Comparing overlapping batch means and standardized time series under multiple replications in parallel

Edjair Mota[*], Adam Wolisz
Technical University of Berlin, Germany
Krzysztof Pawlikowski
University of Canterbury, Christchurch, New Zealand
e-mail : mota@ft.ee.TU-Berlin.DE

Keywords : distributed simulation simulation, multiple replications, parallelization, computer networks, sequential analysis

ABSTRACT

The main drawback of quantitative stochastic simulation approach is the usually prohibitive amount of computer time necessary to give reasonable results. *Multiple Replications in Parallel* (MRIP) aims to attack this issue by generating data in parallel, asynchronously, through a local area network of workstations. We investigated the performance of sequential confidence interval procedures based on overlapping batch means and standardized time series, when they are used to estimate mean values of steady-state processes under Akaroa-2, an MMRIP implementation. Reasoning of both methods are explained and sequential versions proposed and implemented as well.

1 INTRODUCTION

The main drawback of quantitative stochastic simulation approach is the usually prohibitive amount of computer time necessary to give reasonable results. Assuming there is no problem with model verification and validation phases, time-consuming problem arises from the necessity of a large amount of observations to yield an accurate result.

[*]Partially supported by ISAE/Fundação Getulio Vargas

Multiple Replications in Parallel (MRIP) aims to attack this issue by generating data in parallel, asynchronously, through a local area network of workstations. Despite the increasing computer power, there will be always a necessity of statistical analysis of data if accuracy of results is an issue. There are two ways of analyzing data : (i) fixing *a priori* the amount of computer time or the maximum number of observations to be collected; or (ii) analyzing data sequentially as they are produced during simulation, and stopping the simulation when a predefined accuracy is achieved.

The former, also known as fixed-sample-size scenario, the final statistical error can not be controlled in advance, since "*... no procedure in which the run length is fixed before the simulation begins can be relied upon to produce a confidence interval that covers the true theoretical value with the desired probability*" (Law and Kelton, 1991). The latter approach, also known as sequential scenario, offers an attractive way to control the precision of results. The importance of sequential procedure is widely recognized and they become more and more often used by practitioners as methods for controlling the precision of simulation results.

In this paper we consider a practical implementation of a sequential confidence interval procedures (CIP) based on *Standardized Time*

Series (STS) proposed by (Schruben 1983), and compare it against a sequential version of *overlapping batch means* (OBM), proposed by (Meketon and Schmeiser 1984). The main criteria of comparison are the final coverage of results, the average run length given by the number of observations required to stop the simulation, and the variability of the final confidence intervals. We also consider *en passant* other issues related to MRIP.

2 BATCHING TECHNIQUES

The problem of constructing a confidence interval for the steady-state mean of performance parameter, has offered a vast field of research and the main practical problem is due the presence of strong correlations among observations of simulated processes, also in steady state. Positive correlations, very common in queuing systems, denote negative bias and, thus, the final confidence interval half-width can be underestimated, which leads frequently to the final coverage lower than the nominal confidence level. By coverage we mean the frequency that the final confidence intervals contain the true value being estimated.

Theoretically, steady state occurs in the limit when the run length increases to infinity, but in practice there is a point from which one can assume observations have almost the same distribution. Even though, observations are usually strong correlated and classical statistics techniques can not be applied directly.

To get rid of this tactical problem, several methodologies for estimation of confidence interval on the mean of sequence of correlated observations have been proposed over the last decades. The main difference among these methodologies is the way they estimates variance of the mean. For a thorough review of this and other related problems refer to (Pawlikowski 1990).

Batching methods are supposed to be intuitive and ease to implement, although many issues are still open; especially if one considers their applications under distributed stochastic simulation. The fundamental assumption of batch means methods is that there exists a number $B*$ of batches, or equivalently a batch size M^*, such that the batch means are i.i.d. normal random variables (Schmeiser 1982).

2.1 Sequential CIP based on OBM

The main issue of every batch-means-based procedure is the determination of the minimum batch size M^* that leads to (almost) uncorrelated batch means. After collecting N observations, one should divide them into B batches of size M (N=B.M), and test the batch means against correlation, by applying a test of stationarity proposed for example in (Schruben 1982). If the test fails, the batch size is increased, more observations are collected and the correlation test is repeated. If the test succeeds, observations are grouped in a somewhat different way, where each observation initiates an (overlapped) batch of size M^*. The procedure continues collecting more observations, each observation giving rise to an overlapped batch, and batch means are used to estimate the parameter of interest.

Overlapping Batch Means – OBM, has been considered to be the most promising techniques among batching-based methods of simulation analysis under *Multiple Replications in Parallel* (Mota 1999). Nevertheless, it still suffers of an additional burden relative to the length of the phase that the most convenient size of batches are being searched for, namely the batch size determination phase – BSD.

2.2 Sequential CIP based on STS

This approach takes another order of ideas to generate confidence intervals for steady-state simulation, that can still be based on batching. Instead of standardizing a single scalar, e.g. the sample mean of an output time series Y_i (i=1,2, ..., N), Schruben (Schruben1983) suggests the standardization of the entire time series, defined as

$$T_i(t) \equiv \frac{\lfloor Mt \rfloor (\bar{Y}_{i,M} - \bar{Y}_{i,\lfloor Mt \rfloor})}{\sigma\sqrt{M}}, \; 0 \leq t \leq 1 \quad (1)$$

where $\bar{Y}_j \equiv \sum_{i=1}^{j} Y_k/j$, j=1, ..., N, and $\lfloor . \rfloor$ denotes the greatest integer function.

The transformed series converges asymptotically to a standard Brownian bridge process, whose properties are used to construct a confidence interval.

After standardizing each observation one can find random variables A_i, the asymptotic scaled sum of $T_i(t)$, for each batch by means of

$$\begin{aligned} A_i &= \sigma\sqrt{M} \sum_{k=1}^{M} T_i(t) \quad (2) \\ &= \sum_{k=1}^{M} \sum_{j=1}^{k} (\bar{Y}_i - \bar{\bar{Y}}) \end{aligned}$$

A simplification can be found toward facilitating a sequential procedure. From (2) and (1) one can find that

$$\begin{aligned} A_i &= \sigma\sqrt{M} \sum_{k=1}^{M} \Big[\frac{k(\bar{Y}_{i,k} - \bar{Y}_{i,M})}{\sigma\sqrt{M}} \Big] \\ &= \sum_{k=1}^{M} k \Big[\frac{1}{k} \sum_{j=1}^{k} Y_{(i-1)M+j} - \frac{1}{M} \sum_{j=1}^{M} Y_{(i-1)M+j} \Big] \\ &= \sum_{k=1}^{M} \Big[\sum_{j=1}^{k} Y_{(i-1)M+j} - \frac{k}{M} \sum_{j=1}^{M} Y_{(i-1)M+j} \Big] \end{aligned}$$

$$A_i = \sum_{k=1}^{M} \Big[j - \frac{M+1}{2} \Big] Y_{(i-1)M+j} \quad (3)$$

By computing the statistic

$$A = \sum_{i=1}^{B} \frac{12 {A_i}^2}{(M^3 - M)} + M(\bar{Y}_i - \bar{\bar{Y}})^2$$

an asymptotically valid combined **c**lassical-sum interval estimator $[\bar{\bar{Y}} \pm \mathbf{H}]$ can be constructed for performance parameter μ, considering that $V_{csum} = \frac{A}{2B-1}$, and $\mathbf{H} = t_{2B-1,1-\alpha/2}\sqrt{\frac{A}{2B-1}}$.

3 AN MRIP IMPLEMENTATION

We used Akaroa-2, an implementation of a simple yet effective approach for speeding up sequential simulation known as *Multiple Replications in Parallel* (Pawlikowski 1994).

Akaroa-2 is a user-friendly simulation package, written in C++, designed for automated parallelization of ordinary simulation models and fully automated control of accuracy of the final results. It permits a simulation model be executed on different processors in parallel, trying to produce IID observations by initiating each replication with a nonoverlapping stream of pseudorandom numbers. All series of replicated simulations are executed using strictly nonoverlapping sequences of pseudo-random numbers provided by exhaustively tested generators.

Essentially, a master process (*Akamaster)* is started on a processor that acts as a manager, while one or more slave processes (*akslave*) are started on processors that take part in the simulation experiment, forming a pool of simulation engines. Akaroa-2 takes care of the fundamental tasks of launching the same simulation model on the processors belonging to that pool, controlling the whole experiment and offering an automated control of the accuracy of the simulation output.

At the beginning, stationary tests due to (Schruben 1982) are applied locally within each replication, to determine the onset of steady state conditions in each time-stream separately, and the sequential version of the CIPs are used to estimate the variance of local estimators at consecutive checkpoints.

Each simulation engine keeps on generating output observations. These observations are collected by a local analyzer, who is responsible for determining the locations of checkpoints, that is, the instants at which a group of observations is sufficient to yield a reasonable estimate to the global analyzer, residing in the processor running akmaster. The global analyzer calculates a global estimate, based on local estimates delivered by individual engines, and a simulation manager verifies if the required precision is reached. When it happens then the overall simulation is finished. Otherwise, more local observations are required, so simulation engines continue their activities.

By achieving a checkpoint, the current local estimate and its variance are sent to the global analyzer which computes the current value of the global estimate and its precision. A checkpoint is associated with determining an estimate of type

$$(N_i, X_i, V_i),$$

where X_i and V_i are the sample mean and variance, respectively, obtained by applying one of the methods of analysis mentioned above. Particularly, we are interested in methods based on batching approach, as they are practical, though many issues remain open, especially regarding its implementation in a parallel simulation environment such as that created by Akaroa-2.

There is a global analyzer in Akaroa-2 for each performance parameter being estimated, which averages estimates coming from processors. The grand mean and its variance can be found by

$$\bar{\bar{X}} = \frac{\sum N_i.X_i}{\sum N_i} \tag{4}$$

$$Var[\bar{\bar{X}}] = \frac{\sum N_i^2.Vi}{\sum N_i^2} \tag{5}$$

More details can be found in (Mota 2000).

4 PERFORMANCE COMPARISON

In order to compare the performance of sequential CIPs based on OBM and STS, we investigated their performance to Akaroa-2, through the simulation of the mean waiting time of four queuing systems with increasing coefficient of variation of the service times, namely, $M/D/1, M/E_4/1, M/M/1$ and $M/H_2/1$. We are going to compare here the CIP's performance considering the most correlated process, namely $M/H_2/1$, as it imposes more difficulty in the analysis.

The number of processors used was P=2,6 and 10, traffic intensity 0.90 and stopping rule adopted was a 5%-relative precision. It is worthwhile to say that adding more processors reduces always the simulation completion time, but it remains a question whether this would lower the quality of the results.

The main criterion of comparison is the coverage of results, and we have applied the sequential coverage analysis proposed by (Pawlikowski et al. 1998). That is, coverage analysis is initiated only after collecting a minimum number of bad confidence intervals (e.g. 200). Too short simulation runs are, then, discarded, and the sequential coverage analysis stops when the relative precision of the half-width of the confidence interval for the coverage is less than or equal 5%. Instead of investigating coverage analysis for a single confidence level, we followed the suggestion of (Schruben 1980) and investigated for a range of values varying from 0.1 to 0.95.

Fig. 1 shows the coverage function for both CIPs. When the underlying assumptions are satisfied, a perfect procedure would yield final coverage following the 45-grade-straight-line. Values below that line indicate an underestimated confidence intervals, while values above that line indicate that the procedure collected more observations than it is needed, that is, a waste of resources.

We can observe that OBM behaves almost ideally, whereas STS seems to be not so effective for lower values of confidence level. They are both robust procedures, however, as they converge to the ideal case as long as the run length increases.

Fig. 2 shows the average total number of observations collected by the end of the simulation. The magnitude of this value was almost equivalent for both procedures, but for lower values of confidence level, STS required less observations to achieve the desired precision, what could in some sense explain the degradation in coverage for that region. Nevertheless, STS presents a very interesting behavior for a CIP under MRIP. One could expect that adding more processors should result in a corresponding decrease in the total number of observations necessary to stop the simulation. This is clearly visible in the case of STS.

Fig.3 summarizes the variability of the final confidence intervals by looking at the coefficient of variation of the half-width H, given by

$$CoV[H] = \frac{(Var[H])^{1/2}}{E[H]}$$

It is obvious that OBM is more stable than STS, although difference diminishes as the confidence level increases.

5 CONCLUSIONS

In general, in MRIP application, both sequential procedures behave attractively for highly correlated queuing processes, i.e. processes that create the main analytical problem for simulation output data analysis.

OBM is robust and stable, and offers a very attractive alternative when applied under MRIP. Other variants of *batch means* wait to form a complete batch before sending an estimate to the global analyzer, that is, before yielding a checkpoint. With OBM, this granularity can be as short as 1, and one can achieve higher degrees of speedup. A possible drawback seems to be the overload put on the global analyzer. This issue is being currently investigated.

Although STS produces less stable confidence intervals than OBM for lower values of confidence levels, it offers an attractive feature for being used under MRIP. Namely its BSD phase (the detection of normality of $A_i's$) is shorter than that of the OBM (the test of independence of the batch means). It means that parallelization of STS can be more efficient than of the OBM. The question of granularity could be resolved if one applies the same idea of overlapped batches. It is being investigated and will appear soon in a future work.

References

[1] V. Yau K. Pawlikowski and D. McNicke. Distributed stochastic discrete-event simulation in parallel time streams. In *Winter Simulation Conference*, pages 723–730, New York, December 1994. ACM Association for Computing Machinery.

[2] M.S. Meketon and B. Schmeiser. Overlapping batch means: something for nothing ? In *Winter Simulation Conference*, pages 227–230. IEEE Computer Society Press, December 1984.

[3] E. Mota, A. Wolisz, and K. Pawlikowski. Sequential batch means techniques for mean value analysis in distributed simulation. *Proceedings 13th European Simulation Multiconference, Warsaw*, 1999.

[4] Edjair Mota. *Performance of Sequential Batching-based Methods of Output Data Analysisis in Distributed Steady-state Stochastic Simulation.* PhD thesis, Department of Computer Science, Technical University of Berlin, under submission.

[5] K. Pawlikowski. Steady-state simulation of queueing processes: a survey of problems and solutions. *ACM computing surveys*, 22, 2:83–121, 1990.

[6] K. Pawlikowski, D. McNickle, and G. Ewing. Coverage of confidence intervals in steady-state simulation. *Journal Simulation Practice and Theory*, 1998.

[7] B. Schmeiser. Batch size effects in the analysis of simulation output. *Oper. res.*, 30:556–568, 1982.

[8] L. Schruben. Confidence interval estimation using standardized time series. *Operations Research*, 31, 6:1090–1108, 1983.

[9] L. W. Schruben. A coverage function for interval estimators of simulation response. *Managem. sci.*, 26, 1:18–27, 1980.

[10] L. W. Schruben. Detecting initialization bias in simulation output. *Operations Research*, 30, 3:569–590, 1982.

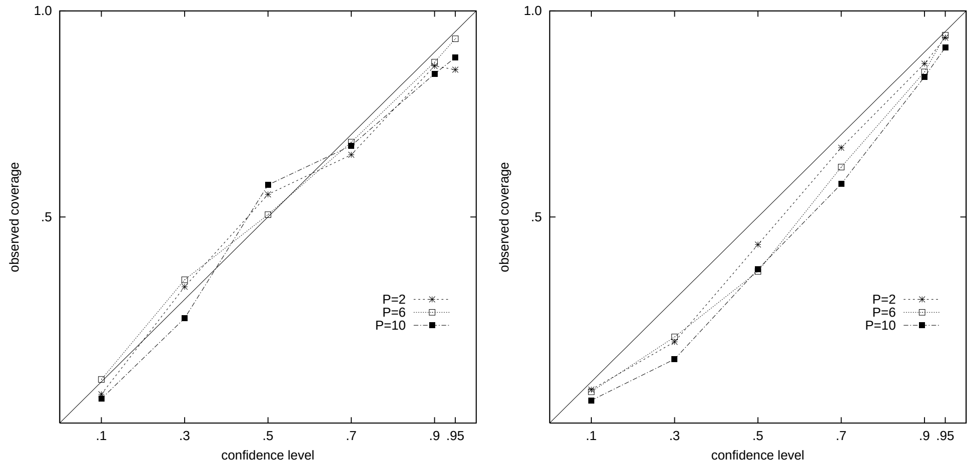

Fig. 1: Coverage function: OBM (L) and STS (R)

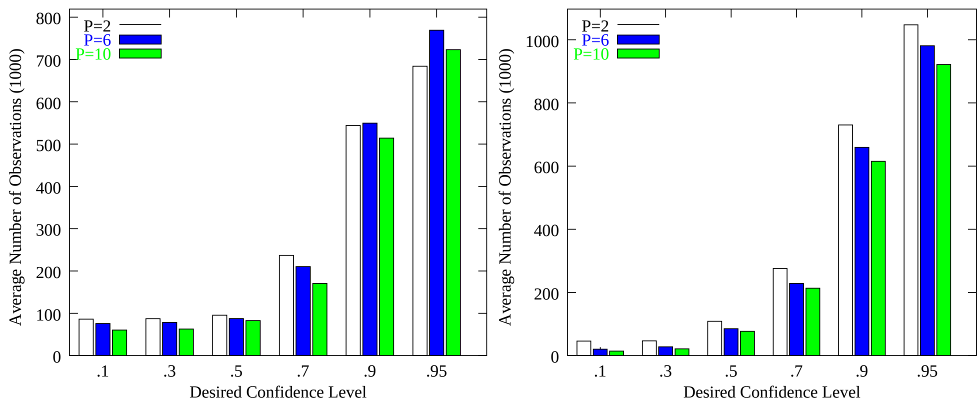

Fig. 2: Average total number of observations: OBM (L) and STS (R)

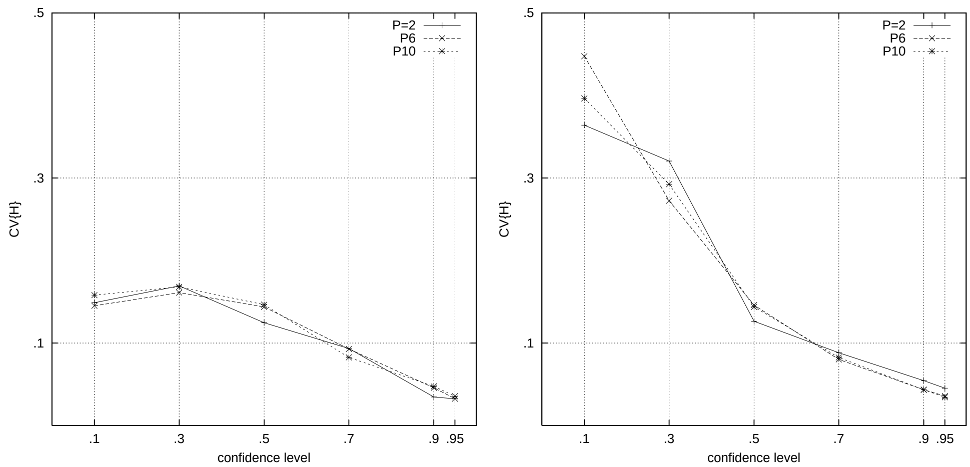

Fig. 3: Coefficient of variation of H: OBM (L) and STS (R)