Proceedings of the 2002 Winter Simulation Conference
E. Yücesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes, eds.

A BATCH MEANS PROCEDURE FOR MEAN VALUE ESTIMATION OF PROCESSES EXHIBITING LONG RANGE DEPENDENCE

Andrés Suárez-González
José C. López-Ardao
Cándido López-García
Miguel Rodríguez-Pérez
Manuel Fernández-Veiga
María Estrella Sousa-Vieira

ETSE de Telecomunicación
Universidade de Vigo
36200 Vigo, SPAIN

ABSTRACT

Mean value estimation of processes exhibiting *Long Range Dependence* (LRD) requires a different approach than the techniques applied to those exhibiting *Short Range Dependence* (SRD), except for the independent replication method. We describe a nonoverlapping batch means method able to deal with LRD processes, the LRD Batch Means method. This method exploits the behavior of Asymptotically Second-order Self-Similar processes: their aggregated processes become well approximated by *Fractional Gaussian Noise* (FGN) processes for large aggregation levels. Once tested positively this similarity, the method produces a correlation-adjusted confidence interval from an empirical approximation of the distribution of the standardized average for the particular case of FGN processes. Afterwards, we measure its performance over both LRD and SRD processes.

1 INTRODUCTION

Several recent traffic measurement studies have convincingly shown the presence of self-similarity in modern high-speed networks (Leland et al. 1993; Garrett and Willinger 1994) phenomenon characterized by exhibiting LRD. The presence of LRD in input traffic may have a drastic impact on network performance (Likhanov, Tsybakov, and Georganas 1995; Erramilli, Narayan, and Willinger 1996). What is more, it might happen that performance metrics (e.g. delay) inherit the LRD behavior of the input traffic (López-Ardao, Suárez-González, and López-García 1998), opening new statistical problems related to mean value estimation.

Although there are several methods for the mean value estimation of covariance stationary processes (see for example Bratley, Fox, and Schrage 1987), the only "classical" method usable when the process exhibits LRD is "independent replication". However, this method requires the deletion of the transient period in every sample path. Since transient periods are expected to be larger when simulating systems using *Asymptotically Second-order Self-Similar* (ASSS) processes exhibiting LRD than when using classical ones, this method will be usually much less efficient than another one using only one sample path.

Although there exist methods able to estimate the mean value of ASSS processes (Beran 1994, pp. 161–164), they need to make assumptions about the spectral density function of the underlying process and require to store and process the whole sample path. In this paper, we propose the *LRD Batch Means* (LRDBM) method in order to adapt the nonoverlapping batch means method approach to the computation of confidence intervals (CIs) of ASSS processes without any further assumption about the underlying process. Afterwards, we check the performance of the LRDBM method when dealing with both LRD and SRD processes.

The rest of the paper is organized as follows. In Section 2, we review the main concepts related to LRD and self-similarity over which the LRDBM method builds up. In Section 3 we present a brief overview of the LRDBM method. A more complete description of the method may be found in (Suárez-González et al. 2002b). In Section 4 we evaluate experimentally its performance over both LRD and SRD processes. Finally, in Section 5 we summarize the conclusions about the proposed method.

2 LRD AND SELF-SIMILARITY

We limit our exposition to those stochastic processes of interest in simulation studies of computer networks, that is, those discrete-time processes with finite variance and mainly positive autocorrelated (for a more thorough survey see for example Beran 1994).

Let $X = \{X_k; k = 1, 2, \dots\}$ be a covariance stationary stochastic process with autocorrelation function $\mathrm{r}_k \equiv \mathrm{Cov}(X_i, X_{i+k}) / \mathrm{Var}(X)$. It is said that X exhibits LRD when its autocorrelation function is not summable, i.e., $\sum_{k=1}^{\infty} \mathrm{r}_k = \infty$, like in those processes whose autocorrelation function decays hyperbolically:

$$\exists \beta \in (0,1) \left| \lim_{k\to\infty} \frac{\mathrm{r}_k}{k^{-\beta}} = \mathsf{c}_\mathrm{r} \in (0,\infty) \right. . \tag{1}$$

Conversely, it is said that X exhibits SRD when its autocorrelation function is summable, as in those processes whose autocorrelation function either decays hyperbolically as in equation (1) but with $\beta > 1$, or it decays exponentially:

$$\exists \alpha \in (0,1) \left| \lim_{k\to\infty} \frac{\mathrm{r}_k}{\alpha^k} = \mathsf{c}_\mathrm{r} \in (0,\infty) \right. .$$

Let $X^{(l)}$ be the aggregated process of X (with aggregation level l), obtained by averaging the original process X over non-overlapping blocks (batches) of size l, $X^{(l)} = \left\{\overline{X_i}[l]; i = 1, 2, \dots\right\}$, where:

$$\overline{X_i}[l] = \frac{1}{l} \cdot \sum_{j=(i-1)\cdot l+1}^{i\cdot l} X_j,$$

and with autocorrelation coefficients $\mathrm{r}_k^{(l)}$.

The strictly stationary process X is called exactly self-similar, with self-similarity parameter H (defined by Hurst 1951), if for all l, its finite-dimensional distributions are identical to those of the aggregated process, $X^{(l)}$, scaled by $l^{1-\mathsf{H}}$, that is, $X \overset{\mathrm{d}}{=} l^{1-\mathsf{H}} \cdot X^{(l)}$.

Another less strict definition involves exclusively the second-order moments (Cox 1984). So, the covariance stationary process X is called *second-order self-similar* (SSS) if the aggregated process $X^{(l)}$ scaled by $l^{1-\mathsf{H}}$ (from now on we will restrict to $1/2 \le \mathsf{H} < 1$) has the same variance and autocorrelation as X for all l, that is, if the aggregated processes exhibits the same correlation structure as the original stochastic process. Any of the following conditions is sufficient for second-order self-similarity:

- The autocorrelation function for every lag $k \ge 1$ is given by the expression

$$\mathrm{r}_k = \mathrm{g}_k^{\mathsf{H}} \equiv \frac{1}{2}\left[(k+1)^{2\mathsf{H}} - 2k^{2\mathsf{H}} + (k-1)^{2\mathsf{H}}\right] \tag{2}$$

  For $\mathsf{H} = 1/2$ the process is uncorrelated so that $\mathrm{r}_k = 0$ for every $k \ge 1$, whereas for $\mathsf{H} \in (1/2, 1)$ we can see that (Cox 1984):

$$\lim_{k\to\infty} \frac{\mathrm{g}_k^{\mathsf{H}}}{k^{2\mathsf{H}-2}} = \mathsf{H} \cdot (2\mathsf{H} - 1),$$

  that is, it decays hyperbolically as in (1), hence the process exhibits long-range dependence.
- The variances satisfy:

$$\mathrm{Var}\left(X^{(l)}\right) = \mathrm{Var}(X) \cdot l^{2\mathsf{H}-2} \quad \forall l > 1. \tag{3}$$

  Therefore, the classical central limit theorem does not hold for $\mathsf{H} \in (1/2, 1)$.

If the process is Gaussian, self-similarity and second-order self-similarity are equivalent. A FGN process, the sequence of increments of a Fractional Brownian Motion (Mandelbrot and Ness 1968) process, is of this kind, that is, it is Gaussian and exactly self-similar. For $1/2 < \mathsf{H} < 1$ it exhibits LRD, while for $\mathsf{H} = 1/2$ it is simply the classical Gaussian noise process, the sequence of increments of the classical Brownian Motion process.

The process is called asymptotically second-order self-similar (Cox 1984) if expression (2) is satisfied asymptotically as l tends to infinity:

$$\lim_{l\to\infty} \mathrm{r}_k^{(l)} = \mathrm{g}_k^{\mathsf{H}} \quad \forall k \ge 1 \tag{4}$$

Therefore, an ASSS process with $\mathsf{H} \in (1/2, 1)$ will exhibit LRD.

In the same way, an ASSS process also satisfies asymptotically expression (3):

$$\frac{\mathrm{Var}\left(X^{(l\cdot k)}\right)}{\mathrm{Var}\left(X^{(k)}\right)} \xrightarrow[k\to\infty]{} l^{2\mathsf{H}-2} \quad \forall l \ge 1 \tag{5}$$

The following theorem (Tsybakov and Georganas 1997) states a sufficient condition on a stochastic process to be asymptotically second-order self-similar exhibiting LRD:

Theorem 2.1 *A covariance stationary process X with variance σ^2 and autocorrelation function decaying hyperbolically:*

$$\exists \beta \in (0,1) \left| \lim_{k\to\infty} \frac{\mathrm{r}_k}{k^{-\beta}} = \mathsf{c}_\mathrm{r} \in (0,\infty) \right. ,$$

satisfies:

$$\frac{\mathrm{Var}\left(X^{(l)}\right)}{\mathsf{c}_\mathrm{r}\cdot\sigma^2\cdot l^{2\mathsf{H}-2}} \xrightarrow[l\to\infty]{} \frac{1}{\mathsf{H}(2\mathsf{H}-1)} \quad \textit{with} \quad \mathsf{H} = 1-\frac{\beta}{2},$$

and is asymptotically second-order self-similar.

If this process X is also Gaussian, we have the following central limit theorem for self-similar processes (Samorodnitsky and Taqqu 1994):

Theorem 2.2 *If X is an stochastic process satisfying the condition of theorem 2.1, and is also Gaussian with mean value μ, then:*

$$\frac{X^{(l)}-\mu}{\sqrt{\mathrm{Var}\left(X^{(l)}\right)}} \xrightarrow[l\to\infty]{\mathrm{d}} Y,$$

where Y is an standard FGN process (mean value 0 and variance 1).

Although this theorem holds only for Gaussian processes, Taqqu and Teverovsky (1998) claim that it is a good approximation for the non Gaussian case, as it is the case with the M/G/∞ processes (Cox 1984) and non-Gaussian Auto-Regressive Fractional Integrated Moving Average processes (Granger and Joyeux 1980).

At the other hand, for $\mathsf{H} = 1/2$ we have the classical central limit theorem (Billingsley 1968).

3 OVERVIEW OF LRDBM

The LRDBM method is a sequential procedure that tries to estimate the mean value μ of an ASSS process X —with Hurst parameter $\mathsf{H} \in [1/2, 1)$ and variance σ^2— from one sample path of growing length n. It stores a maximum number m of batches with increasingly aggregation level l (power of two) until a goodness of fit test for its spectral density and another one for its marginal distribution are unable to distinguish it from a sample path of an FGN process. Then, when both tests fail to reject the null hypothesis $\mathcal{H}_o \equiv$ "is an FGN process", the method computes a correlation-adjusted estimation of the variance of the average upon an estimation of the Hurst parameter, and finally constructs a CI for the mean value from an empirically fitted distribution of the normalized average of FGN processes.

Since the process X is required to be ASSS, we will have that the autocorrelation structure (hence spectral density) of the batches of the sample path of the process X will tend to that of an SSS process as the aggregation level increases, as given by expression (4). Therefore, the natural approach to the estimation of the Hurst parameter is to use the aggregated Whittle estimator (see for example Leland et al. 1993), that is, the parametric discrete Whittle estimator (see for example Beran 1994, pp. 116–119) with the parametric spectral density of SSS processes, over a sample path of size m of the aggregated process $X^{(l)}$, $\widetilde{\mathsf{H}}[m,l]$. In order to obtain a significant enough estimation of the Hurst parameter (Suárez-González et al. 2002b) we have chosen the number m of stored batches to be 4096.

As goodness of fit test for the spectral density of the aggregated process $X^{(l)}$ we use the "flat zone" goodness of fit test (Suárez-González et al. 2002b) of the null hypothesis "the spectral density of $X^{(l)}$ is that of an SSS process": it computes the absolute value difference of two estimations two steps apart of the Hurst parameter, $\widetilde{\mathsf{H}}[4096,l]$ and $\widetilde{\mathsf{H}}[4096,l/4]$, and uses an empirically computed rejection region $\mathsf{R} = (-\infty,-0.024903) \cup (0.024903,+\infty)$ for an approximate 0.05 significance level. Although heuristic in its nature, we have checked its higher power to reject the null hypothesis for non FGN processes than the more theoretical test proposed by Beran (1994, pp. 201–206).

Once we consider the correlation structure of the aggregated process $X^{(l)}$ well approximated by that of an SSS process, we may estimate the variance of the average, $\sigma^2_{\overline{X}[n]} \equiv \mathrm{Var}\left(\overline{X}[n]\right)$, from property (5):

$$\sigma^2_{\overline{X}[n]} = \mathrm{Var}\left(X^{(m\cdot l)}\right) \xrightarrow[l\to\infty]{} \mathrm{Var}\left(X^{(l)}\right)\cdot m^{2\mathsf{H}-2}, \qquad (6)$$

applying a correlation-adjusted estimation of the variance for an SSS process Y (Beran 1994, p. 156):

$$\widetilde{S^2}_Y\left[n,\widetilde{\mathsf{H}}\right] \equiv S^2_Y[n]\cdot\frac{n-1}{n-n^{2\widetilde{\mathsf{H}}-1}}, \qquad (7)$$

where $S^2_Y[n]$ is the common quasi-variance estimator and $\widetilde{\mathsf{H}}$ is an estimator of the Hurst parameter. Hence, we get:

$$\widetilde{S^2}_{X^{(l)}}\left[m,\widetilde{\mathsf{H}}\right] = S^2_{X^{(l)}}\left[m,\widetilde{\mathsf{H}}\right]\cdot\frac{m-1}{m-m^{2\widetilde{\mathsf{H}}-1}}. \qquad (8)$$

In this way we may derive an estimator of $\sigma^2_{\overline{X}[n]}$ using expressions (6) and (8):

$$\begin{aligned}\widehat{S^2}\left[m,l,\widetilde{\mathsf{H}}\right] &\equiv \widetilde{S^2}_{X^{(l)}}\left[m,\widetilde{\mathsf{H}}\right]\cdot m^{2\widetilde{\mathsf{H}}-2} \\ &= \frac{1-m^{-1}}{m^{2-2\widetilde{\mathsf{H}}}-1}\cdot S^2_{X^{(l)}}[m]\,.\end{aligned} \qquad (9)$$

This estimator is still biased, although much less than the common quasi-variance estimator.

Taqqu and Teverovsky (1998) have shown that the result of theorem 2.2 is a good approximation for non-Gaussian marginal distributions. Therefore, we test the goodness of fit of the marginal distribution of the aggregated process $X^{(l)}$ to a Gaussian distribution, using a Kolmogorov-Smirnov test with composite hypothesis (mean value and variance estimated from the m batches) with a significance level of 0.05 under independence. Of course, the actual significance level when the null hypothesis is true (X is an FGN process) will not match this value, although it will tend as m increases to a number different of the unity (Beran 1994, pp. 197–201) for any value of H.

Once we have checked both the goodness of fit of the spectral density by an SSS process and the marginal distribution by a Gaussian random variable, and in order to compute a CI for the mean value, we approximate the distribution of the standardized average:

$$Z \equiv \frac{\overline{X}[n] - \mu}{\sqrt{\widehat{S^2}\left[m, l, \widetilde{\mathsf{H}}\right]}},$$

considering $X^{(l)}$ an FGN process. Nevertheless, since we will be using a moderate number of blocks m (both for quick convergence and low storage requirement), our estimators —$\widetilde{\mathsf{H}}$ and $\widehat{S^2}\,[m, l]$— will not be consistent (we will not let m tend to infinity as n does), and so, a standard normal (as given by theorem 2.2 for Gaussian processes) will not be a good enough approximation of the distribution of Z, even if X is actually an FGN process. Beran (1994, p. 163) proposes to approximate analytically the distribution of Z, through the distribution of some random variables function of two independent Gaussian random variables, whose quantiles have to be estimated by Monte Carlo simulation, though. Instead of this approach, we have conducted an intensive Monte Carlo simulation of the actual Z variable with X an actual FGN process for different values of the Hurst parameter, and fitted their empirical distribution by a t-Student one. In this way, we limit the stored information to a table of freedom degrees.

3.1 LRDBM Implementation

In practice, we have implemented the LRDBM method with some modifications in order to get either the first CI as soon as possible, or a higher quality CI.

First, the number of batches used to compute the estimator of the variance of the average given by expression (9), m_σ, will be the power of two which minimizes its bias and for which the Kolmogorov-Smirnov test fails to reject the null hypothesis for $X^{(l_\sigma)}$ ($n = l_\sigma \cdot m_\sigma$). Also, as the storage size will be $m = 4096 > m_\sigma$, we will be able to compute another estimation $\widehat{S^2}\left[m_\sigma, l_\sigma, \widetilde{\mathsf{H}}\right]$ without the need of l_σ to be a power of two. As a matter of fact, m_σ will be a function of the number of batches selected to estimate H, m_{H}, as stated below.

With respect to the empirically adjusted distribution of Z, we have fitted it by a t-Student for $m_{\mathsf{H}} \in \{2^k; k = 10, 11, 12\}$, $m_\sigma \in \{2^j; j = 6, \ldots, 10\}$ and $\mathsf{H} \in \{0.5 + i \cdot 0.05; i = 1, \ldots, 9\}$, selecting its number of degrees of freedom such that it minimizes the maximum absolute error between its 100 quantiles and those obtained empirically from 72000 sample paths of each FGN process, generated using the exact method proposed by Crouse and Baraniuk (1999). In this same Monte Carlo study, we have also estimated the value of m_σ which minimizes the bias of the estimator from equation (9), giving rise to the values $m_\sigma = 64$ for $m_{\mathsf{H}} \in \{1024, 2048\}$ and $m_\sigma = 128$ for $m_{\mathsf{H}} = 4096$. For these three $(m_\sigma, m_{\mathsf{H}})$ pairs, the worst (best) fitted case has been for $\mathsf{H} = 0.95$ (0.95), $m_\sigma = 64$ (128) and $m_{\mathsf{H}} = 1024$ (4096) with a maximum absolute error among the 100 quantile pairs (empirical versus t-Student one) of 0.0697888 (0.0198527) for 7 (16) degrees of freedom. We also let m_σ to be 32 for $m_{\mathsf{H}} = 1024$, since the estimator of the variance of the average with this pair exhibits an acceptable small bias. The number of degrees of freedom of the t-Student distribution actually used to compute the CIs is obtained by linear interpolation for $\widetilde{\mathsf{H}}$ of the stored values for m_σ and m_{H} completed with $m_\sigma - 1$ for $\mathsf{H} = 0.5$ and 1 for $\mathsf{H} = 1$.

Second, since the "flat zone" test implies a quite large initial sample size n for a first CI computation, and given that for some ASSS processes with large low-lag autocorrelation structure the estimator $\widetilde{\mathsf{H}}\,[4096, l]$ tends initially to decrease (Suárez-González et al. 2002b), we will try to detect this behavior and permit to compute a CI without passing the "flat zone" test and possibly using a higher aggregation level for the Hurst parameter estimation ($\widetilde{\mathsf{H}}\,[m_{\mathsf{H}}, n/m_{\mathsf{H}}]$ with m_{H} equal to 1024, 2048 or 4096). This way we try to compute conservative CIs with less sample values than the required for the "flat zone" test to fail to reject.

Third, since for some ASSS processes with low low-lag autocorrelation structure the estimator $\widetilde{\mathsf{H}}\,[4096, l]$ tends initially to increase giving rise to low-quality CIs the first time the "flat zone" test fails to reject the null hypothesis (Suárez-González et al. 2002b), we will try to detect this behavior and react using a more pessimistic Hurst estimator: $\max\left(\widetilde{\mathsf{H}}\,[4096, l]\,, \widetilde{\mathsf{H}}\,[2048, 2l]\,, \widetilde{\mathsf{H}}\,[1024, 4l]\right)$. This way we try to counteract the subestimation of the correlation structure (hence the estimation of the variance of the average) and produce higher quality CIs.

The resulting LRDBM implementation is made up by two algorithms:

- Every time the sample size n is a power of two multiple of the number of stored batches $m = 4096$, the Algorithm 1 verifies the expected behavior of

$\widetilde{\mathsf{H}}[4096, l]$ and updates accordingly the Hurst estimation, $\widetilde{\mathsf{H}}$, the number of batches used to compute the variance of the average, m_σ, and the next number of batches to compute the Hurst parameter, m_H.

- Every time the sample size n is, either a power of two multiple of the number of stored batches $m = 4096$, or it is higher than the next sample size check point and, at the same time, a multiple of both m_σ and the new l aggregation level for the next $m = 4096$ batches, the Algorithm 2 computes a new estimation of the variance of the average, the degrees of freedom of the t-Student distribution used to approximate the distribution of Z, and finally the half-width of the new CI for the desired quality $1-\alpha$. If this half-width is lower than the desired half-width relative to the average, then it returns its value and quits. If not, it computes a new gross estimation of the needed samples and selects the midpoint between it and the current sample size as the next check point.

The source code is available at `<ftp://ftp-gris.det.uvigo.es/pub/LRD/LRDBM-src.tgz>`.

4 PERFORMANCE MEASURE

In this section we present the measured performance of the LRDBM method applied over some selected LRD and SRD processes. For those processes exhibiting SRD, we also show the behavior of the *Automated Simulation Analysis Procedure* (ASAP) procedure proposed by (Steiger and Wilson 2000), since it is also an autocorrelation corrected batch means procedure which gives similar or better CI quality than other nonoverlapping batch means methods.

For the first set of tests, we select a process that arises naturally in teletraffic as the limiting case for aggregation of on/off sources (Likhanov, Tsybakov, and Georganas 1995), the M/G/∞ process. Moreover, its sample generation method is exact and quite efficient —$\mathcal{O}(n)$ with n the sample size— and, even more useful, at the same time it is a sequential procedure (its sample size has not to be set beforehand). Specifically, we use the discrete-time process proposed by Suárez-González et al. (2002a), M/S/∞, since its correlation structure is simple and quite flexible (two parameters).

As the second set of tests (exclusively SRD) we have selected those queue waiting time processes where Steiger and Wilson (2000) have measured the worst coverage for ASAP. Namely, an M/M/1 queue with *Last Input First Output* (LIFO) service and a computer model with one central server (CPU) and two peripheral units, both previously used by Law and Carson (1979).

Algorithm 1 Compute $\widetilde{\mathsf{H}}$, m_H and m_σ

Require: $n = 2^{12+k}$
  $m \leftarrow 2^{12} = 4096$
  $l \leftarrow 2^k = n/m$
  {set initial values}
  if $n = 4096$ **then**
    $maxm_\sigma, m_\sigma, grow, flatzone, incm_\mathsf{H} \leftarrow 0$
    $nextm_\mathsf{H} \leftarrow 1024$
  {compute highest m_σ value}
  for all $p \in \{5, 6, 7\}$ such that $2^p > maxm_\sigma$ **do**
    if K_S$\left(X^{(2^{12+k-p})}\right)$ fails to reject $\mathcal{H}_o$ **then**
      $maxm_\sigma \leftarrow 2^p$
  {check behavior of $\widetilde{\mathsf{H}}[m, l]$ estimator}
  if $\left|\widetilde{\mathsf{H}}[m, l] - \widetilde{\mathsf{H}}[m, l/4]\right| < 0.024903$ **then**
    $flatzone \leftarrow flatzone + 2$
  else
    if $flatzone > 0$ **then**
      $flatzone \leftarrow flatzone - 1$
    if $\widetilde{\mathsf{H}}[m, l] > \widetilde{\mathsf{H}}[m/2, 2l] > \widetilde{\mathsf{H}}[m/4, 4l]$ **then**
      $grow \leftarrow grow - 1$
    else if $\widetilde{\mathsf{H}}[m, l] < \widetilde{\mathsf{H}}[m/2, 2l] < \widetilde{\mathsf{H}}[m/4, 4l]$ **then**
      $grow \leftarrow grow + 1$
  {compute m_σ and new $\widetilde{\mathsf{H}}$ estimation if possible}
  $m_\mathsf{H} \leftarrow nextm_\mathsf{H}$
  if $m_\mathsf{H} = 4096$ **then**
    $optm_\sigma \leftarrow 128$
  else
    $optm_\sigma \leftarrow 64$
  if $grow < 0$ or $flatzone > 0$ **then**
    $m_\sigma \leftarrow \min\left(maxm_\sigma, optm_\sigma\right)$
  if $grow \leq 0$ **then**
    $\widetilde{\mathsf{H}} \leftarrow \widetilde{\mathsf{H}}[m_\mathsf{H}, n/m_\mathsf{H}]$
  else
    $\widetilde{\mathsf{H}} \leftarrow \max\left(\widetilde{\mathsf{H}}[k, n/k] : m_\mathsf{H} \leq k = 2^i \leq m\right)$
  {adjust next m_H}
  if $m_\mathsf{H} < m$ **then**
    if $grow < 0$ **then**
      if $\widetilde{\mathsf{H}}[m_\mathsf{H}, n/m_\mathsf{H}] \geq \widetilde{\mathsf{H}}\left[2m_\mathsf{H}, \frac{n}{2m_\mathsf{H}}\right]$ **then**
        $incm_\mathsf{H} \leftarrow 1$
    else if $flatzone > 0$
    and $\widetilde{\mathsf{H}}[m_\mathsf{H}, n/m_\mathsf{H}] < \widetilde{\mathsf{H}}\left[2m_\mathsf{H}, \frac{n}{2m_\mathsf{H}}\right]$ **then**
      $incm_\mathsf{H} \leftarrow 1$
  if $incm_\mathsf{H} = 1$ and
  ($maxm_\sigma > optm_\sigma$ or ($m_\mathsf{H} = 1024$ and $maxm_\sigma = 64$))
  then
    $nextm_\mathsf{H} \leftarrow 2m_\mathsf{H}$
    $incm_\mathsf{H} \leftarrow 0$

Algorithm 2 Compute CI

Require: $m_\sigma \neq 0$ and (
$n = 2^{12+k}$
or ($n >$ *nextn* and $n \div l$ and $n \div m_\sigma$))
$\{l = 2^i$ with $i = \lceil \log_2 n \rceil - 12\}$
$l_\sigma \leftarrow n/m_\sigma$
$var \leftarrow \widehat{S^2}[m_\sigma, l_\sigma]$ {use $\widetilde{\mathsf{H}}$ computed in Algorithm 1}
$df \leftarrow$ degrees_of_freedom $\left(m_\sigma, m_{\mathsf{H}}, \widetilde{\mathsf{H}}\right)$
$t \leftarrow$ t_Student $(1 - \alpha/2, df)$
$rhalfCI \leftarrow t \cdot \sqrt{var}/\overline{X}[n]$
{check quality requirements}
if $rhalfCI < goal\_rhCI$ **then**
return *rhalfCI* and quit
else
$nextn \leftarrow (n +$ extrapolate$(n, rhalfCI, goal_rhCI))/2$

For all the experiments of this section, 1000 steady state simulations have been run for any given set of parameters. The set of requirements for the half width of the 90% CI relative to the mean value estimation has been {10, 4%, 2%, 1%}. The 95% confidence interval for the coverage (given a method with exact CI estimation for the mean value) is (0.884, 0.916), considering approximately Gaussian the binomial distribution with $\mathsf{p} = 0.9$ and $\mathsf{n} = 1000$. None of the experiments of this section has shown an estimated coverage of the LRDBM procedure lower than 0.884. The tables for each experiment show for each CI requirement (rhCI req.), the average sample size ($\overline{n}$), the estimated coverage (cvrg), the average half width of the 90% CI relative to the mean value estimation (av.rhCI), its estimated coefficient of variation ($\mathrm{C_{rhCI}}$), that is, the estimated $\sqrt{\mathrm{Var}\,(\mathrm{rhCI})}/\,\mathrm{E}\,(\mathrm{rhCI})$ value, and, in the LRDBM case, the estimated bias ($\widetilde{\mathsf{H}} - \mathsf{H}$) shown by the used Hurst estimation (av.$\widetilde{\mathsf{H}}$bias).

4.1 M/S/∞ Processes

The autocorrelation structure of these family of processes is given by:

$$\mathrm{r}_k = \begin{cases} 1 - \dfrac{\alpha - 1}{\mathsf{m}\alpha} \cdot k & \forall k \in (0, \mathsf{m}] \\ \dfrac{1}{\alpha} \cdot \left(\dfrac{\mathsf{m}}{k}\right)^{\alpha - 1} & \forall k \geq \mathsf{m}, \end{cases}$$

hence the process is ASSS exhibiting LRD of Hurst parameter $\mathsf{H} = (3 - \alpha)/2$ for $1 < \alpha < 2$ from Theorem 2.1 ($\beta = \alpha - 1$), and it exhibits SRD ($\mathsf{H} = 0.5$) for $\alpha > 2$. The parameter $\mathsf{m} > 0$ is selected to get a desired $\mathrm{r}_1 \in (0, 1)$. In any case, the marginal distribution of any M/G/∞ process with arbitrary service time random variable of finite mean value is Poisson.

The mean value selected for all the M/S/∞ processes generated was 20. All the simulations have been started in steady state. For the LRD cases, we have selected $\mathrm{r}_1 = 0.9$ (0.1) and $\mathsf{H} = 0.6$ (0.8) as representative of strong (weak) correlation and of weak (strong) LRD behavior. For the SRD case, we have selected $\mathrm{r}_1 = 0.9$ and $\alpha = 2.1$, as representative of strong low-lag correlation and slow correlation decay.

Table 1 shows the performance of the LRDBM method for an M/S/∞ process with $\mathsf{H} = 0.6$ and $\mathrm{r}_1 = 0.9$. All of the CIs have been computed after forecasting a decreasing behavior of the Hurst estimation (*grow* < 0 in Algorithm 1). This decreasing behavior is a direct consequence of the autocorrelation structure of the process being higher than that of the SSS process of the same Hurst parameter: $\mathrm{r}_i > \mathrm{g}_i^{0.6} \quad \forall i > 1$. As shown by the average sample size for the non-requirement case ($\simeq 5000$), this behavior is detected quite quickly.

We observe in this and all the following tables, how the estimated bias of the Hurst estimator decreases with the average sample size (we get closer to the asymptotic behavior of the process), as expected. This behavior of the Hurst estimation drives that of the estimated coverage: it tends to be the nearer to the asked CI quality (90%) when the bias of the former is lower. Nevertheless, there is a light glitch of this behavior in Table 1 (also in Table 5 below), when we compare the $\pm 2\%$ and $\pm 1\%$ CI requirement entries. This has almost certainly been due to the stochastic nature of the estimators, since from the total 58 CIs that failed to contain the actual mean value for these two cases, there were only 4 of them common to both cases (same seed, hence same sample path). At the same time, we observe that the estimated variation coefficient of the computed CI relative to the average ($\mathrm{C_{rhCI}}$) also diminishes with the sample size.

Table 1: M/S/∞ Process with $\mathsf{H} = 0.6$ and $\mathrm{r}_1 = 0.9$

rhCI req.	none	$\pm 4\%$	$\pm 2\%$	$\pm 1\%$
$\overline{n}$	$5.0 \cdot 10^3$	$3.1 \cdot 10^4$	$9.5 \cdot 10^4$	$4.1 \cdot 10^5$
cvrg	1.0	0.986	0.966	0.978
av.rhCI	$\pm 79.4\%$	$\pm 3.63\%$	$\pm 1.88\%$	$\pm 0.95\%$
$\mathrm{C_{rhCI}}$	0.514	0.102	0.0622	0.0477
av.$\widetilde{\mathsf{H}}$bias	0.361	0.142	0.088	0.055

Table 2 shows the performance of the LRDBM method for an M/S/∞ process with $\mathsf{H} = 0.8$ and $\mathrm{r}_1 = 0.1$. Conversely to the previous process, this one has a correlation structure always lower than a second-order self-similar process of the same Hurst parameter, that is, $\mathrm{r}_i < \mathrm{g}_i^{0.8} \quad \forall i > 1$. This fact causes the increasing behavior of the Hurst estimation, and, as a consequence, all of the CIs in this experiment have been computed after the "flat zone" test has been unable to reject the null hypothesis (*flatzone* > 0 and *grow* ≥ 0 in Algorithm 1). This is a harder condition to achieve than that of decreasing $\widetilde{\mathsf{H}}$ behavior, and this explains the high value of the average sample size for the first computable

CI. As a matter of fact, already in this non-requirement case the half length CI relative to the average is lower than our stricter requirement (±1%). There are actually 708 out of the 1000 experiments which are identical for each one of the requirements, that is, their computed relative half length CI in the non-requirement case is already lower than ±1%. We may observe also that the proposed method is able to compute good quality CIs even though the Hurst estimation is negatively biased.

Table 2: M/S/∞ Process with $\mathsf{H} = 0.8$ and $r_1 = 0.1$

rhCI req.	none	±4%	±2%	±1%
$\overline{n}$	$3.2 \cdot 10^6$	$3.2 \cdot 10^6$	$3.3 \cdot 10^6$	$4.4 \cdot 10^6$
cvrg	0.897	0.897	0.897	0.901
av.rhCI	±0.89%	±0.89%	±0.88%	±0.76%
C_{rhCI}	0.498	0.492	0.438	0.254
av.$\widetilde{\mathsf{H}}$bias	−0.038	−0.038	−0.038	−0.038

Table 3 shows the performance of the LRDBM and ASAP methods over an M/S/∞ process with $\alpha = 2.1$ and $r_1 = 0.9$, that is, over an SRD M/S/∞ process. The behavior of the LRDBM method in this case is qualitatively equal to the case showed in Table 1, although the weaker autocorrelation structure in this SRD case implies a lower average sample size for any given requirement. The ASAP method has clear trouble keeping pace with this extreme correlation structure, although the process exhibits SRD.

Table 3: M/S/∞ Process with $\alpha = 2.1$ and $r_1 = 0.9$

rhCI req.	none	±4%	±2%	±1%
		LRDBM		
$\overline{n}$	$6.6 \cdot 10^3$	$1.8 \cdot 10^4$	$4.0 \cdot 10^4$	$1.1 \cdot 10^5$
cvrg	1.0	0.996	0.991	0.978
av.rhCI	±44.6%	±3.34%	±1.8%	±0.939%
C_{rhCI}	1.02	0.14	0.0941	0.0626
av.$\widetilde{\mathsf{H}}$bias	0.397	0.218	0.159	0.112
		ASAP		
$\overline{n}$	$1.6 \cdot 10^3$	$2.0 \cdot 10^3$	$7.6 \cdot 10^3$	$3.0 \cdot 10^4$
cvrg	0.811	0.797	0.824	0.808
av.rhCI	±4.01%	±3.54%	±1.87%	±0.946%
C_{rhCI}	0.216	0.116	0.0865	0.0693

4.2 Queue Waiting Time

The first queue system is the M/M/1/LIFO. We have taken $\rho = 0.8$ as utilization of the server, and $\mathrm{E}(S) = 4$ as the mean value of the service time random variable, yielding a mean queue waiting time (between arrival to the system queue and receiving service) of 16. Each one of the simulations has been started in the empty state.

Table 4 shows the behavior of both the LRDBM and ASAP methods applied to this system, estimating the mean value of the queue waiting time process. In this case, 770 out of the 1000 first CIs computed by the LRDBM method had to wait for the Kolmogorov-Smirnov test to fail to reject the Gaussian hypothesis, with 49438.72 as average sample size in all the 1000 experiments for the first time to fail. In fact, the Kolmogorov-Smirnov test has failed only for a maximum value of $maxm_\sigma = 32$ in 910 times out of these 1000 experiments (first CI), while in the M/S/∞ experiments it has happened ($maxm_\sigma = 32$) less than 7 times out of every 1000 set of experiments. Obviously, the M/S/∞ process has an almost Gaussian distribution (Poisson), while the marginal distribution of queue waiting time processes has an exponential tail (see for example Kleinrock 1975). Comparing both the LRDBM and ASAP methods in this case: initially, the former has a mildly conservative behavior while the latter clearly gets a quite low coverage; asymptotically, both have a fairly similar performance.

Table 4: M/M/1/LIFO Queue Waiting Time Process

rhCI req.	none	±4%	±2%	±1%
		LRDBM		
$\overline{n}$	$7.2 \cdot 10^4$	$3.6 \cdot 10^5$	$1.1 \cdot 10^6$	$3.8 \cdot 10^6$
cvrg	0.963	0.934	0.911	0.906
av.rhCI	±19.5%	±3.7%	±1.88%	±0.953%
C_{rhCI}	0.577	0.0802	0.0559	0.0426
av.$\widetilde{\mathsf{H}}$bias	0.167	0.043	0.017	0.009
		ASAP		
$\overline{n}$	$5.3 \cdot 10^3$	$1.9 \cdot 10^5$	$9.0 \cdot 10^5$	$3.7 \cdot 10^6$
cvrg	0.75	0.848	0.904	0.875
av.rhCI	±20.9%	±3.8%	±1.89%	±0.944%
C_{rhCI}	0.311	0.0606	0.0609	0.0534

The second queue system is a computer model with one CPU and two peripheral units, the third of the four cases of the second computer model used by Law and Carson (1979). The system has a fixed number of jobs $\mathsf{N} = 8$. The probability that the job is routed to the first (second) peripheral unit is 0.9 (0.1). All service times are exponentially distributed. The service rate at the CPU is 1.0, and at the first (second) peripheral unit is 0.45 (0.05). After getting service at one of the peripheral units, the job leaves the system and is immediately replaced by another job joining the CPU queue. The process whose mean value will be estimated is made up of the jobs response time (between arrival at the CPU queue and departure from the system). The mean response time with these parameters is 18.279. The initial state is 5 jobs at the CPU and 1 (2) jobs at the first (second) peripheral unit.

Table 5 shows the performance of both the LRDBM and ASAP methods in this case. In the LRDBM case, only 185 of the 1000 experiments without CI requirement have waited for the Kolmogorov-Smirnov test to fail. Clearly, the tail behavior of the marginal distribution of the queue waiting time process will be less strong, since we have a limited number of jobs in the system. Of the remaining 815 cases, 62 have computed its first CI after forecasting a decreasing $\widetilde{H}$ behavior, while the remaining 753 cases have waited to the "flat zone" test to fail to reject the null hypothesis. The used Hurst estimator tends to begin increasing, for afterwards start decreasing. This behavior has also been observed, although less extended in time, in the previous M/M/1/LIFO queue system. As a side effect, this tends to force the method to use $m_H = 4096$ (414 out of the 1000 simulations for $\pm 1\%$ requirement) before it really has arrived to an stable behavior (estimated bias for the Hurst estimation is still 0.054), producing an over pessimistic CI. Comparing both methods, the behavior of the LRDMB method is quite conservative, while the ASAP method gets nearer to the desired coverage but from an optimistic approach.

Table 5: Computer Model Response Time Process

rhCI req.	none	$\pm 4\%$	$\pm 2\%$	$\pm 1\%$
		LRDBM		
$\overline{n}$	$7.8 \cdot 10^4$	$7.9 \cdot 10^4$	$1.1 \cdot 10^5$	$2.8 \cdot 10^5$
cvrg	0.95	0.944	0.959	0.947
av.rhCI	$\pm 2.9\%$	$\pm 2.7\%$	$\pm 1.78\%$	$\pm 0.92\%$
C_{rhCI}	0.426	0.34	0.157	0.101
av.$\widetilde{H}$bias	0.086	0.086	0.086	0.054
		ASAP		
$\overline{n}$	$2.5 \cdot 10^3$	$8.3 \cdot 10^3$	$3.3 \cdot 10^4$	$1.5 \cdot 10^5$
cvrg	0.825	0.816	0.832	0.882
av.rhCI	$\pm 7.44\%$	$\pm 3.71\%$	$\pm 1.9\%$	$\pm 0.954\%$
C_{rhCI}	0.259	0.0846	0.0674	0.0569

5 CONCLUSIONS

In this paper we have presented the LRDBM method designed to perform the steady-state mean estimation of LRD processes, although it is also applicable to SRD processes. Actually, the LRDBM implicitly performs a test of independence, that is, when the Hurst estimator takes the value 0.5, but it does not fix its result for the rest of the simulation. Although its approach is heuristic, its quality is supported by a quite intensive Monte Carlo simulation study using an exact method of generation of FGN samples (72000 per process).

Unlike other batch means methods proposed for SRD processes, the LRDBM method tries to keep itself in the safe side and produce conservative CIs. This is quite important when dealing with LRD processes, but may be a desirable property even when the LRDBM method is applied to SRD processes: it is not clear the utility of a mean estimation method whose real coverage gets clearly lower than the asked quality in quite simple simulation studies, where we are able to compute the actual mean value. This fact also has been pointed out by Steiger and Wilson (2000), and will be addressed by a future modification of the ASAP procedure.

We have measured the performance of the LRDBM method for a selected set of LRD and SRD processes. In these tests, the LRDBM method has performed well, producing conservative CIs, and even getting quite accurate coverage for a classical tough test, the M/M/1/LIFO queue system.

Nevertheless, the LRDBM method may be further improved at least in the following ways: (i) limiting the history for the forecasting of the Hurst estimator behavior, in order to discard a possible noisy initial transient period, like that observed in the studied computer model; (ii) studying the usefulness of other Hurst estimators, in order to use an even lower number of batches than 1024; (iii) trying to extrapolate the previous computed values of the Hurst estimator in order to be able to compute sooner a CI, mainly when it exhibits an increasing behavior.

ACKNOWLEDGMENTS

This work was partially supported by the project TIC2000-1126 of the *Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica*.

REFERENCES

Beran, J. 1994. *Statistics for long-memory processes*, Volume 61 of *Monographs on Statistics and Applied Probability*. Chapman & Hall.

Billingsley, P. 1968. *Convergence of probability measures*. New York: John Wiley & Sons.

Bratley, P., B. L. Fox, and L. E. Schrage. 1987. *A guide to simulation*. 2nd ed. Springer-Verlag.

Cox, D. R. 1984. Long range dependence: A review. In *Statistics: An Appraisal*, ed. H. A. David and H. T. David, 55–74. Iowa State University Press.

Crouse, M. S., and R. G. Baraniuk. 1999. Fast, exact synthesis of Gaussian and nonGaussian long-range-dependent processes. Submitted to IEEE Transactions on Information Theory. Draft version available at <http://www.dsp.rice.edu/publications/pub/fftMC.ps.gz>.

Erramilli, A., O. Narayan, and W. Willinger. 1996. Experimental queueing analysis with long-range dependent

packet traffic. *IEEE/ACM Transactions on Networking* 4:209–223.

Garrett, M. W., and W. Willinger. 1994. Analysis, modeling and generation of self-similar VBR video traffic. In *Proc. ACM SIGCOMM '94*, 269–280. London, UK.

Granger, C. W. J., and R. Joyeux. 1980. An introduction to long-range time series models and fractional differencing. *Journal of Time Series Analysis* 1:15–30.

Hurst, H. E. 1951. Long-term storage capacity of reservoirs. *Trans. Am. Soc. Civil Engineers* 116:770–799.

Kleinrock, L. 1975. *Queueing systems*. Wiley & Sons.

Law, A. M., and J. S. Carson. 1979. A sequential procedure for determining the length of a steady-state simulation. *Operations Research* 27:1011–1025.

Leland, W. E., M. S. Taqqu, W. Willinger, and D. V. Wilson. 1993. On the self-similar nature of Ethernet traffic. *Computer Communication Review* 23:183–193.

Likhanov, N., B. Tsybakov, and N. Georganas. 1995. Analysis of an ATM buffer with self-similar ("fractal") input traffic. In *Proc. IEEE INFOCOM '95*, 985–992. Boston, MA.

López-Ardao, J., A. Suárez-González, and C. López-García. 1998. On the effects of long-range dependence on the performance of FDDI. In *Proc. 7th. IEEE Workshop on Computer-Aided Modeling, Analysis and Design of Communication Links and Networks (CAMAD'98)*. Sao Paulo, Brazil.

Mandelbrot, B. B., and J. W. V. Ness. 1968, October. Fractional brownian motions, fractional noises and applications. *SIAM Review* 10 (4): 422–437.

Samorodnitsky, G., and M. S. Taqqu. 1994. *Stable non-gaussian random processes: Stochastic models with infinite variance*. Chapman & Hall.

Steiger, N. M., and J. R. Wilson. 2000. Experimental performance evaluation of batch means procedures for simulation output analysis. In *Proceedings of the 2000 Winter Simulation Conference*, ed. J. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, Volume 1, 627–636.

Suárez-González, A., J. López-Ardao, C. López-García, M. Fernández-Veiga, R. Rodríguez-Rubio, and M. Sousa-Vieira. 2002a. A new heavy-tailed discrete distribution for LRD M/G/∞ sample generation. *Performance Evaluation* 47 (2–3): 197–219.

Suárez-González, A., J. López-Ardao, C. López-García, M. Fernández-Veiga, and M. Sousa-Vieira. 2002b. Sequential mean estimation of LRD processes. In preparation, draft available at available at <ftp://ftp-gris.det.uvigo.es/pubLRD/MeanLRD.ps.gz>.

Taqqu, M. S., and V. Teverovsky. 1998. On estimating the intensity of long-range dependece in finite and infinite variance time series. In *A Practical Guide to Heavy Tails*, ed. R. Adler, R. Feldman, and M. S. Taqqu. Birkhauser.

Tsybakov, B., and N. D. Georganas. 1997. On self-similar traffic in ATM queues: Definitions, overflow probability bound, and cell delay distribution. *IEEE/ACM Transactions on Networking* 5 (3): 397–409.

AUTHOR BIOGRAPHIES

ANDRÉS SUÁREZ-GONZÁLEZ is an associate professor in the *Departamento de Enxeñería Telemática* at *Universidade de Vigo*. He received a Ph.D. degree in telecommunication engineering from *Universidade de Vigo* in 2000. He is a member of ACM. His current research interests include simulation methodology and analysis of stochastic systems. His e-mail address is <asuarez@det.uvigo.es>.

JOSÉ C. LÓPEZ-ARDAO is an associate professor in the *Departamento de Enxeñería Telemática* at *Universidade de Vigo*. He received a Ph.D. degree in telecommunication engineering from *Universidade de Vigo* in 1999. He is a member of ACM.

CÁNDIDO LÓPEZ-GARCÍA is an associate professor in the *Departamento de Enxeñería Telemática* at *Universidade de Vigo*. He received a Ph.D. degree in telecommunication engineering from *Universidad Politécnica de Madrid* in 1995.

MIGUEL RODRÍGUEZ-PÉREZ is an assistant researcher in the *Departamento de Enxeñería Telemática* at *Universidade de Vigo*. He received a telecommunication engineering degree from *Universidade de Vigo* in 2001. Nowadays, he is working toward his Ph.D. in the *Departamento de Enxeñería Telemática* at *Universidade de Vigo*.

MANUEL FERNÁNDEZ-VEIGA is an associate professor in the *Departamento de Enxeñería Telemática* at *Universidade de Vigo*. He received a Ph.D. degree in telecommunication engineering from *Universidade de Vigo* in 2001. He is an associate member of IEEE.

MARÍA ESTRELLA SOUSA-VIEIRA is an assistant professor in the *Departamento de Enxeñería Telemática* at *Universidade de Vigo*. She received a telecommunication engineering degree from *Universidade de Vigo* in 1998. Since then, she is working toward her Ph.D. in the *Departamento de Enxeñería Telemática* at *Universidade de Vigo*.