

Rare Event Simulation

Estimation of rare event probabilities with the naive Monte Carlo techniques requires a prohibitively large number of trials in most interesting cases.

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 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 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 into a refined simulation technique under the name RESTART.

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.

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 $\gamma = E_f[g(X)]$. This can now be rewritten:

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

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

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

with variance

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

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.

The *optimal* $f(x)$ is the one that minimizes the variance. Two general guidelines can be given:

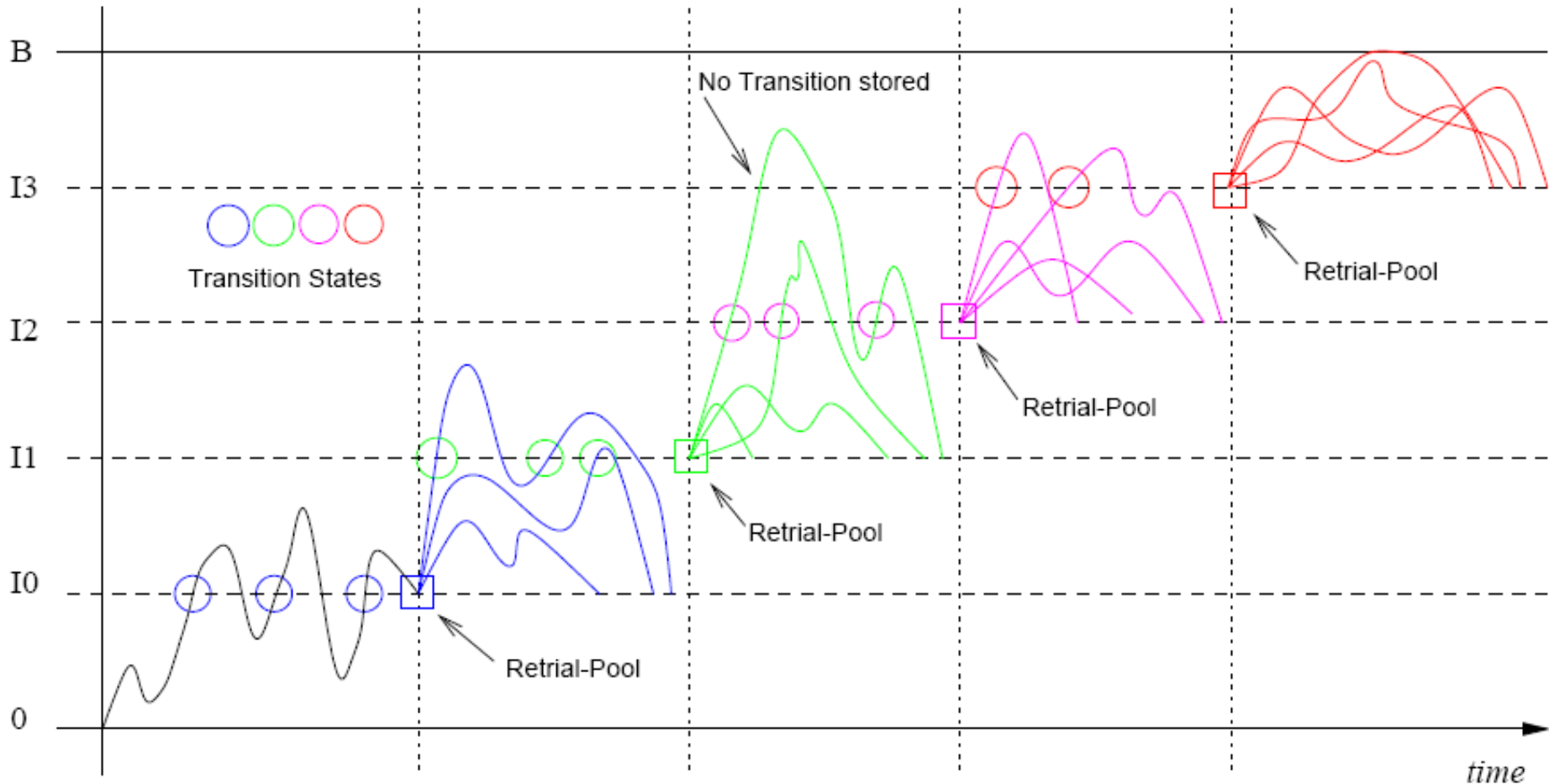
- if $g(x)f(x) > 0$ then $f^*(x) > 0$, otherwise $L(x)$ infinite
- $(g(X)L(X) - \gamma)$ must be small, $f^*(x) \approx g(x)f(x)/\gamma$.

Die RESTART-Methode

The RESTART (REpetitive Simulation Trials After Reaching Thresholds) method is a simple simulation method for the estimation of small probabilities. It was introduced in Villen-Altamirano (1991) and enhanced in Villen-Altamirano (1994), but it is similar to an older technique called *splitting* proposed in Kahn et al. (1951).

The basic idea of the RESTART method is to consider the rare event as the intersection of a nested sequence of events. The probability of the rare event is thus the product of conditional probabilities, each of which can usually be estimated much more accurately than the rare event itself, for a given simulation effort.

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



Step-by-Step RESTART

One of the main questions for any RESTART implementation is *how* and *when* to restart the simulation, in order to achieve the most accurate results for a fixed simulation effort.

RESTART for overflow probabilities:

The basic setting is the following: Consider a Markov process $\mathcal{X} := (X_t, t \geq 0)$ with space E , and let f be a real-valued function on E . Define $Z_t := f(X_t)$, for all $t \geq 0$. Assume for definiteness that $Z_0 \geq 0$. For any *threshold* or *level* $L > 0$, let T_L denote the first time that the process $Z := (Z_t, t \geq 0)$ hits the set $[L, \infty)$; and let T_0 denote the first time, after 0, that Z hits the set $(-\infty, 0]$. We assume that T_L and T_0 are well-defined (possibly infinite) stopping times with respect to the history of \mathcal{X} .

We are interested in the probability, γ say, of the event $D_L := \{-T_L < T_0\}$, i.e., the probability that Z up-crosses level L before it down-crosses level 0. Note that γ depends on the initial distribution of X .

An exact analysis of γ is often not possible. A standard way to estimate by simulation is the following. Generate independently r realizations (sample paths) of the Markov process X . Each path $x^{i\leftarrow} := (x_t^{i\leftarrow})$ defines a realization $z^{i\leftarrow} := (z_t^{i\leftarrow})$ of Z . Let I_i be the indicator that $z^{i\leftarrow}$ up-crosses level L before it down-crosses level 0. An unbiased estimate for γ is given by

$$\hat{\gamma} := \frac{1}{r} \sum_{i=1}^r I_i$$

For small values of γ this method is not very efficient.

We can see this by examining the *relative error* (RE) of the corresponding estimator which is defined as

$$\text{RE}(\hat{\gamma}) := \frac{\sqrt{\text{Var} \hat{\gamma}}}{\text{E} \hat{\gamma}} = \sqrt{\frac{1 - \gamma}{r \gamma}}.$$

Note that the relative error tends to infinity as γ tends to 0.

An alternative way to estimate γ is based on the following observation: If $L > K$ then $D_L \subset D_K$, where D_K denotes the event that Z up-crosses level K before it down-crosses level 0.

Therefore, we have by basic conditional probability,

$$\gamma = p_1 p_2$$

$$\text{with } p_1 := \text{P}(D_K) \quad \text{and} \quad p_2 := \text{P}(D_L | D_K).$$

Hence, if we estimate both p_1 and p_2 and multiply the results, we obtain an estimate for γ . When p_1 and p_2 are considerably larger than γ , this estimation procedure is likely to be more efficient than the standard method in the above equation.

Moreover, the same arguments may be used when we divide the interval $[0, Z]$ into *multiple* subintervals, instead of just two.

Fixed Effort RESTART

We partition the interval $[0, L)$ into m subintervals $[L_0, L_1)$, $[L_1, L_2)$, \dots , $[L_{m-1}, L_m)$, with $0 =: L_0 < L_1 < \dots < L_m := L$. Let D_z denote the event that process Z reaches level L_z before returning to 0. It is assumed that Z actually *hits* all thresholds $L_1 < \dots < L_m$ if event D_m occurs. Then $D_1 < D_2 < \dots < D_m$ is a nested sequence of events, decreasing to D_m . And, with $p_1 := P(D_1)$, $p_2 := P(D_2 | D_1)$, \dots , we have

$$\gamma = p_1 \cdot p_2 \cdot \dots \cdot p_m$$

We wish to estimate at the k th stage ($k = 1, \dots, m$) the conditional probability p_k . We do this by generating a *fixed* number of samples $I_1^{(k)}, \dots, I_{r_k}^{(k)}$ of the indicator that process Z reaches level L_k before returning to 0, starting from level L_{k-1} . We call r_k the simulation *effort* at stage k , and refer to this RESTART implementation as the *Fixed Effort* (FE) method.

Estimators of $p_1 \llcorner \llcorner \llcorner p_m$ are given by

$$\hat{p}_k := \frac{R_k}{r_k}, \quad k = 1, \dots, m, \quad \text{where} \quad R_k = \sum_{i=1}^{r_k} I_i^{(k)}$$

is the total number of successes at the k th stage. Moreover, the natural estimator of γ is

$$\hat{\gamma} := \prod_{k=1}^m \hat{p}_k.$$

Obviously these are unbiased estimators.

Fixed Splitting RESTART (more standard)

Every time when level L_k from L_{k-1} is reached, n_k , a fixed number, of samples in level L_k are simulated. Hence, the effort in level k , $r_k = n_k R_{k-1}$, is random.

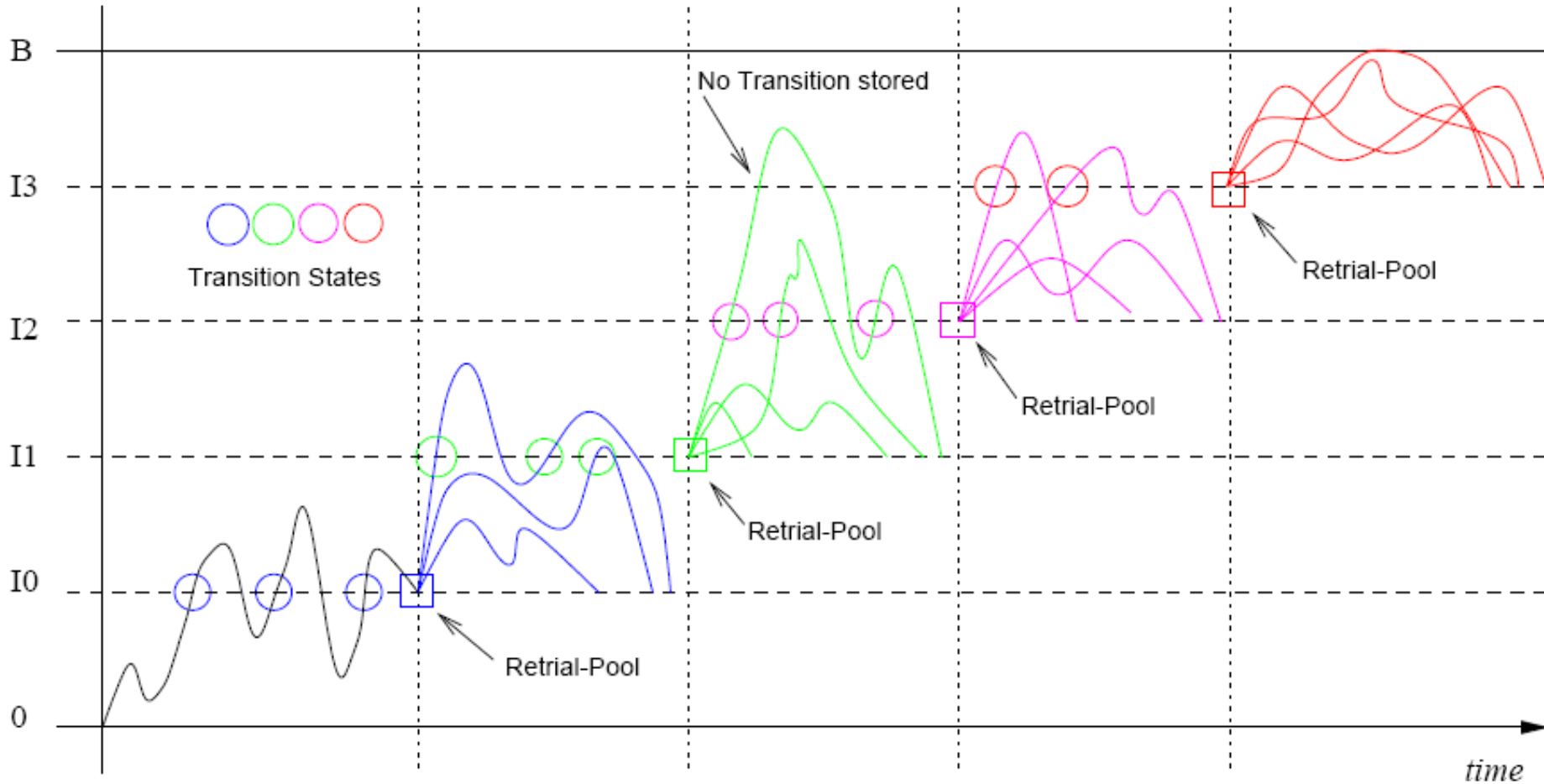
The estimator for γ is

$$\hat{\gamma} := \frac{R_m}{\prod_{k=1}^m n_k}$$

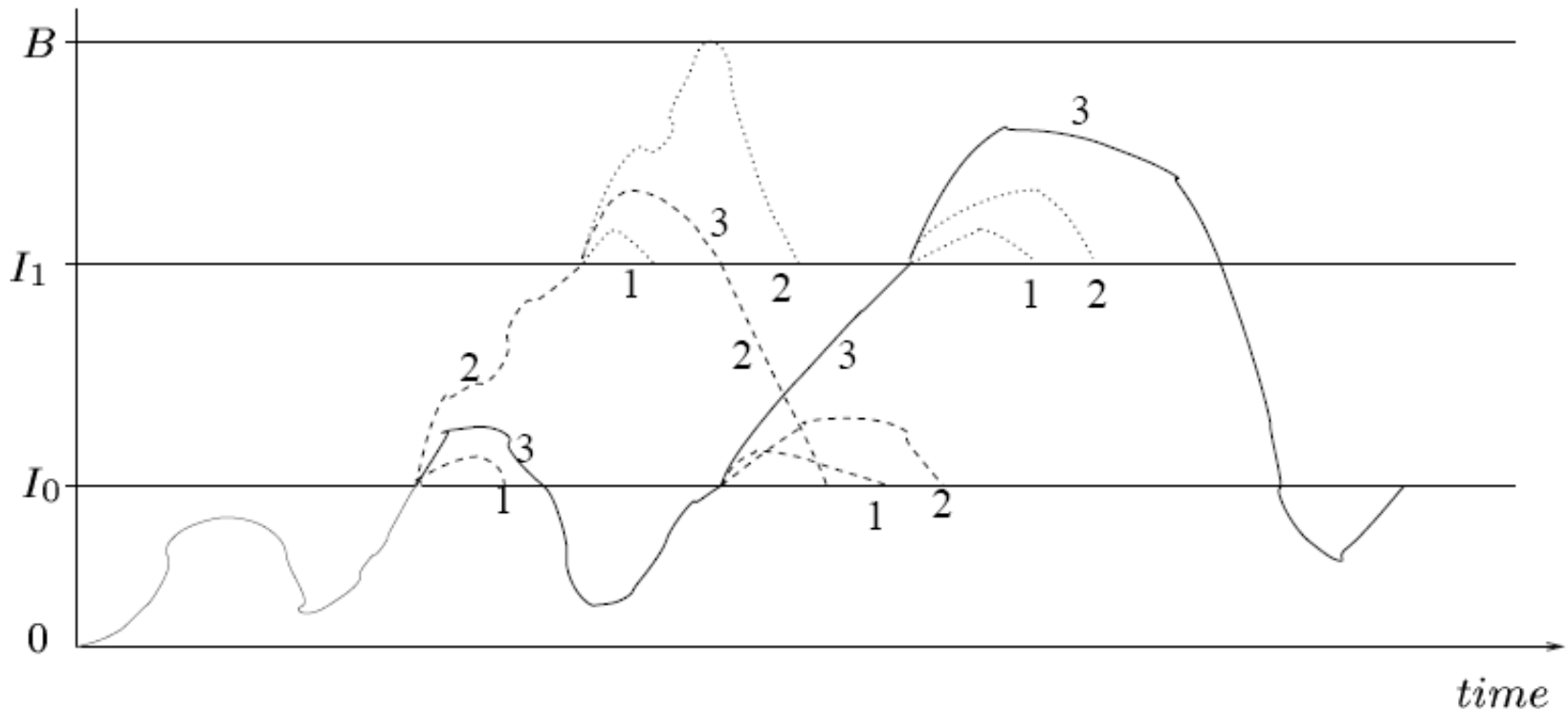
$$(n_1 = r_1).$$

Single Step vs. Global Step

In any simulation experiment involving RESTART we have two choices: either we simulate “stage-by-stage” or “root-by-root”. In the first case we complete all the paths starting from a certain stage before we move to the next one. This is called the *Single Step* approach.



In the second case, we generate all the offspring originating from a single root (i.e. a state in the upper stage which is reached immediately from the lower state) before we move to the next root. We call this the *Global Step* approach.



Literatur:

C. Görg, E. Lamers, O. Fuß, P. Heegaard: Rare Event Simulation, chapter in: Modelling and Simulation Environment for Satellite and Terrestrial Communication Networks, Proceedings of the European COST Telecommunications Symposium (Cost256), Chapter 21, pp. 365-396, Kluwer, Boston, Dordrecht, New York, 2001.

M.J.J.Garvels und D.P.Kroese, a comparison of RESTART implementations. Proceedings of the 1998 Winter Simulation Conference, D.J.Medeires, E.F.Watson, J.S.Carson, and M.S.Manivannan, eds.