JAERI－Conf 2003－019

Automated Suppression of the Initial Transient in Monte Carlo Calculations based on Stationarity Detection using the Brownian Bridge Theory

Yann RICHET[1*] – Olivier JACQUET[1] – Xavier BAY[2]
[1] *Institut de Radioprotection et de Sûreté Nucléaire, BP 17, 92262 Fontenay-Aux-Roses Cedex, France*
[2] *Ecole Nationale Supérieure des Mines de Saint Etienne, 158, cours Fauriel, 42023 Saint-Etienne, France*

The accuracy of a criticality Monte Carlo (MC) calculation requires the convergence of the k-effective series. Once the convergence is reached, the estimation of the k-effective eigenvalue must exclude the initial transient of the k-effective series. The present paper deals with a post-processing algorithm to suppress the initial transient of a criticality MC calculation, using the Brownian Bridge theory.

KEYWORDS: Monte Carlo calculation, Initial Transient, Stationarity Detection, Brownian Bridge.

1. Introduction

The initial transient of a criticality MC calculation induces a bias in the k-effective estimate. A "positive" transient corresponds to an initial overestimation and induces a positive bias, whereas a "negative" transient corresponds to an initial underestimation and induces a negative bias.

For instance the fourth benchmark of the NEA Source Convergence Expert Group[1] is designed to produce a transient in the k-effective series due to a bad initialization of the sources distribution (Fig. 1).

Estimation of the k-effective ($\sigma_{keff} \approx 0.00240$):

- 1.10327 without transient suppression
- 1.11682 with an empirical transient suppression of 170 observations

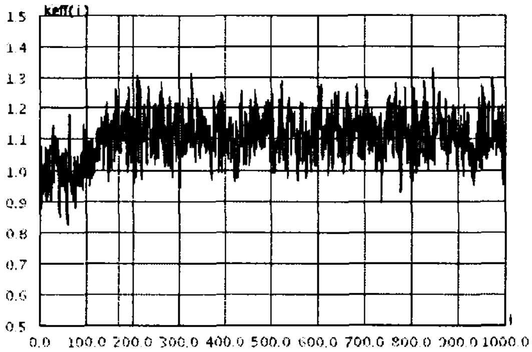

Fig. 1 Cycle k-effective series from NEA Source convergence benchmark 4

This example shows a bias of more than 0.01300 in the k-effective estimate, and the k-effective MC uncertainty of 0.00240 is not sufficient to take into account this bias. The post-processing algorithms detailed in this paper are designed to suppress this kind of transient and the induced k-effective bias.

2. Scalar series used to detect the initial transient

The initial transient observed in the k-effective series $\{k(i)\}_{1\le i\le N}$ is obviously a consequence of the source distribution (or eigenvector) convergence process from the initial guess to the fundamental mode. Since it is much easier to study the convergence in a series of scalar values than in a series of vectors, and because the cycle k-effective series is used for the estimation of the k-effective eigenvalue, a first approach to suppress the initial transient – using the Brownian bridge[2] theory – focused on the stationarity detection of the cycle k-effective series.[3]

However the apparent convergence of k-effective series does not strictly imply the convergence of the source distribution. Recently, a stationarity diagnostic based on the Shannon entropy of source distribution, using the two-sample F test, was proposed.[4] An advantage of this method lies in the use of a more representative scalar value of sources distribution than k-effective.

Nevertheless the stationarity of Shannon entropy does not rigorously imply sources distribution stationarity and moreover stationarity of sources distribution does not guarantee the convergence towards the fundamental mode. Finally, it should be noticed that the transient suppression aims only at obtaining a stationary series and cannot guarantee the convergence of the sources distribution in any way.

As a consequence, it is necessary to run a sufficient number of generations to ensure that the MC powering algorithm has converged to the true eigenvalue, or to use an improved powering method which guarantees that the most reactive parts of the system are correctly sampled after the transient.[5]

Keeping these points in mind, the stationarity detection can be based on the cycle k-effective series or on the cycle Shannon entropy series. So, let us define a series $x=\{x(i)\}_{1\le i\le N}$ of N observations standing for either a k-effective series or a Shannon entropy series resulting from a MC calculation.

*Corresponding author, Tel. +33-1-58-35-88-84, Fax. +33-1-46-57-29-98, E-mail: yann.richet@irsn.fr

3. Series and bridge processes

The underlying idea of the stationarity detection is to consider a process revealing the stationarity or non-stationarity of the series.

Let us consider $\bar{x}(j)=\frac{1}{j}\sum_{i=1}^{j} x(i)$.

Basically, when the series x is stationary, the series $\{j(\bar{x}(j)-\bar{x}(N))\}_{1\le j\le N}$ has small values and is centered around 0 (see Fig. 6). On the contrary, in the case of a negative (resp. positive) initial bias in the series x, the series $\{j(\bar{x}(j)-\bar{x}(N))\}$ is not centered around 0 and has a negative (resp. positive) peak (see Fig. 7).

Let us introduce the "series bridge" of x:

- $BS_x\left(\frac{j}{N}\right)=\frac{1}{\tau\sqrt{N}}\, j(\bar{x}(j)-\bar{x}(N))$

  for $1\le j\le N$, with $\tau^2=\lim_{N\to\infty} N\,Var(\bar{x}(N))$,
- $BS_x(0)=0$.

In order to analyze the stationarity of the x series, G. Vassilacopoulos[6] suggests to work with the "ranks series" of x in place of the x series. This ranks series is defined as $\{R_x(i)\}_{1\le i\le N}$ where $R_x(i)$ is the rank of $x(i)$, i.e. the ascending sort order of $x(i)$ among all x values. It can be proved that the $\{R_x(i)\}$ process is stationary if the x series is stationary.

In the same way that $\bar{x}(j)$ let us define $\bar{R}_x(j)=\frac{1}{j}\sum_{i=1}^{j} R_x(i)$. Another "bridge" process can thus be based on the "ranks series" such that:

- $BR_x\left(\frac{j}{N}\right)=\frac{1}{\sigma\sqrt{N}}\, j(\bar{R}_x(j)-\bar{R}_x(N))$,

  for $1\le j\le N$, where $\sigma=\frac{1}{2\sqrt{3}}\sqrt{N(N+1)}$ is the sample standard deviation of the R_x series.
- $BR_x(0)=0$.

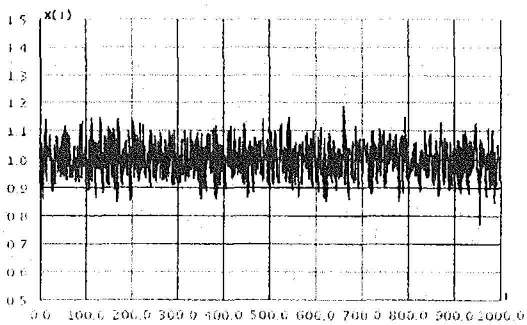

Fig. 3 A stationary series

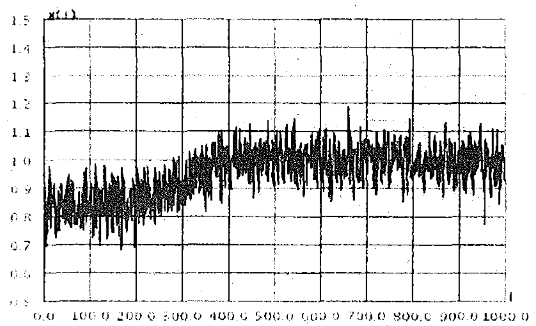

Fig. 2 A non stationary series

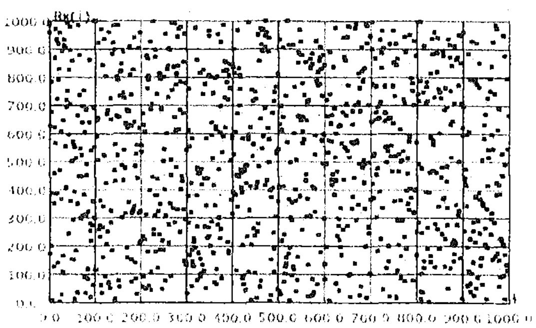

Fig. 5 Ranks of a stationary series

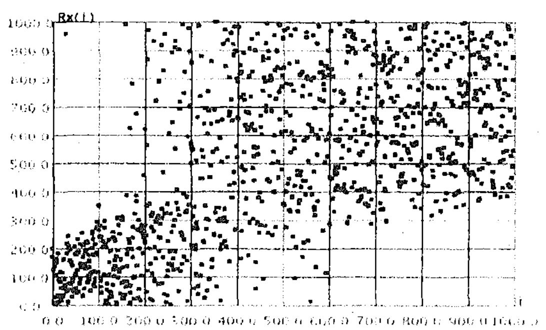

Fig. 4 Ranks of a non stationary series

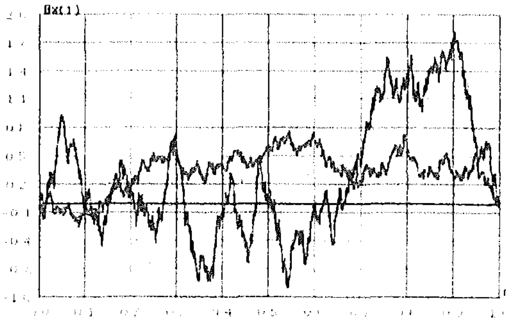

Fig. 6 Bridges of a stationary series

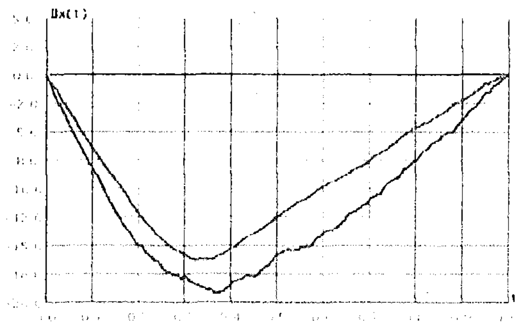

Fig. 7 Bridges of a non stationary series

Fig. 3 to 7 show respectively a stationary series, its ranks series and its bridges and a non-stationary series, its ranks series and its bridges. Note that the bridges ranges are very low for stationary series and reach high values in case of non stationary series.

The terminology 'bridge" comes from the fact that all the bridges processes equal 0 when $t = 0$ and $t = 1$, thus the curves look like bridges between the points (0,0) and (1,0) (see Fig. 6, 7).

It can be proved that bridges processes tend to the Brownian bridge, when the x series is weakly dependent − i.e. when two $x(i)$ observations widely separated from each other in the series are almost independent (in the sense of phi-mixing[2]) − and as N tends to infinity.

4. Characterization of stationarity

The bridges processes defined previously can be useful to determine if a series is stationary or not. Indeed it is possible to determine the distribution of some characteristic values (called statistics and denoted as S) of the bridges of stationary series.

Let us give some examples of such statistics used in literature, B_x being BS_x or BR_x:

$$Sv_{min}(B_x)=min(B_x)\ ,\ Ss_{min}(B_x)=\frac{B_x(t_{min})^2}{t_{min}(1-t_{min})}\ ,$$

$$Sv_{max}(B_x)=max(B_x)\ ,\ Ss_{max}(B_x)=\frac{B_x(t_{max})^2}{t_{max}(1-t_{max})}$$

where t_{min} (resp. t_{max}) is the value at which the minimum (resp. maximum) of $B_x(t)$ occurs.

But other characteristic values of the bridges can be mentioned:

$$Sr_{min}(B_x)=\frac{-B_x(t_{min})}{B_x(t_{max})}\ ,\ Sr_{max}(B_x)=\frac{-B_x(t_{max})}{B_x(t_{min})}\ ,$$

$$Ssr_{min}(B_x)=\frac{Ss_{min}(B_x)}{Ss_{max}(B_x)}\ ,\ Ssr_{max}(B_x)=\frac{Ss_{max}(B_x)}{Ss_{min}(B_x)}\ ,$$

$$min_{0\le t\le 1}\left(\frac{B_x(t)^2}{t(1-t)}\right)\ ,\ Sad(B_x)=\sum_{j=0}^{N}\frac{B_x\left(\frac{j}{N}\right)^2}{\frac{j}{N}\left(1-\frac{j}{N}\right)}$$

and $Sfc(B_x)=t_0$ where t_0 is the first value strictly greater than $\frac{1}{N}$ where $B_x(t_0-\frac{1}{N})\cdot B_x(t_0)\le 0$.

In fact, the characterizations of minimum and maximum values of bridges are two symmetrical point of views. Thus, the distributions based either on minimum or maximum are identical.

Note that when B_x is a Brownian bridge (i.e. bridge of a weakly dependent series):

- G. Vassilacopoulos shows that Sv_{min} and Sv_{max} both follow a Kolmogorov-Smirnov distribution[6]
- L. Schruben shows that Ss_{min} and Ss_{max} both follow a χ_3^2 distribution[7]
- X. Bay shows that Sr_{min} and Sr_{max} both follow a distribution having this cumulative density: $F_R(r)=1-\frac{\pi r}{(1+r)^2}\left(\frac{1+r}{\pi}-\cot\left(\frac{\pi}{1+r}\right)\right)$

All the statistics (denoted as S^-) designed to characterize the minimum value of the bridges can only detect negative transients (for instance Sv_{min}, Ss_{min}, Sr_{min}). In the same way, all the statistics (denoted as S^+) designed to characterize the maximum value of the bridges can only detect positive transients (for instance Sv_{max}, Ss_{max}, Sr_{max}). A new kind of statistic based on two symmetrical statistics S^- and S^+ can be defined as $max[S^-,S^+]$. This kind of statistic detects both positive and negative transients. We only tested such statistics in our paper:

- **Schruben statistic** based on the series bridge[7]: $Ss(BS_x)=max\{Ss_{min}(BS_x),Ss_{max}(BS_x)\}$

This statistic is correctly estimated only if τ^2 is well known, but τ^2 estimation is very influenced by the non stationarity of the series. One way suggested to improve the τ^2 estimation accuracy is to estimate τ^2 just on the last half of the series. The two following statistics avoid the estimation of this parameter.

- **Supremum ratios statistic** based on the series bridge: $Sr(BS_x)=max\{Sr_{min}(BS_x),Sr_{max}(BS_x)\}$
- **Vassilacopoulos statistic** based on the ranks bridge[6]: $Sv(BR_x)=max\{Sv_{min}(BR_x),Sv_{max}(BR_x)\}$

Nevertheless all the statistics distributions are established for theoretical Brownian bridges, and the k-effective or Shannon entropy series are all the more far from weakly dependent processes (see section 3) as the length of the series N is low. In fact the series resulting of MC calculations have various finite lengths and are somewhat autocorrelated. In this paper, the two main parameters considered as influencing the statistics and their distributions are the length N and the first autocorrelation coefficient ρ of the series (see Fig. 8, 9). The determination of theoretical statistics distributions versus N and ρ has to be achieved.

However an empirical estimation of each statistic distribution $\Delta_{N,\rho}$ is possible by estimating this statistic for a great number of series with the same values of N and ρ. For the purpose of our study, a plan of experiments has been defined to evaluate distributions of the $Sr(BS_x)$, $Ss(BS_x)$, $Sv(BR_x)$ statistics for $N\in\{25,100,500,1000\}$ and $\rho\in\{0.0,0.1,0.2\}$. Practically, when an estimation of $\Delta_{N,\rho}$ has to be performed with N or ρ not included in the previous plan of experiments, the distributions of the statistics are linearly interpolated. This interpolation could be improved using a more dense plan of experiments and a better model than a simple linear interpolation.

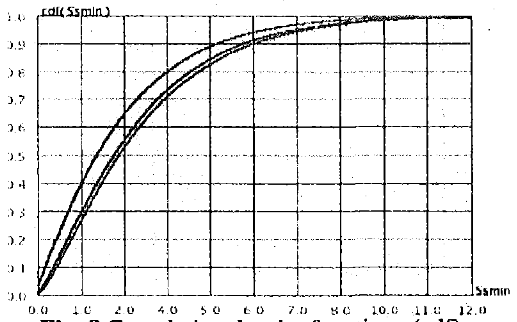

Fig. 8 Cumulative density functions (cdf) of Ss_{min} statistic for N = {50, 500, 5000}, ρ = 0.0

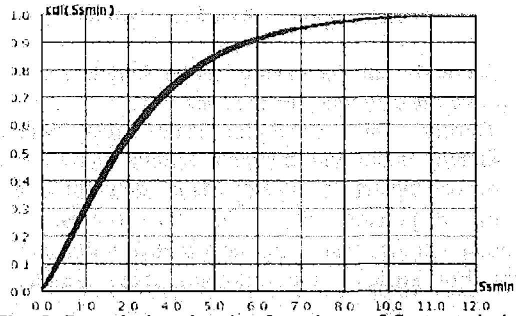

Fig. 9 Cumulative density functions of Ss_{min} statistic for ρ = {0.0, 0.2, 0.4}, N = 1000

5. Stationarity testing

Considering a series x, an extreme statistic value of a bridge (BS_x or BR_x) points out the non stationarity of this series. Thus, a simple statistical test can be based on each statistic previously defined:

- S being the statistic value obtained on the x series of length N and first autocorrelation coefficient ρ,

  $\Delta_{N,\rho}$ being the theoretical distribution of S for series of length N and first autocorrelation coefficient ρ,

  α being the level of significance of the test (for instance 10%),

the stationarity hypothesis is accepted if:

$$cdf_{\Delta_{N,\rho}}(S(B_x)) < 1-\alpha$$

where $cdf_{\Delta_{N,\rho}}$ is the cumulative density function of Δ distribution (see Fig. 10).

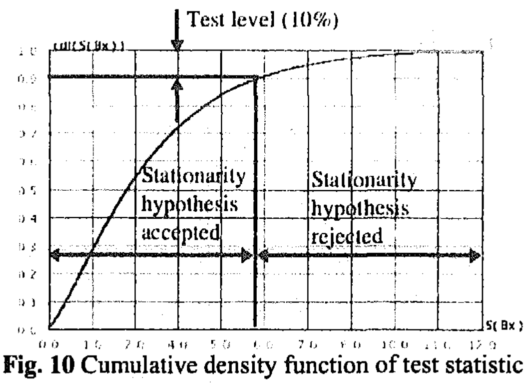

Fig. 10 Cumulative density function of test statistic and stationarity hypothesis acceptance region

6. Transient suppression procedure

Let us define a series x of N observations from which we want to remove the initial transient. Tests are performed at first on the entire series. If the x series is not considered as stationary, the series is truncated (of a certain number of initial observations) and the tests are performed again. This iterative procedure is repeated as long as the test hypothesis is rejected and the number of remaining observations is statistically sufficient.[8] Another way to truncate the series at each iteration could be to suppress $n = t.N$ observations, t being the value (t_{min} or t_{max}) where the considered bridge $B_x(t)$ reaches its supremum.

Of course, the efficiency of this iterative procedure depends on the statistic test performed at each iteration.

The methodology of truncation and various stationarity tests (not only based on Brownian bridge theory) are implemented in the OPOSSUM[9] post-processing tool of the MC code MORET 4[10].

7. Validation

The validation of this transient suppression methodology and the comparison of underlying stationarity tests are performed on a plan of experiments of artificial and although realistic series.

The series are modeled (see Fig. 11) as the sum of a stationary series and of a transient: $x(i) = x_s(i) + T(i)$, where:

- x_s is a stationary Gaussian autoregressive process: $x_s(i+1) = \rho\, x_s(i) + \epsilon(i+1)$, where $(\epsilon(i))_{1 \le i \le N}$ are independent Gaussian random variables of mean 0 and variance σ_ϵ^2. The parameter τ (introduced in the series bridge definition in section 3) of the x_s series equals $\tau = \sigma_\epsilon / (1-\rho)$.
- T is a transient function defined as:

$$T(i) = \begin{cases} A, & for\ 1 \le i \le a \\ A \cdot (b-i)/(b-a), & for\ a \le i \le b \\ 0, & for\ b \le i \le N \end{cases}$$

The series are chosen to reveal different transients, different lengths and different ranges while the stationarity tests have the same level of significance (10%).

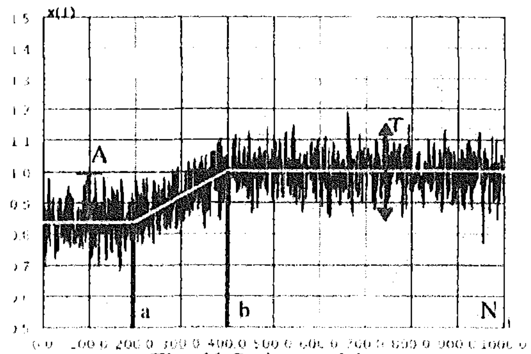

Fig. 11 Series model

The parameters defining the series are (see Fig. 11):

- $N \in \{100, 500, 1000\}$,
- $\frac{a}{N} \in \{0.0, 0.25, 0.5, 0.75\}$,
- $\frac{b}{N} \in \{0.25, 0.5, 0.75\}$ (with $b \geq a$),
- $\tau \in \{0.01, 0.02, 0.03\}$,
- $\rho \in \{0.0, 0.1, 0.2\}$,
- $\overline{T} \in \{-0.00500, -0.01000, -0.02000\}$,

where $\overline{T} = \frac{1}{N}\sum_{i=1}^{N} T(i) = A \cdot \frac{a+b-1}{2N}$ is the mean of the transient (equal to the induced bias).

The efficiency indicators used to compare the tests are $\frac{\overline{T_{n,N}}}{\overline{T}}$ and $\frac{n}{b}$ where n is the number of observations finally truncated and $\overline{T_{n,N}} = \frac{1}{N-n+1}\sum_{i=n}^{N} T(i)$ is the remaining bias. Basically, $\frac{\overline{T_{n,N}}}{\overline{T}}$ is the remaining bias over the initial bias and $\frac{n}{b}$ is the truncation length over the transient length.

For each series parameters set, the truncation is performed on 1000 series to obtain a statistical distribution of the two efficiency indicators. Finally, these distributions obtained with each stationarity test are compared.

Note that Schruben test is performed with an exact value of τ whereas it should be estimated. In the same way, all the tests are performed with an exact value of ρ whereas it should be estimated.

Firstly, regarding efficiency indicators means distributions (see Fig. 12, 13), it appears that Schruben, supremum ratios and Vassilacopoulos tests give good results. However, it seems that Schruben test generally tends to lightly more overestimate the transient length than the two other tests, maybe due to a bad estimation of empirical statistic distributions which alters the levels of significance.

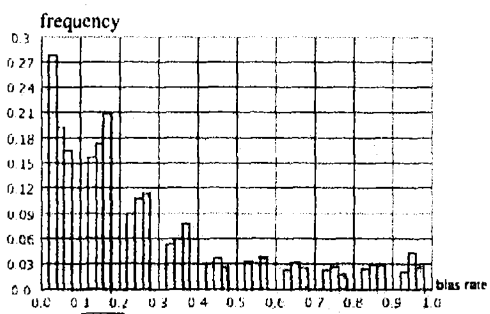

Fig. 12 $\frac{\overline{T_{n,N}}}{\overline{T}}$ (remaining bias over initial bias) distributions for Schruben, supremum ratios and Vassilacopoulos based truncations

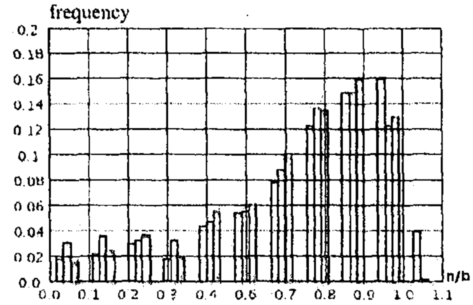

Fig. 13 $\frac{n}{b}$ (truncation length over transient length) distributions for Schruben, supremum ratios and Vassilacopoulos based truncations

Secondly an increasing series length seems to improve the efficiency of the tests.

Thirdly, the supremum ratios test becomes inefficient when the transient length increases and when the transient amplitude (A) is comparable to the series dispersion (τ). On the contrary, the Vassilacopoulos test is the most robust in this extreme case. For instance, the following series parameters: $N=1000$, $\frac{a}{N}=0.5$, $\frac{b}{N}=0.75$, $\tau=0.02$, $\rho=0.2$, $\overline{T}=-0.00500$ give a transient difficult to identify with the naked eye (see Fig. 16) which although biases the series mean estimation of -0.00500. The number of observations suppressed and the remaining bias rate follow these distributions (Fig. 14, 15):

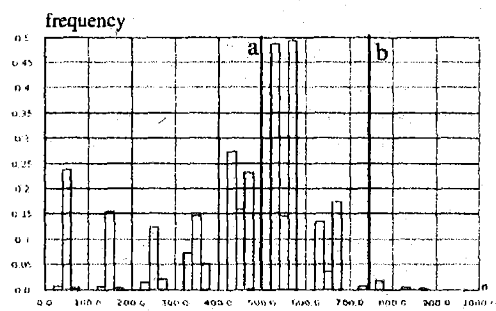

Fig. 14 Statistical distributions of the truncation n using Schruben, supremum ratios and Vassilacopoulos based tests

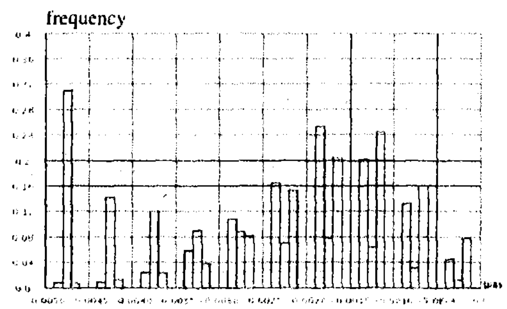

Fig. 15 Statistical distributions of the remaining bias $\overline{T_{n,N}}$ using Schruben, supremum ratios and Vassilacopoulos based tests

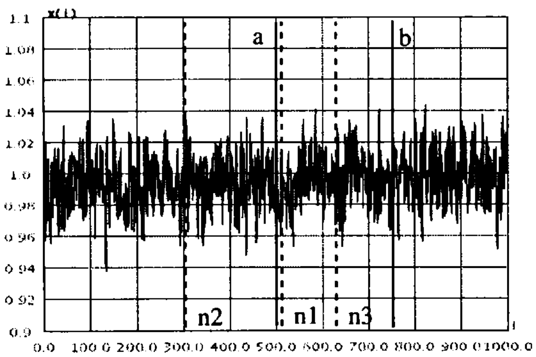

Fig. 16 Series from the plan of experiments with $N=1000$, $\frac{a}{N}=0.5$, $\frac{b}{N}=0.75$, $\tau=0.02$, $\rho=0.2$, $\overline{T}=-0.00500$, leading to a -0.00500 bias on the k-effective estimation

The following Table 1 gives the number of observations suppressed and the remaining bias for each test used with the automated truncation on the series Fig. 16.

Table 1 Truncation and remaining bias

Test used	n	$\overline{T_{n,N}}$
Schruben	n1=511	-0.00186
Supremum ratios	n2=302	-0.00370
Vassilacopoulos	n3=631	-0.00061

This result confirms the efficiency of an automated truncation whereas an empirical truncation would have probably give a bad result in terms of bias gains.

8. Conclusion and prospects

This study on bridges based statistical tests shows good results in terms of transient bias gains, and even when transients are not visible to the naked eye.

This preliminary study on automated transient suppression allowed to validate the iterative truncation methodology and some stationarity tests. However, it is necessary to perform an exhaustive study of all statistics detailed in section 4 (on the series but also on the ranks series). For instance, a Schruben statistic performed on the rank series may combine the efficiency of Schruben and Vassilacopoulos tests that we noticed thanks to the plan of experiments. It will also be necessary to build a more efficient model for statistics distributions than the simple linear interpolation used.

Finally this plan of experiments and a set of practical cases could be used to compare and validate all these improvements.

References

1) R. Blomquist et al., 'OECD/NEA Source Convergence Benchmark Program: Overview and Summary of Results", ICNC 2003.
2) P. Billingsley, "Convergence of probability measures", John Wiley & Sons, 1968.
3) O. Jacquet et al., 'Eigenvalue uncertainty evaluation in MC calculations, using time series Methodologies", in A. Kling et al. (Eds), Advanced Monte Carlo for Radiation Physics, Particle Transport Simulation and Applications, Springer-Verlag, Berlin Heidelberg, 2001.
4) T. Ueki, F.B. Brown, 'Stationarity Diagnostic Using Shannon Entropy in Monte Carlo Criticality Calculation I: F test". ANS Winter Meeting, Washington, Nov 2002.
5) J. Miss, O. Jacquet, A. Nouri, 'Development and Comparison of Monte Carlo Techniques Implemented in the MORET 4 Code for the Calculation of Loosely Coupled Systems", in A. Kling et al. (Eds), Advanced Monte Carlo for Radiation Physics, Particle Transport Simulation and Applications, Springer-Verlag, Berlin Heidelberg, 2001.
6) G. Vassilacopoulos, 'Testing for Initialization Bias in Simulation Output". Simulation Councils (1989)
7) L. Schruben, "Detecting Initialization Bias in Simulation Output". Operations Research, vol. 30, May-June 1982.
8) P. Heidelberger and P. Welch, 'Simulation Run Length Control in the Presence of an Initial Transient". Operations Research, vol. 31, Nov-Dec 1983.
9) L. Heulers et al., "The Graphical User Interface for CRISTAL V1", ICNC 2003.
10) O. Jacquet, J. Miss, G. Courtois, 'MORET: version 4.B, a multigroup Monte Carlo criticality code", ICNC 2003.