The AA$^{\mathcal{C}}$-Method for the Evaluation of Simulation Results

Norbert Th. Müller
Abteilung Informatik — Universität Trier
D-54286 Trier, Germany
E-mail: mueller@uni-trier.de

Abstract: In this paper, we present and analyze an algorithm for the evaluation of simulation results, called the AA$^{\mathcal{C}}$-method. It is an enhancement of the LRE-algorithm, derived from the sojourn time analysis of that algorithm.

The AA$^{\mathcal{C}}$ method is based on weaker assumptions than the LRE and thus usually creates confidence intervals that are more reliable that delivered by the LRE-algorithm.

Keywords: Discrete event simulation, LRE-algorithm, confidence interval, sojourn times

1 Introduction

In this paper, we present and analyze an enhancement of the LRE-algorithm derived from the sojourn time analysis in [Mu00], called the AA$^{\mathcal{C}}$-method.

The classical method of evaluation of simulations surely is the batch means method, see e.g. [BFS87], giving confidence intervals to express the precision of the simulation results. Unfortunately, it is not suited for rare event simulation techniques, as they do not produce the necessary large batches.

As an alternative, the LRE-algorithm has been introduced in [Schr84]. Its name is based on the 'Limited Relative Error' that has been chosen to describe the precision of the results. A novel analysis of this algorithm, based on sojourn times, has been given in [Mu00]. This analysis allowed to enhance the basic LRE idea with confidence intervals, thus making it possible to replace the name giving Limited Relative Error by this more prominent error measure. This slightly enhanced algorithm will be called LRE/sojourn in the following.

The AA$^{\mathcal{C}}$-method as a further and significant enhancement of the LRE idea is based on weaker assumptions and thus usually creates confidence intervals that are more reliable than those delivered by the LRE-algorithm.

2 The LRE-algorithm

The batch means method (e.g. [BFS87]) tries to construct independent samples from a given time series by building batches that tend to have lesser autocorrelation than the original series. So essentially it aims on the reduction of autocorrelation. It results in confidence intervals for the mean of the simulation results.

In contrast, the LRE-algorithm, as introduced in e.g. in [Schr84] and used in a variety of papers later on, e.g. [Schr99, GöSc96, Gö97], measures the autocorrelation and tries to deduce the precision of the simulation results from this autocorrelation (instead of reducing it). Additionally, instead of confidence intervals, the 'Limited Relative Error' has been introduced as an error measure. The reliability of the LRE method has been questioned e.g. in [BBK], where attempts were made to improve on it.

In this paper we will enhance the analysis of LRE given in [Mu00]. First, we will briefly recall the idea behind the LRE-algorithm. Consider a jump process $X(t)$ with state space S. Given a subset $A \subseteq S$, the problem is to determine

$$P(A) := P\{X(t) \in A\} \tag{1}$$

As $X(t)$ is a jump process, we may derive two time series: $\{S_i \mid i \in \mathbb{N}\}$ of the states of the process, and $\{T_i \mid i \in \mathbb{N}\}$ of the sojourn times in states S_i. The above settings apply to processes in continuous time as well as in discrete time. For the latter we simply always have $T_i = 1$

Let Y_i be the time series defined by $Y_i = 1 \Leftrightarrow S_i \in A$.

In order to answer our original question of estimating how long the process $X(t)$ stays in A, the following values are computed from $X(t)$ (i.e. they are measured during a simulation of the process):

- sum $s = s(n) = \sum_{i \le n} T_i$ of the first n sojourn times, i.e. the total time for the first n process steps.

- sum $r = r(n) = \sum_{i \le n, Y_i = 1} T_i$ of sojourn times in states $S_i \in A$, i.e. the amount of time the process spends in A during the first n steps.
- quantity $a = a(n)$ of transitions from $Y_i = 1$ to $Y_{i+1} = 0$ for $1 \le i \le n$, i.e. the quantity of observed pairs (S_i, S_{i+1}) with $S_i \in A, S_{i+1} \in A^c$.

In this chapter, n will be treated as fixed, and we will simply use a, s, and r instead of $a(n)$, $s(n)$ and $r(n)$.

Between subsequent transitions Y_i=1 $\rightarrow$ Y_{i+1}=0 and Y_j=1 $\rightarrow$ Y_{j+1}=0 there must be a third transition Y_k=0 $\rightarrow$ Y_{k+1}=1, $i < k < j$. So it is not necessary to measure the following values, as they can be deduced with sufficient precision from n, a, r:

- time $v = s - r$ spent in states from A^c.
- quantity $c \approx a$ of transitions from Y_i=0 to Y_{i+1}=1 for $1 \le i \le n$, i.e. the quantity of observed pairs (S_i, S_{i+1}) with $S_i \in A^c, S_{i+1} \in A$.

Please note that the approximation $c \approx a$ from above may be wrong by at most 1. So in the following we will simply use $c = a$.

The following approximation is immediate:

$$P(A) \approx r/s \tag{2}$$

To assess the precision of this approximation, we consider the alternating time series (Y_i). Here the central assumption of the LRE-algorithm (for the case of discrete time processes $X(t)$ was as follows:

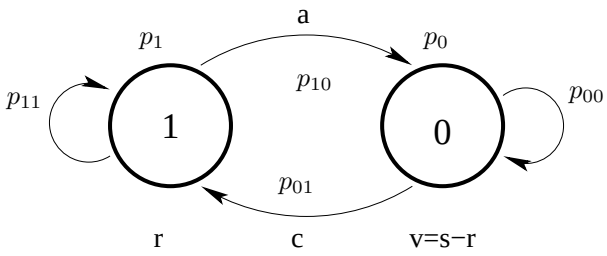
- (0-LRE) Assume that the time series $(Y_i)_{i \in \mathbb{N}}$ is a discrete homogenous Markov chain with just the two states 0 and 1!

Please note that in the discrete case we simply have $s(n) = n$. In [Gö97, GöSc96, Schr99] this chain is called the $F(x)$-equivalent 2-Node Markov chain. Such a chain is characterized by its one step transition probabilities:

$$p_{ij} = \text{Prob(transitions starting in } i \text{ lead to } j)$$

where $i, j \in \{0, 1\}$. Resulting from this, there are the steady state probabilities p_0 and p_1 for being in state 0 resp. in state 1. So the central assumption of the LRE algorithm is $P(A) = p_1$ and $P(A^c) = p_0$ for these values derived from the Markov chain.

Later in this paper we will consider the implications of this central assumption and also discuss its validity. But first we recall the results from [Schr99] for this 2-Node chain. The following graph contains the transition probabilities as well as the steady state probabilities together with the measured (or derived) values:



The following estimates are obvious (please remember $s(n) = n$):

$$p_1 = P(Y_i = 1) \qquad \approx \tfrac{r}{s} \tag{3}$$

$$p_0 = P(Y_i = 0) \quad \approx \tfrac{v}{n} = \tfrac{s-r}{s} \tag{4}$$

$$p_{10} = P(Y_i = 0 \mid Y_{i-1} = 1) \qquad \approx \tfrac{a}{r} \tag{5}$$

$$p_{01} = P(Y_i = 1 \mid Y_{i-1} = 0) \quad \approx \tfrac{c}{v} \approx \tfrac{a}{s-r} \tag{6}$$

The analysis of the LRE algorithm from [Schr99] shows that the set of values p_1 that 'fit' to s, r, a is approximately normal distributed for sufficiently large values of s, r, v:

$$p_1 \sim N(\frac{r}{s}, \sigma^2) \tag{7}$$

with

$$\sigma^2 \approx \frac{rv}{s^3} \cdot (\frac{2rv}{as} - 1) \tag{8}$$

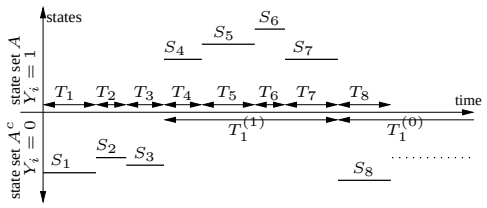
Instead of defining confidence intervals, the use of *limited relative error (LRE)* $\frac{\sigma}{r/s}$ was proposed in [Gö97, Schr99]. Because $p_0 + p_1 = 1$, σ also is an error measure for the approximation v/s of p_0. If r/s is significantly larger than v/s, then the relative error $\frac{\sigma}{v/n}$ is bigger than $\frac{\sigma}{r/n}$. In [Gö97, Schr99], the authors suppose to perform a simulation until both values are below 0.05. For the validity of the approximations (that are derived via an application of the central limit theorem), they give only very crude 'large sample conditions': $n > 1000$, $r, v > 100$, $a, c > 10$.

3 Sojourn time analysis

Unfortunately, the analysis in [Schr99] is strictly concentrated on the statistics of 2-Node Markov chain and so it does not give hints how good the approximations are in case that condition (0-LRE) does not hold. In the following we refine the analysis from [Mu00] based on sojourn times.

The alternating time series $Y_1, Y_2, Y_3, \ldots$ together with the sojourn times T_i define two sequences $(T_j^{(1)})$ and $(T_j^{(0)})$ of sojourn times in states 1 and 0 (corresponding to the state set A and its complement A^c in the original process $X(t)$).

We neglect the first, partial sojourn time and start building the $(T_j^{(\cdot)})$ after the first change in the values of Y_i.



When considering the LRE-algorithm, we see that using the 2-Node Markov chain to model the behavior of the simulation contains three basic assumptions concerning these sojourn times:

- (1-LRE) The sojourn times $(T_j^{(1)})$ for state 1 are supposed to have identical geometrical distribution $T^{(1)}$ given by probability p_{10}.
- (2-LRE) The sojourn times $(T_j^{(0)})$ for state 0 are supposed to have identical geometrical distribution $T^{(0)}$ given by probability p_{01}.
- (3-LRE) The random sequences $(T_i^{(1)})$ and $(T_i^{(0)})$ are supposed to be mutually independent.

In the following we will relax the assumptions on the distributions assumed by the LRE method, but still use the third one:

- (1-AA$^{\mathcal{C}}$) The sojourn times $T_j^{(1)}$ in state 1 are supposed to be identically distributed with mean $\mu_1 := E[T_j^{(1)}]$ and variance $\sigma_1^2 := Var(T_j^{(1)})$.
- (2-AA$^{\mathcal{C}}$) The sojourn times $T_j^{(0)}$ in state 0 are supposed to be identically distributed with mean $\mu_0 := E[T_j^{(0)}]$ and variance $\sigma_0^2 := Var(T_j^{(0)})$.
- (3-AA$^{\mathcal{C}}$) The random sequences $(T_i^{(1)})$ and $(T_i^{(0)})$ are supposed to be mutually independent.

Condition 1-LRE implies 1-AA$^{\mathcal{C}}$. In that case, the properties of the geometric distribution imply

$$\mu_1 = 1/p_{10} \tag{9}$$

and

$$\sigma_1^2 = (\mu_1 - 1)/\mu_1^3 = p_{11}/p_{10}^2 \tag{10}$$

The same holds for A^c: 2-LRE implies 2-AA$^{\mathcal{C}}$with

$$\mu_0 = 1/p_{01} \tag{11}$$

and

$$\sigma_0^2 = (\mu_0 - 1)/\mu_0^3 = p_{00}/p_{01}^2 \tag{12}$$

In the following we will assume that the three conditions (1-AA$^{\mathcal{C}}$), (2-AA$^{\mathcal{C}}$), and (3-AA$^{\mathcal{C}}$) hold, and we will always explicity mention what the implications form the stronger (1-LRE), (2-LRE) would be.

μ_1 and μ_0 are the mean sojourn times, so we have

$$P(A) = \frac{\mu_1}{\mu_1 + \mu_0} \tag{13}$$

Now suppose that we perform a simulation exactly until the first a samples of sojourn times $T_j^{(1)}$ in state 1 and the first a sojourn times $T_j^{(0)}$ in state 0 are completed. We still measure the same three values s, a, r as above (with the slight modification that we start measuring at a transition from A^c to A). In contrast to the previous chapter, where n was fixed and a and r were depending on n, now a is fixed and $s = s(a)$ and $r = r(a)$ are depending on a. We still use $v = s - r$. In fact we have

$$r = \sum_{j=1}^{a} T_j^{(1)} \tag{14}$$

$$v = \sum_{j=1}^{a} T_j^{(0)} \tag{15}$$

Now consider

$$R_a^{(1)} := 1/a \cdot \sum_{j=1}^{a} T_j^{(1)} \tag{16}$$

and

$$R_a^{(0)} := 1/a \cdot \sum_{j=1}^{a} T_j^{(0)} \tag{17}$$

$R_a^{(1)}$ is the mean value of a identically distributed random variables, that are independent because of (3-AA$^{\mathcal{C}}$). If a is large enough, we may apply the central limit theorem stating that $R_a^{(1)}$ has approximately normal distribution with mean μ_1 and variance σ_1^2/a.

Our measured value r/a is indeed just a sample $r/a = 1/a \cdot \sum_{j=1}^{a} T_j^{(1)}$ of $R_a^{(1)}$, so we may deduce confidence intervals for μ_1

$$P(\ |\mu_1 - r/a| < z \cdot \sigma_1/\sqrt{a}\ ) \leq \alpha(z) \tag{18}$$

using the well known function $\alpha(z) = norm(z) - norm(-z)$ for the width of confidence intervals at confidence level z.

A similar result holds for $R_a^{(0)}$ and μ_0:

$$P(\ |\mu_0 - v/a| < z \cdot \sigma_0/\sqrt{a}\ ) \leq \alpha(z) \tag{19}$$

By (3-AA$^{\mathcal{C}}$) the random variables $R_a^{(1)}$ and $R_a^{(0)}$ are independent. But the fraction

$$\frac{R_a^{(1)}}{R_a^{(1)} + R_a^{(0)}} \tag{20}$$

of two (almost) normal distributed random variables is certainly *not* normal distributed! However, if both $\mu_1 > 0$ and $\mu_0 > 0$, the fraction will be approximately normal distributed, if the variances $\sigma_1/\sqrt{a}$ and $\sigma_0/\sqrt{a}$ are small:

$$\begin{aligned}
&\frac{R_a^{(1)}}{R_a^{(1)} + R_a^{(0)}} - \frac{\mu_1}{\mu_1 + \mu_0} \\
&= \frac{(R_a^{(1)} - \mu_1) \cdot \mu_0 - \mu_1 \cdot (R_a^{(0)} - \mu_0)}{(R_a^{(1)} + R_a^{(0)}) \cdot (\mu_1 + \mu_0)} \\
&\approx \frac{(R_a^{(1)} - \mu_1) \cdot \mu_0 - \mu_1 \cdot (R_a^{(0)} - \mu_0)}{(\mu_1 + \mu_0) \cdot (\mu_1 + \mu_0)} \\
&= R_a^{(1)} \cdot \frac{\mu_0}{(\mu_1 + \mu_0)^2} - R_a^{(0)} \cdot \frac{\mu_1}{(\mu_1 + \mu_0)^2} \qquad (21)
\end{aligned}$$

The relative(!) error of this approximation tends to 0, if $R_a^{(1)}$ approaches μ_1 and $R_a^{(0)}$ approaches μ_0. So the precision of the approximation will increase with a, as the variances σ_1^2/a and σ_0^2/a converge to zero.

In consequence, for large a the distributions must be very similar, as soon as the variances are small compared to μ_1 and μ_0. As the class of normal distributed random variables is closed under finite sums and multiplication with a constant value, (21) is obviously normal distributed. Its mean is 0, and its variance turns out to be

$$\sigma_a^2 := \frac{\sigma_1^2 \mu_0^2 + \sigma_0^2 \mu_1^2}{a \cdot (\mu_1 + \mu_0)^4} \qquad (22)$$

Putting all things together, we get the following summary valid for the more general AA^c-conditions:

- For large a, (20) is approximately(!) normal distributed with mean $P(A)$ and variance (22), and we have the sample r/s of (20):

$$r/s = \frac{r/a}{s/a} = \frac{r/a}{r/a + v/a}$$

- It is now clear to see where large sample conditions are necessary: a should allow application of the the central limit theorem:

$$a > 30 \qquad (23)$$

  This is significantly larger than the corresponding $a = 10$ from the original LRE analysis dealing essentially with geometric distributions, where such a small $a \approx 10$ was reasonable.

  Furthermore the approximation (21) should be sufficiently precise:

$$\sigma_1/\sqrt{a} << \mu_1 \quad , \quad \sigma_0/\sqrt{a} << \mu_0 \qquad (24)$$

Now suppose that the stronger LRE-conditions are given: Substituting the values from (9) to (12), we see that (22) and (8) are identical w.r.t. the approximations (5) and (6):

$$\sigma_a^2 \approx \frac{rv}{s^3} \cdot (\frac{2rv}{as} - 1) \approx \sigma^2$$

So our analysis leads to the same results as [Schr99], but now r/s turns out to be a sample from an (almost) normal distribution, while in the original paper the variance of p_1 was constructed in a fairly complicated way with an a-posteriori argument from s, a, r.

4 The AA^{c}method

The analysis from the previous section immediately leads to the following modification of the LRE-algorithm. We will call this the AA^c-method:

During a simulation, measure the following values starting immediately *after* the first transition from A^c to A:

- quantity a of transitions from A^c to A.
- time $s = s(a)$ necessary for these a full sojourn times in A and in A^c.
- sojourn time $r = r(a)$ of the values $Y_i = 1$ for $1 \leq i \leq n(a)$, i.e. the time $X(t)$ spends in A during these $n(a)$ steps.
- sums s_1 of the squares of the a sojourn times in A
- sums s_0 of the squares of the a sojourn times in A^c

So in contrast to the LRE method, we additionally measure the sojourn time squares and use a special trigger to start the measurements.

From these measurements we get the following approximations for the values μ_0, μ_1, σ_0^2, σ_1^2, and σ_a^2: Again let $v = v(a) := s(a) - r(a)$ be the time spend in the complementary set A^c.

$$\mu_1 \approx \frac{r}{a} \qquad (25)$$

$$\mu_0 \approx \frac{s-r}{a} \qquad (26)$$

$$\sigma_1^2 \approx \frac{s_1}{a} - \mu_1^2 \approx \frac{s_1}{a} - \frac{r^2}{a^2} \qquad (27)$$

$$\sigma_0^2 \approx \frac{s_0}{a} - \mu_0^2 \approx \frac{s_0}{a} - \frac{v^2}{a^2} \qquad (28)$$

$$\sigma_a^2 \approx \frac{v^2 s_1 + r^2 s_0 - 2r^2 v^2/a}{s^4} \qquad (29)$$

Due to the (almost) normal distribution of r/s shown in the previuos section, this leads to the following confidence intervals for $P(A)$:

$$P\left(|P(A) - \frac{r}{s}| \; < \; z \cdot \frac{\sqrt{v^2 s_1 + r^2 s_0 - 2r^2 v^2/a}}{s^2} \right) \leq \alpha(z) \qquad (30)$$

5 Discussion and experimental results

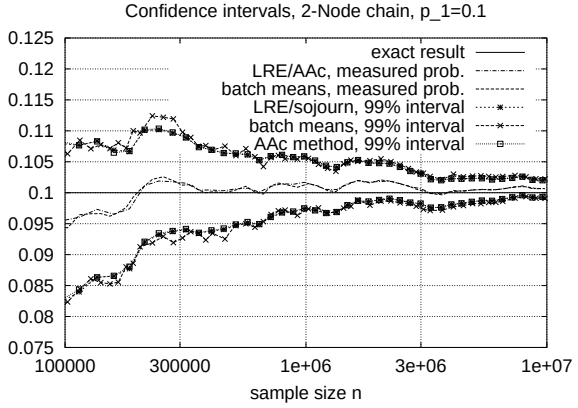
Our sojourn time analysis (22) shows that there is a big influence of the sojourn time variance on the resulting confidence intervals. If the real variances σ_1^2 and σ_0^2 are smaller than given by the LRE-conditions as in (10,12), the LRE-algorithm will result in unnecessarily large intervals, while variances σ_1^2 and σ_0^2 bigger than that lead to intervals that are too small.

In the following we present the results of a few tests to illustrate this behavior. Using three different types of sequences (Y_i) derived from discrete time Markov processes $X(t)$ and for a range of samples from 10^5 to 10^7 we measured 99% confidence intervals for three different methods of evaluating a simulation:

1. an implementation of the LRE-algorithm using (8) to define confidence intervals (called LRE/sojourn),
2. an implementation of the Law-Carson algorithm for the batch means method (see e.g. [BFS87]), and
3. the AA^c-method as described above.

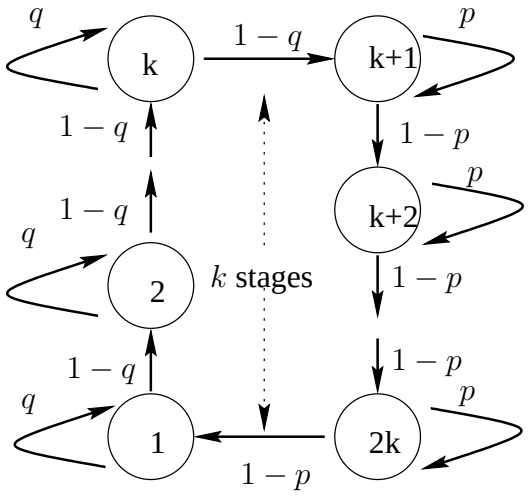
The intervals due to the batch mean method do not rely on the LRE hypotheses, so they vary correctly with the different variances of the sojourn times. Additionally, the mean values for LRE and AA^{c}are always identical because of their construction. The mean values for batch method are slighty different, as they were measured at slighty different points in time.

We first start with a setting that fits to (0-LRE), i.e. the time series (Y_i) indeed came from a 2-Node Markov chain. We chose $p_1 = 0.1$ and $p_{10} = 0.06$. Please note that the other transition probabilities are uniquely defined by p_1 and p_{10}.

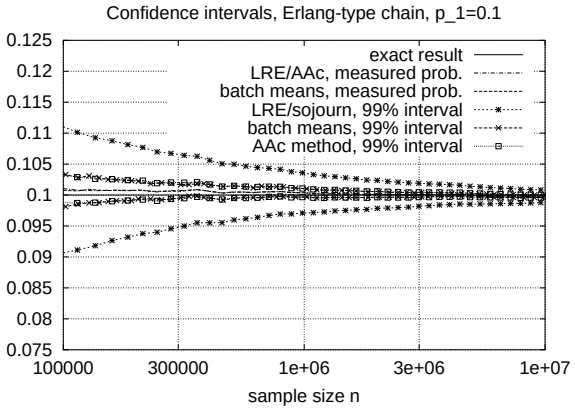


As expected, the confidence intervals turned out to be almost identical for this setting.

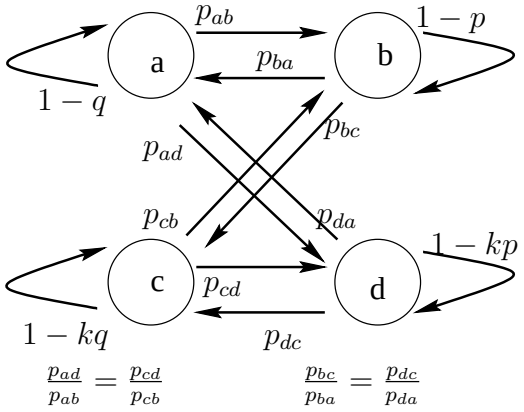
The second test used a series with sojourn times of significant smaller variance. It is based on a 2k-Node Markov chain, where the left nodes build set A while the others correspond to set A^c. This leads to sojourn times of a discrete hypo-exponential type distribution with small variance.



We tried this example with $k = 10$ and $p = 0.99$. As expected, the confidence intervals of LRE are much larger than for the other two methods that are on a par.

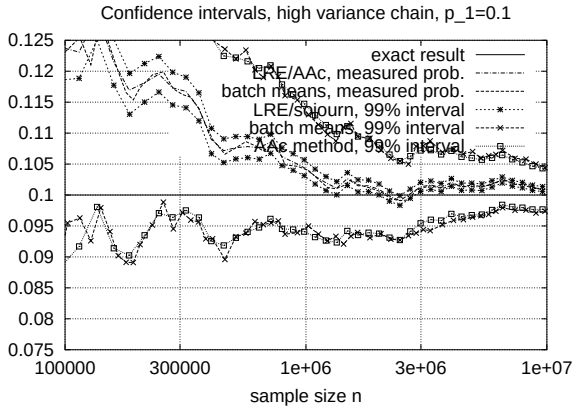


The third test used a 4-Node chain with highly variant sojourn times:



We tried an example with $k = 100$, so the sojourn times are build by mixtures of two geometric distributions differing by a factor of 100. These sojourn times have a kind of discrete hyper-exponential type distribution. As expected, the LRE confidence intervals are way too small

in this case, while the other two methods again give similar results:



Please note that in all examples, the assumption (3-LRE)=(3-AA^C) about the independency of the sojourn times is valid! This is true also for any Markovian birth-death process, so (1-LRE) and (2-LRE) seem to be more critical.

Obviously, condition (0-LRE) will not be true in most real simulation scenarios. On the other hand, the one step transition probabilities lead to correct values for the LAG-1-autocorrelation of the series (Y_i). Unless additional data is measured from the simulation (like has been done in the new AA^C-method), the assumption of Markovian behavior implies estimates on the higher type autocorrelations that can hardly be improved.

6 Conclusion

From the examples, the disadvantages of the 2-Node approach can be clearly seen. On the other hand, this approach can be easily incorporated into rare event simulations, see e.g. [Gö97]. Here the batch means method is not applicable at all. So regardless of these above problems, the LRE algorithm is still very useful in this field.

The new AA^C method shows the possible improvements coming essentially from the sojourn time analysis. At the time, we are working on integrating the method into scenarios of rare event simulations. Initial results on the gained confidence intervals for non-geometric sojourn times are very promising.

References

[BBK] Below, Kai & Battaglia, Lorenzo& Killat, Ulrich, *RESTART/LRE Simulation: The Reliability Issue*, Second International Workshop on Rare Event Simulation, RESIM99, Enschede, The Netherlands

[BFS87] Bratley, Paul & Fox, Bennett L.& Schrage, Linus E., *A Guide to Simulation* (Springer, New York, 1987)

[GöSc96] Görg, C., Schreiber, F., *The RESTART/LRE Method for Rare Event Simulation* (1996 Winter Simulation Conference, pp390-397, Coronado, California USA)

[Gö97] Görg, C., *Verkehrstheoretische Modelle und stochastische Simulationstechniken zur Leistungsanalyse von Kommunikationsnetzen* (Habilschrift, Aachen, 1997)

[Mu00] Müller, N.Th., *An analysis of the LRE-Algorithm using Sojourn Times* (14th European Simulation Multiconference, 2000, Ghent (ESM 2000))

[Schr84] Schreiber, F., *Time Efficient Simulation: The LRE-algorithm for producing empirical distribution functions with limited relative error* (AEÜ, 38 93-98, 1994)

[Schr99] Schreiber, F., *A simplified Formula Basis for the LRE-Algorithm Time Efficient Simulation* (MMB '99, Trier)