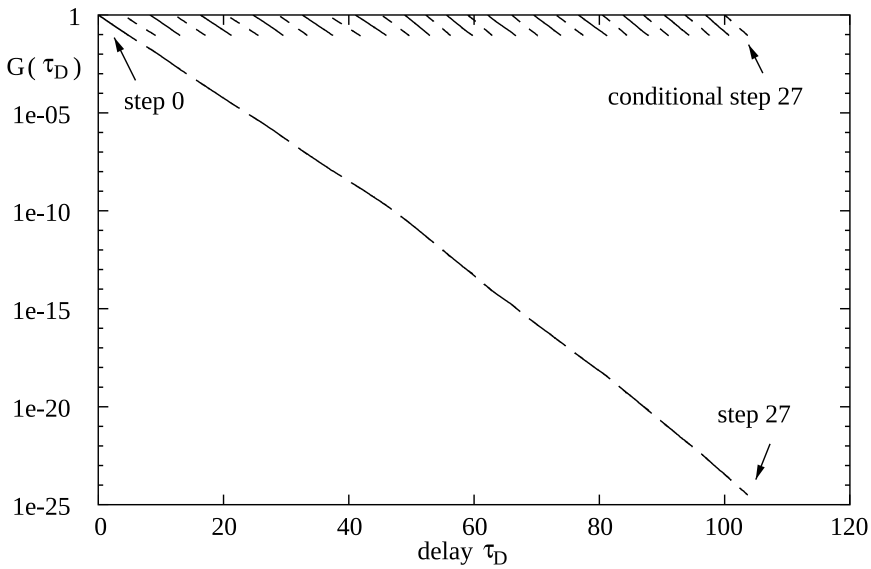
In the current implementation the delay is measured from the point of arrival to a specified point in the connection. This means that delays are measured which involve several nodes. The FIFO queuing discipline is responsible for the correlation of the delay times that is exploited by the RESTART mechanism, which defines delay thresholds. These methods are all part of the RESTART/LRE simulation toolkit MuSICS[7].

There are certain differences between measuring a discrete random variable like the arrival occupancy α and a continuous random variable like the delay time τ_D. In the context of rare event importance splitting methods the thresholds for the delays need to be seen as a relation since the next delay can assume any positive value, whereas the arrival occupancy can increase only by one. Although, generalizing this to bulk arrivals leads to the same type of relation also for the arrival occupancy. The near optimal thresholds can be found in the same way with the step-by-step method [16].

The possibilities of the method for the evaluation of delay time distributions is shown by an M/M/1/100 example in figure 7. For this example a probability of 10^{-25} could be verified by simulation. The example also demonstrates the method showing the conditional complementary distribution functions. This explains the speed-up of the method as the simulation itself is carried out in each step at a much higher level of probabilities.

For a discussion of the parameters of the RESTART/LRE algorithm see the

[7]**Mu**lti **S**tep **I**mportance **S**plitting **C**lass library **S**ystem



LRE Error $d = 3\%$, offered load $\rho = 0.5$

Figure 7: M/M/1/N=100: Complementary Distribution Function $G(\tau_D)$

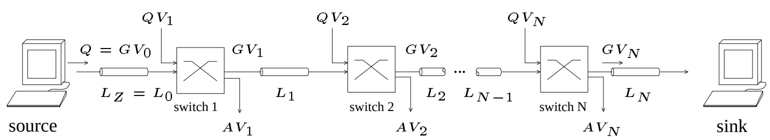


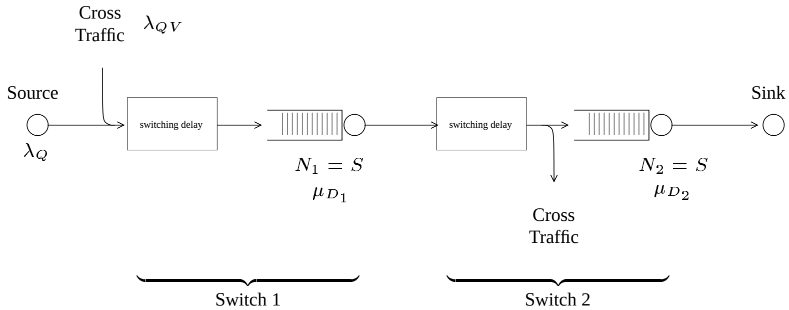
Figure 8: ATM Reference Connection: Model

appendix and [16].

7 ATM REFERENCE CONNECTION

7.1 Models

Figure 8 shows the principle structure of an ATM connection. The cells are transferred from the source to the sink via the switches $1...N$. Besides the reference traffic Q the additional cross traffic ($QV_1 \ldots QV_N$) adds to the load of the switches and links ($L_0 \ldots L_N$). Part of the cross traffic ($AV_1 \ldots AV_N$) leaves the connection at each switch so that in general without making any restricting assumptions the traffic streams ($GV_1 \ldots GV_{N-1}$) are composed of the reference traffic and parts of the cross traffic. Only the last part of the connection is different, here only the reference traffic remains minus the cells that were lost during the transfer.



Traffic Model for a Connection with 2 Switches

Figure 9: ATM Reference Connection

This very general model is simplified for the evaluations in section 7.2 as described in the following.

The ATM switch can be viewed as a system composed of three subsystems:

- connection control (switching delay) can, e.g., be modeled by an *infinite server* with a constant delay [49] or by a single server queueing system. More measurements, e.g., also the coefficient of variation, are needed to refine this part of the model.
- output buffers (queuing delay)
  Due to the (constant) transmission delay, cells have to be buffered in a queue. The size of this queue determines the cell loss rate. Cells are served according to the FIFO principle.
- outgoing connection (transmission delay)
  Transmitting the cells on the outgoing connection line accounts for another delay[8] Δt_N of the cell which depends on the bandwidth of the connecting line, e.g., $\Delta t_N = 2.7\mu s$ for a 155 Mbit/s link. This corresponds to a deterministic server (D) as the cells have a constant length of 53 *byte*.

The next step simplifies the cross traffic. We assume that cross traffic cannot be split up. This means that the complete cross traffic that enters a node has to leave it at some point of the connection as one traffic stream.

In the examples the connections consist of 2 switches. The buffers have a capacity of $N_1 = N_2 = S = 100$ cells. A model of an ATM reference connection with these simplifications is given in figure 9. In section 7.2 the model used for the switching delay is a */M/1 queuing system.

[8] Propagation delay is not considered in the context of this paper.

The cells of the reference traffic as well as the cross traffic both use switch 1 and link 1. The cross traffic leaves the system after switch 2, which means that they are routed to another outgoing line of the switch.

This model allows the evaluation of the occupancy as well as the delay probabilities in form of the respective complementary distribution function.

7.2 Simulation Results

Variables investigated for the ATM reference connection include:

- arrival occupancy of the output buffer
- delay time measured before the cell enters the queue
- delay time measured after the cell has been sent out on the connection line

The RESTART/LRE simulation toolkit MuSICS allows the evaluation of these variables. Besides investigations of the occupancy and thus cell loss probabilities due to buffer overflow, delay measurements are needed to guarantee the correct quality of service for real time services.

The following parameters were used for the model:

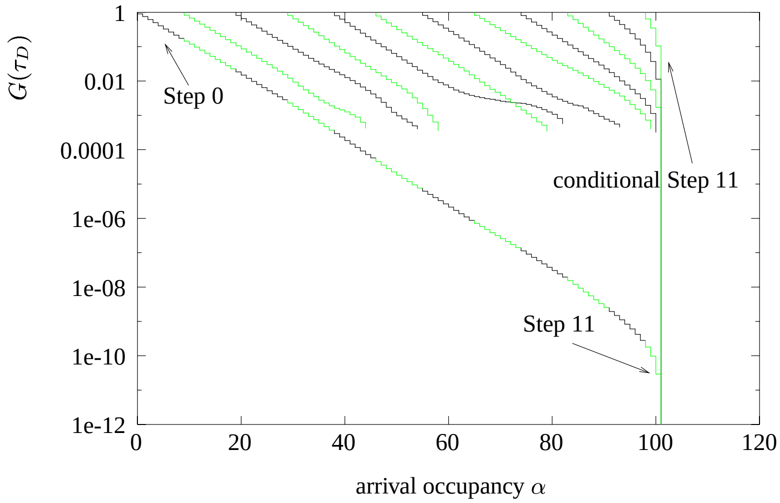
- neg. exp. arrival rates: λ_Q: reference traffic, λ_{QV_i}: cross traffic
- transmission delays (all links 155 *Mbit/s*) : $\Delta t_1 = \Delta t_2 = \Delta t_N = \frac{53 \cdot 8 bit}{155 \cdot 10^6 bit/s} = 2.74 \mu s$
- switching delays: $\tau_{S_1} = \tau_{S_2} = \tau_S$
- d: LRE error, usually values of 3%, 5%, and 10% have been used.

The load for the model is varied with the arrival rate of the reference traffic and the cross traffic.

An example experiment is shown in figures 10, 11, and 12. Figure 10 shows the occupancy of the output buffer of the first switch. This type of simulation also allows to quantify the loss probability, which for the first switch is $G_B = 2.9 \cdot 10^{-11}$ (with $B = 101$), the last value of the curve. The conditional complementary distribution functions in the upper part of the diagram shows that each simulation also produces results in the next interval, which are of course less accurate than the ones obtained in the following steps.

These simulations were performed with the step-by-step approach, which automatically finds the optimal position and number of steps.

The next two figures 11 and 12 show the complementary distribution function of the delay time $G(\tau_D)$ for the reference traffic stream and the cross



$\lambda_Q = 296736s^{-1}$, $\lambda_{QV} = 32906s^{-1}$, LRE error $d = 10\%$, $\bar{\tau}_S \approx 14\mu s$

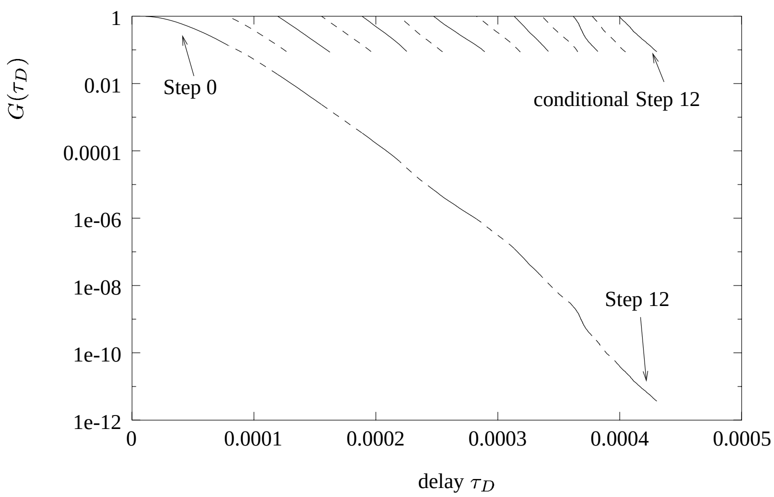
Figure 10: CDF $G(\alpha)$: Output Buffer Switch 1

traffic respectively. It should be pointed out that a straightforward simulation of all these models with results in these regions of very rare events would have taken years whereas all these simulations were performed within less than an hour on a standard workstation.

The complementary distribution function (CDF) now allows to quantify statements about the probability of exceeding given delay limits. Assuming that the maximum delay τ_{max} is defined by the value for which the exceeding probability is 10^{-8} then $\tau_{max} = 340\mu s$ in figure 11. The cell delay variation (CDV) can then be calculated as the difference between τ_{max} and $\tau_{min} = 10.48\mu s$.

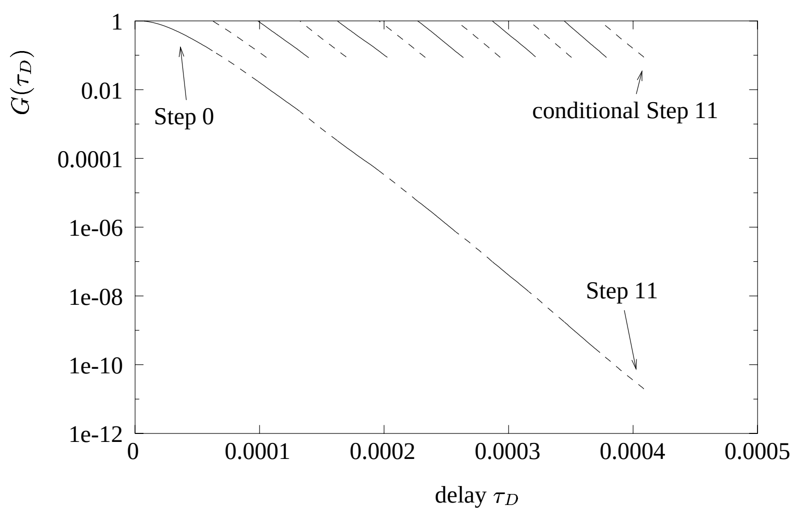
8 CONCLUSION

The methods used for rare event simulation have progressed enormously in the last few years. They have all been successfully used for small to medium sized problems. Future applications include the performance study of IP based networks with QoS support, e.g., differentiated services, in the context of fixed and mobile networks for existing and new (end-to-end and access) services. Further extensions are needed to employ rare event simulation methods in a more general way and especially to integrate them into tools for standard use.



$\lambda_Q = 296736s^{-1}$, $\lambda_{QV} = 32906s^{-1}$, LRE error $d = 3\%$, $\bar{\tau}_S \approx 14\mu s$

Figure 11: CDF $G(\tau_D)$: Reference Traffic Delay τ_D – Switch 2



$\lambda_Q = 296736s^{-1}$, $\lambda_{QV} = 32906s^{-1}$, LRE error $d = 10\%$, $\bar{\tau}_S \approx 14\mu s$

Figure 12: CDF $G(\tau_D)$: Cross Traffic Delay τ_D – Switch 2

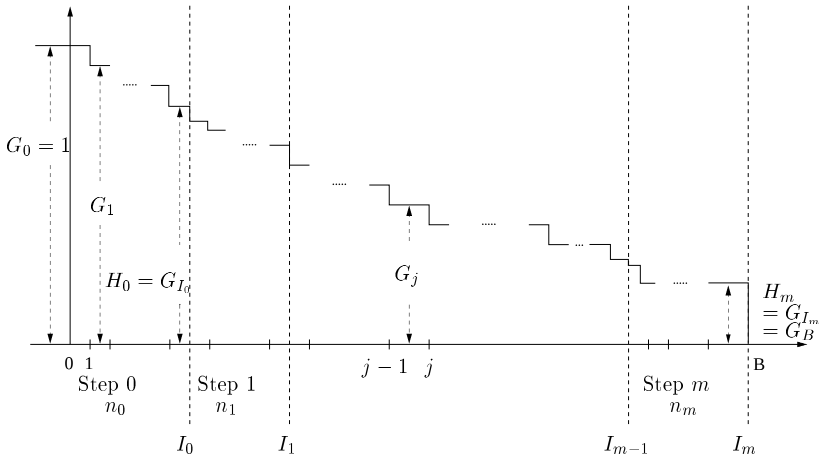


Figure 13: Partitioning of the CDF due to multiple thresholds

ACKNOWLEDGEMENT

The authors gratefully acknowledge the support of this work by Prof. Dr.-Ing. F. Schreiber and Prof. Dr.-Ing. B. Walke, Aachen University of Technology, and the cooperation with Dr. Victor Nicola, University of Twente, Enschede, The Netherlands, and Prof. Dr. Reuven Rubinstein, Technion, Haifa, Israel.

APPENDIX

This appendix gives a summary of results for the step-by-step [15] [16] and global-step strategy [54].

Figure 13 shows how the complementary distribution function is partitioned due to the thresholds. The I_x show the location of the threshold within the random variable and the G_x the corresponding value of the complementary distribution function.

A STEP-BY-STEP RESTART/LRE

In this section, the results for the optimal parameter settings for the minimal number of trials of the step-by-step strategy are given. The notation given in Table 1 is used for splitting the complementary distribution function $G(x)$ into $m + 1$ intervals.

The loss probability is calculated from the results of each step by using the

Table 1: Notation for the multi-step RESTART/LRE and result variables

m	number of steps or thresholds (0: no RESTART, 1: single step RESTART)
n_i	number of trials $n_i = \frac{\gamma_i}{d_S^2 H_{i\mid\alpha\geq I_{i-1}}}$ in step i, $i = 0, \ldots, m$, i.e. m-step RESTART: $m+1$ runs
I_i	interval limits of the thresholds for $i = 0, \ldots, m$ with $I_{-1} = 0$ and $I_m = B$
R_i	splitting factor; number of *restarts* from one system state
G_B, $\tilde{G}_B$	loss probability; G_B: analytical value, $\tilde{G}_B$: value obtained by simulation
$H_i = G_{I_i}$	level of the compl. i for $i = 0, \ldots, m$ with $H_{-1} = 1$ and $H_m = G_B$
P_i	conditional level probabilities $P_i = H_{i\mid\alpha\geq I_{i-1}}$
ρ_i	local correlation $\rho_i = 1 - \frac{\mathsf{P}\{\alpha_n \geq i, \alpha_{n+1} < i\}}{\mathsf{P}\{\alpha_n < i\}(1 - \mathsf{P}\{\alpha_n < i\})}$ for $i = 1, \ldots, B$, see [47]
	α_n is the n^{th} value of the stochastic process for the arrival occupancy α
$cf_i = \frac{1+\rho_i}{1-\rho_i}$	correlation factor for $i = 1, \ldots, B$
$\gamma_i = cf_{I_i}$	threshold correlation factor for $i = 0, \ldots, m$
$d_B = d_{max}$, d_S	d_B: total maximum LRE error; d_S: maximum LRE error in each step

following formula:

$$\begin{aligned} G_B = \mathsf{P}\{\alpha \geq B\} = \\ \mathsf{P}\{\alpha \geq B | \alpha \geq I_{m-1}\}\mathsf{P}\{\alpha \geq I_{m-1} | \alpha \geq I_{m-2}\} \\ \cdots \mathsf{P}\{\alpha \geq I_1 | \alpha \geq I_0\}\mathsf{P}\{\alpha \geq I_0\}\,. \end{aligned} \tag{7}$$

For this method, the number of trials in each step n_i can be calculated by the LRE formula, see Table 1 and [47]. Taking the sum over all steps and then successively using the necessary condition for a minimum, the optimal levels H_i^* and the minimal number of trials n_{min} can be determined as functions of the loss probability $H_m = G_B$:

$$H_i^*(H_m) = \left(\frac{\prod_{j=0}^{i} \gamma_j^{m-i}}{\prod_{j=i+1}^{m} \gamma_j^{i+1}} H_m^{i+1} \right)^{\frac{1}{m+1}}, \tag{8}$$

$$n_{min}(H_m) = \frac{m+1}{d_S^2} \left(\prod_{i=0}^{m} \gamma_i \right)^{\frac{1}{m+1}} \left(\frac{1}{H_m} \right)^{\frac{1}{m+1}}. \tag{9}$$

The error in m steps ($m+1$ intervals) is:

$$d_B^2 = (d_S^2 + 1)^{m+1} - 1 \rightarrow d_S^2 = (d_B^2 + 1)^{\frac{1}{m+1}} - 1\,. \tag{10}$$

Trying to determine the optimal number of steps m_{opt} using (10) leads to transcendental equations that cannot be solved. The following approximation leads to a satisfactory solution:

$$\left(\prod_{i=0}^{m} \gamma_i \right)^{\frac{1}{m+1}} \approx (\gamma^m \gamma_m)^{\frac{1}{m+1}}\,. \tag{11}$$

The reasoning behind this approximation is the fact that the values of γ_i as well as ρ_i form a maximum curve [47] [20] so that it is feasible to approximate all γ_i except for γ_m by the constant $\gamma = cf_M = \gamma_{max}$:

$$\gamma_0 \approx \gamma_1 \approx \gamma_2 \approx \cdots \approx \gamma_{m-1} \approx \gamma = cf_M = \gamma_{max}\,. \tag{12}$$

The resulting approximation for the number of trials n_z^* depends on the maximum total error d_B, the number of steps m, and the loss probability $H_m = G_B$:

$$n_z^*(H_m) = \frac{(m+1)\left(\frac{\gamma^m \gamma_m}{H_m}\right)^{\frac{1}{m+1}}}{(d_B^2+1)^{\frac{1}{m+1}} - 1}\,. \tag{13}$$

The approximation can be checked for specific examples using the exact formula given in [15]. For the optimal levels H_i^* the approximation yields:

$$H_i^*(H_m) \approx \left(\frac{\gamma}{\gamma_m} H_m\right)^{\frac{i+1}{m+1}} \tag{14}$$

and thus

$$H_i^*(G_B) \approx \left(\frac{\gamma}{\gamma_m} G_B\right)^{\frac{i+1}{m+1}} \ . \tag{15}$$

The approximation of the optimal H-levels of the step depends only on G_B, m, i, and the relation of γ to γ_m. The interval limits I_i are indirectly given by the following equation:

$$H_i^*(G_B) = G_{I_i} \ . \tag{16}$$

With another approximation of the error using a Taylor series:

$$d_S^2 = (d_B^2 + 1)^{\frac{1}{m+1}} - 1 \approx \frac{d_B^2}{m+1} \tag{17}$$

the approximate optimal number of trials n^* for m steps is:

$$n^*(m) = \frac{(m+1)^2 \left(\frac{\gamma^m \gamma_m}{H_m}\right)^{\frac{1}{m+1}}}{d_B^2} \ . \tag{18}$$

The necessary condition $\mathrm{d}n^*/\mathrm{d}m = 0$ leads to the solution of the optimal number of steps m_{opt}, where n^* is minimized:

$$m_{opt} = \frac{1}{2} \ln\left(\frac{\gamma_m}{\gamma H_m}\right) - 1 \ . \tag{19}$$

This is a minimum as the 2^{nd} derivative is positive at this point.

The result for the optimal number of trials (events) $n^*(m)$ using equation (18) m and can be approximated with the maximum value of $\gamma = cf_M = \gamma_{max}$. The other values needed are $\gamma_m = cf_B$ and $H_m = G_B$:

$$n^*(m) = \frac{(m+1)^2}{d_B^2} \left(\frac{cf_M{}^m cf_B}{G_B}\right)^{\frac{1}{m+1}} \ . \tag{20}$$

The optimal number of steps m_{opt} in relation to the number of trials given in (20) is:

$$m_{opt} = \frac{1}{2} \ln(\frac{cf_B}{cf_M G_B}) - 1 \tag{21}$$

As the number of steps is an integer between 0 and the maximum $B-1$, the value of m_{opt} has to be rounded and checked by comparing it against the values for 0 and $B-1$:

$$m^* = \begin{cases} 0 & \text{for} \quad m_{opt} < 0\,, \\ \mathsf{round}(m_{opt}) & \text{for} \quad 0 \leq m_{opt} \leq B-1\,, \\ B-1 & \text{for} \quad m_{opt} > B-1\,. \end{cases} \tag{22}$$

An upper resp. lower bound for the number of trials can be found by using the maximum γ_{max} resp. minimum γ_{min} value for γ in equation (13) resp. (18).

Taking into account the typical shape of a maximum curve for the local correlation function ρ_i and also for cf_i (resp. $\gamma_j = cf_{I_j}$), ρ_{max} resp. γ_{max} already yield a very good approximation. A validation is given in [15].

Using the results given in (15) and (19), the conditional probabilities $P_i = H_{i|\alpha \geq I_{i-1}}$ can be shown to evaluate to $P_i = e^{-2}$. This is the same result as given in [54] for the approximate solution (see also next section). The implementation of the step-by-step approach within MuSICS allows two modes of operation: either the thresholds and the maximum overall LRE error can be given as parameters or an automatic version of the method can be used, which successively determines the next threshold using previous information. For the automatic method the simulation is started with a maximum LRE error d_S for each step. In each step the procedure then is the following: increase the given error limit, simulate until the threshold is known, set the error limit to d_S, simulate and store transition states for the next threshold until d_S is satisfied up to the threshold.

B THE GLOBAL-STEP APPROACH

In [54], the results for the optimal gain of the global-step approach are given using variables s_i to describe the statistical behaviour and y_i to describe the restoration costs. s and y are the corresponding geometric means.

The optimal number of *restarts* or splitting factor R_i for each point is:

$$R_i = \frac{1}{(s \cdot y \cdot G_B)^{\frac{1}{m+1}}} \frac{sy}{s_i \cdot y_i} \text{ for } i = 1 \ldots m\,. \tag{23}$$

The optimal levels are given as the conditional probabilities $P_i = H_{i|\alpha \geq I_{i-1}}$:

$$P_1 = (s \cdot y \cdot G_B)^{\frac{1}{m+1}} \frac{s_1}{s \cdot y}\,, \tag{24}$$

$$P_i = (s \cdot y \cdot G_B)^{\frac{1}{m+1}} \frac{s_1 \cdot y_{i-1}}{s \cdot y} \text{ for } 2 \leq i \leq m\,, \tag{25}$$

$$P_m = (s \cdot y \cdot G_B)^{\frac{1}{m+1}} \frac{y_m}{s \cdot y}\,. \tag{26}$$

The gain g_m is defined as the ratio of a metric (simulation costs multiplied with the variance of the estimated value) with and without RESTART. Using the optimal P_i and R_i for a fixed value of m the gain is given by:

$$g_m = \frac{1}{(m+1)^2} \cdot \frac{1}{(s \cdot y \cdot G_B)^{\frac{m}{m+1}}}\,. \tag{27}$$

The optimal number of steps follows as:

$$m_{opt} = -\frac{\ln(s \cdot y \cdot G_B)}{2} - 1 \tag{28}$$

and the optimal gain using the optimal number of steps:

$$g_{opt} = \frac{4}{s \cdot y \cdot G_B[e \cdot \ln(s \cdot y \cdot G_B)]^2}\,. \tag{29}$$

Using approximations for s_i (homogeneous behaviour $s_i = s$) and ignoring the restoration costs y_i ($y_i = y = 1$) and not taking the integer part of m_{opt} the following optimal values are obtained:

$$P_i = e^{-2}\,, \tag{30}$$

$$R_i = e^2\,. \tag{31}$$

Comparing (28) $s \approx \frac{cf_M}{cf_B}$, which remains to be validated.

References

[1] A. A. Akyamaç, Z. Haraszti, J. K. Townsend. *Efficient Rare Event Simulation Using DPR for Multidimensional Parameter Spaces*. Proc. 16th Int. Teletraffic Congress, 1999.

[2] R. Andreassen, P.E. Heegaard, B.E. Helvik. *Importance Sampling for speed-up simulation of heterogeneous MPEG sources*. In *The 13th Nordic Teletraffic Seminar (NTS-13), August, 1996,* pp. 190–203, Trondheim, Norway, 1996.

[3] S. Asmussen, R.Y. Rubinstein. *Steady State Rare Event Simulation in Queueing Model and its Complexity Properties*. In J. Dshalalow, editor, *Advances in Queueing: Theory, Methods and Open Problems*, Vol. I, pp. 429–462. CRC Press, 1995.

[4] Søren Asmussen. *Rare events simulation for heavy–tailed distributions*. Workshop 11.-12. March 1999, University of Twente, Enschede, The Netherlands, 1999.

[5] A.J. Bayes. *Statistical Techniques for Simulation Models*. The Australian Computer Journal, Vol. 2(4), pp. 180–184, 1970.

[6] N. K. Boots. *Quick Simulation Methods for Estimating Ruin Probabilities in Risk Processes with Subexponential Claims*. Workshop 11.-12. March 1999, University of Twente, Enschede, The Netherlands, 1999.

[7] Pieter-Tjerk de Boer. *Ananlysis and efficient simulation of queueuing models of telecommunication systems*. PhD thesis, Universiteit Twente, 2000.

[8] Pieter-Tjerk de Boer, Victor F. Nicola. *Hybrid Importance Sampling Estimation of Consecutive Cell Loss Probability*. International Journal of Electronics and Communications, Vol. 52, No. 3, pp. 133–140, 1998.

[9] M. Devetsikiotis, J.K. Townsend. *Statistical optimization of dynamic importance sampling parameters for efficient simulation of communication networks*. IEEE/ACM Transactions on Networking, Vol. 1, No. 3, pp. 293–305, June 1993.

[10] M.R. Frater. *Estimation of the Statistics of Rare Events in Data Communications Systems*. PhD thesis, The Australian National University, 1990.

[11] Oliver J. Fuß. *Validierung von Dienstgüteparametern von B-ISDN/ATM-Referenzverbindungen durch das RESTART/LRE-Verfahren zur Simulation seltener Ereignisse*. Diplomarbeit, RWTH Aachen, Lehrstuhl Kommunikationsnetze, Januar 1998.

[12] Marnix J.J. Garvels. *The splitting method in rare event simulation*. PhD thesis, Universiteit Twente, 2000.

[13] P. Glasserman, P. Heidelberger, P. Shahabuddin, T. Zajic. *Splitting for Rare Event Simulation: Analysis of Simple Cases*. In *1996 Winter Simulation Conference*, pp. 302–308, Coronado, California, USA, December 1996.

[14] A. Goyal, P. Shahabuddin, P. Heidelberger, V.F. Nicola, P.W. Glynn. *A Unified Framework for Simulating Markovian Models of Highly Dependable Systems*. IEEE Trans. on Computers, Vol. 41, No. 1, pp. 36–51, 1992.

[15] C. Görg. *Verkehrstheoretische Modelle und Stochastische Simulationstechniken zur Leistungsanalyse von Kommunikationsnetzen*, Vol. ABMT 13. Verlag der Augustinusbuchhandlung, Aachener Beiträge zur Mobil- und Telekommunikation, Aachen, 1997. Habilitationsschrift, Lehrstuhl Kommunikationsnetze, RWTH Aachen.

[16] C. Görg, O. Fuß. *Comparison and Optimization of RESTART Run Time Strategies*. AEÜ, Vol. 52, pp. 197–204, 1998.

[17] C. Görg, O. Fuß. *Simulating Rare Event Details of ATM Delay Time Distributions with RESTART/LRE*. Workshop 11.-12. March 1999, University of Twente, Enschede, The Netherlands, 1999.

[18] C. Görg, C. Kelling. *Special Issue on Rare Event Simulation*. Görg, C. and Kelling, Chr. (eds) AEÜ, Vol. 52, 1998.

[19] C. Görg, E. Lamers, R.G. Addie. *Broadband Traffic Modeling: Rare Event Simulation for Gaussian and M/Pareto Processes with the same Autocovariance*. Workshop 11.-12. March 1999, University of Twente, Enschede, The Netherlands, 1999.

[20] C. Görg, F. Schreiber. *The RESTART/LRE Method for Rare Event Simulation*. In *1996 Winter Simulation Conference*, pp. 390–397, Coronado, California, USA, December 1996.

[21] Z. Haraszti. *How to set up my simulator? - empirical guidelines*. Impromptu Session 16th International Teletraffic Congress, 7-10 June 1999, Edinburgh, UK.

[22] P.E. Heegaard. *Efficient Simulation of Network Performance by Importance Sampling*. PhD thesis, NTNU - Norwegian Technical and Scientific University, Trondheim, Norway, 1998.

[23] Poul E. Heegaard. *A Scheme for Adaptive Biasing in Importance Sampling*. International Journal of Electronics and Communications, Vol. 52, No. 3, pp. 172–182, 1998.

[24] P. Heidelberger. *Fast Simulation of Rare Events in Queueing and Reliability Models*. ACM Transactions on Modeling and Computer Simulation, Vol. 5(1), pp. 43–85, 1995.

[25] P. Heidelberger, P. Shahabuddin, V.F. Nicola. *Bounded Relative Error in Estimating Transient Measures of Highly Dependable Non-Markovian Systems*. In S. Ozekici, editor, *Reliability and Maintenance of Complex Systems*, Vol. 154 of *F: Computer and Systems Sciences*, pp. 487–515. Springer, 1996.

[26] B.E. Helvik, P.E. Heegaard. *A Technique for Measuring Rare Cell Losses in ATM Systems*. Teletraffic and Datatraffic, Proceedings 14th ITC, Antibes, Juan-Les-Pins, France, June 6-10, 1994, pp. 917–930, 1994.

[27] A.C.M. Hopmans, J.P.C. Kleijnen. *Importance Sampling in Systems Simulation: a Practical Failure?* In *Mathematics and Computers in Simulation XXI*, pp. 209–220, North-Holland, 1979.

[28] H. Kahn, A.W. Marshall. *Methods of Reducing Sample Size in Monte Carlo Computations*. J. Oper. Res. Society of America, Vol. 1, pp. 263–278, 1953.

[29] C. Kelling. *A Framework for Rare Event Simulation of Stochastic Petri Nets using "RESTART"*. In *1996 Winter Simulation Conference*, pp. 317–324, Coronado, California, USA, December 1996.

[30] Dirk P. Kroese, Victor F. Nicola. *Efficient Simulation of Backlogs in Fluid Flow Lines*. International Journal of Electronics and Communications, Vol. 52, No. 3, pp. 165–171, 1998.

[31] Thomas Kuhlmann, Christian Kelling. *Case Studies on Multidimensional RESTART Simulations*. International Journal of Electronics and Communications, Vol. 52, No. 3, pp. 190–196, 1998.

[32] Eugen Lamers. *Enhancement and Application of the RESTART/LRE-Simulator for Quality of Service Investigations in High Speed Networks*. Diploma thesis, Chair of Communication Networks, RWTH Aachen, June 1999.

[33] P. E. Lassila, J. T. Virtamo. *Efficient Importance Sampling for Monte Carlo Simulation of Loss Systems*. Proc. 16th Int. Teletraffic Congress, 1999.

[34] D. Lieber, R.Y. Rubinstein, D. Elmakis. *Quick Estimation of Rare Events in Stochastic Networks*. IEEE Transactions on Reliability Systems, Vol. 46, No. 2, pp. 254–265, 1997.

[35] Michel Mandjes. *Rare Event Analysis of Communication Networks*. PhD thesis, Vrije Universiteit Amsterdam, 1996.

[36] M.K. Nakayama. *A characterization of the simple failure biasing method for simulation of highly reliable Markovian systems*. ACM Transaction on Modeling and Computer Simulation, Vol. 4, No. 1, pp. 52–88, 1994.

[37] V.F. Nicola, G.A. Hagesteijn. *Efficient Simulation of Consecutive Cell Loss in ATM Networks*. Vol. II. Chapman and Hall, London, 1996.

[38] W. D. Obal II. *Importance Sampling Simulation in UltraSAN*. Simulation, Vol. 62(2), pp. 98–111, 1994.

[39] W. D. Obal II, W. H. Sanders. *Quick Simulation of a Performability Model*. In *Proc. 3rd. Intl. Workshop on Performability of Computer and Communication Systems*, pp. 6–10, Bloomingdale, IL, USA, 1996.

[40] S. Parekh, J. Walrand. *A Quick Simulation Method for Excessive Backlogs in Networks of Queues*. IEEE Transactions on Automatic Control, pp. 54–66, 1989.

[41] Ad Ridder. *Fast Simulation of Discrete Time Queues with Markov Modulated Batch Arrivals and Batch Departures*. International Journal of Electronics and Communications, Vol. 52, No. 3, pp. 127–132, 1998.

[42] Ad Ridder. *Efficient Simulation of Fluid Queues with Many Sources*. Workshop 11.-12. March 1999, University of Twente, Enschede, The Netherlands, 1999.

[43] Reuven Rubinstein. *The Cross-Entropy Method for Combinatorial and Continuous Optimization*. Methodology and Computing in Applied Probability, Vol. 1, pp. 127–190, 1999.

[44] Reuven Rubinstein. *Rare Event Simulation via Cross-Entropy and Importance Sampling*. Workshop 11.-12. March 1999, University of Twente, Enschede, The Netherlands, 1999.

[45] Reuven Rubinstein. *Cross-Entropy for Combinatorial Optimization*. to appear in Encyclopaedia of Management Sciences, 2000.

[46] F. Schreiber. *Effective control of simulation runs by a new evaluation algorithm for correlated random sequences*. AEÜ, Vol. 42, pp. 347–354, 1988. (see also: Proc. 12th Int. Teletraffic Congr. (ITC), Torino, 1988, p. 4.3B.1.1-9).

[47] F. Schreiber, C. Görg. *Stochastic Simulation: a Simplified LRE-Algorithm for Discrete Random Sequences*. AEÜ, Vol. 50, pp. 233–239, 1996.

[48] P. Shahabuddin. *Rare event simulation in stochastic models*. In *Proceedings of the 1995 Winter Simulation Conference*, pp. 178–185, Arlington, Virginia, USA, 1995.

[49] Zarah Siahpoureh. *Vergleich, Modellierung und Implementierung von simulativen und analytischen Modellen für ATM-Referenzverbindungen in Telekommunikationsnetzen*. Diplomarbeit, RWTH Aachen, Lehrstuhl für Kommunikationsnetze, February 1998.

[50] José Villén-Altamirano. *RESTART Method for the Case Where Rare Events Can Occur in Retrials From Any Threshold*. International Journal of Electronics and Communications, Vol. 52, No. 3, pp. 183–189, 1998.

[51] M. Villén-Altamirano, A. Martínez-Marrón, J. Gamo, F. Fernández-Cuesta. *Enhancement of the Accelerated Simulation Method RESTART by considering Multiple Thresholds*. In *Proceedings 14th International Teletraffic Congress*, pp. 797–810. North-Holland, 1994.

[52] M. Villén-Altamirano, J. Villén-Altamirano. *Accelerated simulation of rare events using RESTART method with hysteresis*. In J. Filipiak, editor, *Telecommunication Services for Developing Economies.*, pp. 675–686. Proceedings of the ITC Specialist Seminar, Cracow, Poland, April 1991, Elsevier, Amsterdam, Netherlands, 1991.

[53] M. Villén-Altamirano, J. Villén-Altamirano. *RESTART: a Method for Accelerating Rare Event Simulations*. In J.W. Cohen, C.D. Pack, editors, *Queueing, Performance and Control in ATM*, pp. 71–76. 13th International Teletraffic Congress, Copenhagen, North-Holland, 1991.

[54] M. Villén-Altamirano, J. Villén-Altamirano. *RESTART: A Straightforward Method for Fast Simulation of Rare Events*. In *Proceedings of the 1994 Winter Simulation Conference*, pp. 282–289, 1994.