Journal of Statistics: Advances in Theory and Applications
Volume 1, Number 2, 2009, Pages 169-192

SAMPLING METHODS AND PARALLELISM INTO MONTE CARLO SIMULATION

MEGDOUDA OURBIH-TARI and ABDELOUHAB ALOUI

Department of Mathematics
Laboratory of Applied Mathematics
University of Bejaia
Algeria
e-mail: megourbih@gmail.com
aaloui_abdel@yahoo.fr

Abstract

This paper provides an updated review of Monte Carlo simulation. The use of random sampling, its consequences in simulation, the use of descriptive sampling, its advantages, its limits and its improvement upon the precision of simulation estimates: Refined descriptive sampling. This state of the art gives also an insight into other sampling procedures used in the literature, like Ranked set sampling, Systematic sampling, Stratified sampling, Latin hypercube sampling, *L* Ranked set sampling, Importance sampling and Quasi Monte Carlo methods. Finally, as a future work parallelism of the best sampling procedure is proposed to reduce the time of running simulation experiments.

1. Introduction

Monte Carlo method traces its modern origins and name to the work of Von Neuman and Ulam in the late 1940's. Later, simulation inherited the name of Monte Carlo (MC) method and the term of MC sampling has become synonymous with Random Sampling (RS). This statistical method

2000 Mathematics Subject Classification: 62D05, 62N05.

Keywords and phrases: variance reduction, sampling, Monte Carlo parallelism.

Received May 8, 2009

© 2009 Scientific Advances Publishers

is the purest form of probability sampling. Each member of the population has an equal and known chance of being selected. The random sampling method is commonly used to simulate systems containing stochastic or probabilistic situations. It might also be considered to solve some deterministic problems that cannot be solved analytically. Currently, it is widely used in Monte Carlo simulation (Fishman [19] and Anderson [4]) and has a very large application field; like differential equation integration, matrix inversion, particles transport, fluid mechanics and financial mathematics (Ross [49], Doucet [15] and Robert [47]).

This paper deals with the use of random sampling and its consequences in simulation; an update of various work on Descriptive Sampling (DS), its advantages and its problems as well as the Refined Descriptive Sampling (RDS) recently introduced as a better alternative to RS and DS. We will also expose in this paper, other sampling methods used in the literature like Ranked set sampling, Systematic sampling, Stratified sampling, Latin Hypercube Sampling (LHS), *L* Ranked set sampling, selective sampling, Importance sampling and Quasi Monte Carlo (QMC) methods. Finally, this paper gives an open research issues on the chosen best sampling method used in Monte Carlo Simulation: parallelism.

2. Problem Formulation

In a simulation study, a logical model is built and used as a vehicle for experimentation. The model is illustrated in Figure 1. The distributions of input variables are assumed to be known, while the response variables' distributions are unknown. When the problem is simulated, input random variables are replaced by samples. As a result, response variables are also replaced by samples. Therefore, experiments are carried out on the model and unknown parameters of the response random variables of interest are estimated.

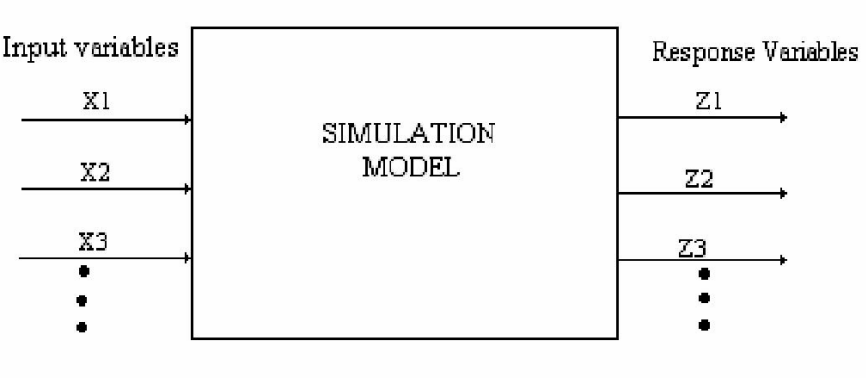

Figure 1. Simulation model representation: A set of input random variables is transformed into a set of output random variables.

When RS procedure is used, the sample values are generated using some Random Number Generators (RNGs) and either the inverse cumulative function distribution or equivalent methods. RNG provides sequences of real numbers in the interval $[0, 1)$. These numbers are viewed as realizations of identically independent distribution $U(0, 1)$ random variables and transformed as needed to generate variates from some probability distributions, and then unknown parameters of the output random variables are estimated. Thus, every Monte Carlo simulation that requires s such uniforms may be regarded as computing a function f defined over the s dimensional unit hypercube $[0, 1)^s$, at some point u determined by the RNG, to estimate a mathematical expectation that can be written as

$$\mu = E(f(U)) = \int_{[0,1)^s} f(u)du.$$

3. Variance Reduction Techniques

Indeed, random sampling can solve a large variety of problems but some variations are present in a randomly generated sample. Already in the late 1930s, several authors studied RS variability, for example, Kendall and Babington [25] argued that two sources of sampling errors: the set effect and the sequence effect are generated using RS method in

most simulation studies. Another source of error was detected by Saliby [51] and was given the name of the set-sequence interaction effect. This type of error cannot be explained by neither the set effect nor the sequence effect, but only by their interaction. In the same reference, empirical results on different simulation problems were given to support the existence of this source of error and the set effect was explained by the observed deviations results between the input sample moments and their corresponding theoretical values.

Several authors, for example, Bratley et al. [8], Ross et al. [48], Tuffin [67] and Law [28] suggested two alternative methods to reduce the sampling errors. One is the use of Variance Reduction Techniques (VRTs), and the other is the use of replicated runs. The most used variance reduction techniques in the literature are control variates, antithetic variates and common random numbers.

Control variates

The method of control variates is defined by the variables used to control the sampling errors called control variables. For instance, if the mean is used to control the sampling errors then, the mean is a control variable. In this method, the idea is to exploit the correlation that might exist between the input variable and the respective output variable in order to obtain estimators values that are more accurate.

In the study of the variability of simulation estimates, Ehrenfeld and Ben Tuvia [18], and Saliby [51] show that the linear response model explains a great deal of sampling variability. The idea of the linear response model came from the need to investigating the relationship between the simulation outputs and the sampled inputs. This technique is equivalent to the linear response model (Saliby [51]), in a way that the use of control variables will provide a more useful interpretation of the regression coefficient. Indeed, if the input parameters are identified as control variables, then the linear response model becomes a regression model.

Antithetic variates

In order to reduce the variance of simulation estimates, Hammersley and Morton [23] introduced the method of antithetic variates This VRT deals with two negatively correlated estimates of an unknown parameter. In the case of both independent estimates, the variance of the overall mean will be smaller than the respective variances. Tocher [66] gave a way of obtaining negatively correlated results by using the complementary random streams, and then, the method of antithetic variates was associated with complementary random streams. Barnett [7] has shown that antithetic variates and control variates are both efficient on an example of the negative exponential law with unit mean and he obtained negatively correlated results with a correlation coefficient of $1 - \frac{\pi^2}{6} = -0.645$ using complementary random streams of the same exponential law. Shannon [56] agreed that antithetic variates is efficient compared with other techniques of variance reduction by considering the advantages of the method and the fact, that is, relatively easy to find negatively correlated unbiased estimators. Cheng [11] shows the ease of implementation, and the effectiveness of the method through numerical examples and proposes a procedure based on both antithetic variates and control variates for generating samples from a normal distribution. Davidson and Mac Kinnon [13] studied these two methods arguing that both work well in practice and greatly reduce the number of replications required to obtain a given level of accuracy.

Common random numbers

Common random variates also known as correlated sampling procedure is a technique of reducing variance applicable for comparing two or more alternatives that could be policies or systems configuration. The principle of the method is to use the difference between the responses of alternatives to estimate the difference between the respective expected values. The objective is to obtain a high and positive correlation between all responses.

Several authors including Shannon [56], Kleijnen [26] and L'Ecuyer and Buist [31] agreed that correlated sampling is more stable compared to other methods of variance reduction and have shown its power on some examples. Saliby [51], who confirmed empirically the existence of a high correlation between the responses when this method is used, gave a justification of common random variates. Moreover, he proved that the positive and high correlation observed between the responses is attributed to the various sources of sampling errors. He also argued that this method controls only partially the set effect, but it is the only technique, where part of the sequence effect can also be controlled.

Replicated runs

Suppose that a MC experiment involves n independent replicated runs. In a given run i, the use of random sequence $u_i \in [0, 1)^s$, $i = 1, \ldots, n$ leads to the following MC unbiased estimator of μ

$$Q_n = \frac{1}{n}\sum_{i=1}^{n} f(u_i),$$

which has a variance of σ^2 / n, where

$$\sigma^2 = \mathrm{Var}(f(U)) = \int_{[0,1)^s} f^2(u)du - \mu^2.$$

The use of replicated runs was criticized by some authors for instance, Shannon [56] and L'Ecuyer [29] argued that the error of the average results is inversely proportional to the square root of runs number, that is, the error $Q_n - \mu$ converges at rate $O(\sigma / \sqrt{n})$.

4. Probability Sampling Methods

Sampling methods are classified as either probability or non-probability. In probability samples, each member of the population has a known non-zero probability of being selected. Probability methods include RS, ranked set sampling, systematic sampling, stratified sampling, Latin hypercube sampling and L ranked set sampling. All these sampling methods can also be considered as VRTs.

Ranked set sampling

Mc Intyre [34] proposes an unbiased selective sampling method using ranked sets called later Ranked Set Sampling (RSS) by Halls and Dell [22]. In this method, we take the largest in the first of n sets, each of n random items, the second largest in the second set, and so on to the smallest in the n-th set. The sample of n items selected in this way is an unbiased sample of the population. This sampling procedure induces stratification of the whole population at the sample level; indeed, it is a random sampling generation from the subpopulations of predominantly criteria without having to construct the subpopulation strata. Each subpopulation has its own distribution. Many types of this sampling procedure are used for estimating the population mean. Dell end Clutter [14] show that ranked set sampling produces unbiased simulation estimates and at least as precise as the RS but Muttlak [37] increases the efficiency of the estimators obtained by RSS just on a simple linear regression model, it follows that RSS still fails to improve on random sampling.

Systematic sampling

Systematic sampling is a probability method often used instead of random sampling. It is also called an N^{th} name selection technique. After the required sample size has been calculated, every N^{th} record is selected from a list of population members. A random starting point must also be selected. As long as the list does not contain any hidden order, this sampling method is as good as the random sampling method. Its only advantage over the RS technique is simplicity. Systematic sampling is frequently used to select a specified number of records from a computer file. When this sampling procedure is used the list can have a periodic arrangement and in this case, it can fare very badly. Unless the list itself is in random order, which it never is, systematic sampling is not better than random sampling.

Stratified sampling

Stratified sampling is commonly used probability method, that is, better than random sampling since it reduces sampling error. A stratum is a subset of the population that shares at least one common

characteristic. The researcher first identifies the relevant stratum and their actual representation in the population. A sample is then taken from each stratum using either RS or Systematic sampling, and when it is a random sample, it is referred to as stratified random sampling. The sample size must be large enough to be reasonably confident that the stratum represents the population. Stratified sampling is often used when one or more of the stratums in the population have a low incidence relative to the other stratum. Rubinstein [50] and Nelson and Schmeiser [38] discussed stratified sampling in the context of MC simulation.

Latin hypercube sampling

McKay et al. [35] proposed Latin hypercube sampling. The latter was suggested as a variance reduction technique in which the selection of sample values is highly controlled, although still letting them to vary. The basis of LHS is a full stratification of the sampled distribution with a random selection inside each stratum. It is a kind of stratified random sampling, where sample values are randomly shuffled among different variables. An input sample is generated by

$$(xh)_{ji} = H^{-1}\left[(i - 1 + R_i) \Big/ n_j\right]$$

$$\text{for } i = 1, 2, \ldots, n_j \text{ and } j = 1, 2, \ldots, k,$$

where R_i stands for an independent random uniform on $[0, 1]$, $i = 1, 2, \ldots, n_j$ and $H^{-1}(R)$, $R \in [0, 1)$ is the inverse cumulative distribution of a particular input variable.

This technique has been extensively used in practice, not only because of its implementation simplicity but also because of its nice properties. Indeed, Mc Kay et al. [35] show that the variance of the LHS estimates is not higher than the sample variance from Monte Carlo sampling, if the function f is monotone in all of its arguments. Hoshino and Takemura [24] extend this result to the case, where the function f is monotone in all but one of its arguments. Stein [59] writes the ANOVA description of the function f and show that asymptotically, the sample variance from LHS is just equal to the variance of the residual term and is lower than the variance of Monte Carlo sampling. Loh [33] extends this

results to the multivariate case, where $f : R^s \to R^d$. Owen [43] shows that LHS satisfies a central limit theorem with the variance equal to the variance of the residual term and in [44] the same author shows that for any n and any function f,

$$\mathrm{Var}_{LHS} \leq \frac{n}{n-1} \mathrm{Var}_{MC}.$$

L Ranked set sampling

A robust ranked set sampling (LRSS) is introduced by Al-Nasser [3] as a generalization for many types of ranked set sampling. The same author argued that LRSS method gives an unbiased estimator for the population mean with minimum variance providing that the underlying distribution is symmetric. However, comparisons between this sampling method, RSS and RS for detecting outliers on some distributions are given in the same reference and the results indicate that the estimator produced by using LRSS is the best. This sampling method still lacks adequate properties, its efficiency is proved in particular case (symmetric distribution) and so, it cannot replace random sampling.

5. Non-probability Sampling Methods

According to Saliby [51], it is not necessary to have a large amount of sampling errors in simulation, if later; we use some techniques to remove part of them. He argued that trying to reduce the variance without knowing a priory its cause is obviously an unsatisfactory approach. As sampling errors are generated by the use of random sampling procedure, he concluded that RS is inefficient and unwise. We also know that the simulation estimates are affected by such sampling errors. Therefore, the former vary between different runs, whereas the model remains unchanged. Consequently, their variability is subject to RS procedure and the random behaviour of an input stochastic variable is not well represented. Given the relationship between the lack of precision of simulation estimates and random sampling method, it was born a remedial new paradigm. The latter says that is not always necessary to generate sample values randomly to describe a stochastic behaviour of an input distribution. As the aim of any sampling procedure is to represent

truly the population distribution from which it is generated so, that any simulation estimates drawn from it can be safely implemented in the real system. Then, new non-random sampling methods were derived from this paradigm. In non-probability sampling, members are selected from the population in some non-random manner. Most of them are presented in this section. These include selective sampling, importance sampling, descriptive sampling, refined descriptive sampling and quasi Monte Carlo methods. All these non-random sampling methods can also be considered as variance reduction techniques.

Selective sampling

Brenner [3] proposed a more restrictive procedure than random sampling called selective sampling. This procedure consists of sampling without replacement. Kleijnen [26] argued that the use of this method will produce biased results, so, no real attention has been given to the method that has never been taken up.

Importance sampling

Importance sampling attempts to concentrate sampling in regions of interest, where "interest" may be related to the variance within the region, the likelihood of observations in the region, and/or the magnitude of observations in the region. This sampling method biases the outputs by altering the probability distributions of the inputs. Traditional fast simulation techniques are usually based on importance sampling. Since the simulation estimates are biased through this sampling procedure, it cannot be compared to random sampling.

Quasi Monte Carlo methods

In quasi Monte Carlo methods, we use deterministic sequences of weak convergence to reduce the variance in MC method. It is also a variance reduction technique. In other words, we replace the random points u_i by a set of points $p_n = \{u_1, u_2, \ldots, u_n\} \subset [0, 1)^s$ that cover the unit hypercube $[0, 1)^s$ more uniformly than typical random points. The two main classes of methods for constructing such point sets are digital nets and integration lattices (L'Ecuyer et al. [30], Niederreiter [39] and Sloan [58]).

Tuffin [68], Owen [44] and Okten [40] using, respectively, a random permutation of QMC methods, randomized nets and sequences and random sampling with low discrepancy sequences, carried out various comparisons between the simulation of MC and QMC producing hybrid methods. Keng Seng and Boyle [60], analyzes the Owen approach of randomized quasi random sequences, suggesting an improvement for high dimensional cases, exhibiting better performance than both MC and QMC methods. Okten [42] shows that an hybrid-Monte Carlo sequence satisfies a central limit theorem and improved its error bounds. The existing literature in this field abounds of theoretical results on QMC, for example, (Morokoff [36], Avramidis and L'Ecuyer [5]).

Descriptive sampling

Later, in an attempt to improve on random sampling procedure, Saliby [52] proposed an alternative approach, called, descriptive sampling. This method is based on a deterministic selection of the input sample values and their random permutation. The values of the descriptive sample do not vary but only their sequences vary between different simulation runs. Unlike the RS, the set of input values are the same for all replicated runs in the simulation. Once the sample size is known, the set values are defined for each input random variable Xj, $j = 1, \ldots, k$ using the inverse method by

$$(xd)_{ji} = H^{-1}\left[(i - 0.5) \Big/ n_j\right]$$

$$\text{for } i = 1, 2, \ldots, n_j \text{ and } j = 1, 2, \ldots, k,$$

where $H^{-1}(R)$, $R \in [0, 1)$ is the inverse cumulative distribution of a particular input variable.

This method has been criticized by Pidd [45], who argued that it can be biased, and its strict operation requires a prior knowledge of the sample size. Some other problems on its data processing implementation were also mentioned by some authors and solved later by Saliby [53], who implemented DS on different models in the three phase discrete event simulation.

Saliby [52] argued that if the sample size is known in advance, the use of DS eliminates completely the set effect. Therefore, the simulation estimates produced will be more precise than those obtained by RS method. In this case, DS offers the advantage of the precision of the simulation estimates and once this procedure is used, there is no need for further variance reduction techniques. Indeed, the assumption on the sample size makes the use of DS difficult, since its determination beforehand is not easy for real simulation problems but in this case, the same author proposed to use again the same sample values to overtake this possible problem.

Although, no mathematical proof on the study of biased estimates was proposed by Saliby [52], all the same, he mentioned its insignificance. He added that, even if this bias exists, it will not overtake the coefficient of correlation between the descriptive variables and then, it will be of a small magnitude.

Refined descriptive sampling

In Tari [63], a study is available on the bias that DS produces based on some stylized input-output transformations that allow the determination of the types of problems, where DS encounters the most bias. Several authors argued that the possibility of producing biased estimators through simulation using DS comes from the fact that the descriptive variables are dependent and negatively correlated variables, which constitutes the main difference between descriptive and random sampling procedures, but this dependence does not increase the risk of biased estimates. Indeed, the problem of bias in DS is because a regular sampling grid is used to select the input values. If the input function is periodic, there is a risk that the set of output values will be biased. However, if the input function is not periodic, there is no risk of sampling bias.

Based on this study, refined descriptive sampling is proposed by Tari and Dahmani [64] to improve on descriptive sampling by reducing significantly its bias. This sampling procedure is based on a block that must be situated inside a generator aiming to distribute regular numbers from subsets of prime number sizes as required by the simulation. We stop the process when the simulation terminates. The generation process

of the subset values is deterministic, whereas the generation of the prime numbers as well as the sequence of the subset values is random. In this method, each run is determined by a block of different prime numbers.

Suppose that m prime numbers have been used in a simulation run. Sample values for any input random variable X are generated as required by the simulation using the inverse method by

$$(xd)_{qi} = H^{-1}\left(\frac{i - 0.5}{p_q}\right) \text{ for } i = 1, 2, \ldots, p_q \text{ and } q = 1, 2, \ldots, m,$$

where p_q, $q = 1, 2, \ldots, m$ are prime numbers generated randomly and $H^{-1}(R)$, $R \in [0, 1)$ is the inverse cumulative distribution of the input variable X.

This approach removes the need to determine in advance the sample size. Mathematical arguments together with a proof of its efficiency are presented in Tari and dahmani [64] by studying a problem, whose input variable is a sinusoidal function.

6. Comparison

Nelson and Schemeiser [38] proposed and illustrated variance reduction taxonomy by considering seven VRTs namely, antithetic variates, common random numbers, control variates, conditional expectations, importance sampling, stratified sampling and post stratified sampling. This taxonomy reduces the confusion existing among variance reduction techniques and provides a common language for communication among researchers and practitioners.

Saliby [51] compared antithetic variates and common random numbers showing that the former is less efficient than the latter. In the same reference, he argued that the assumption of antithetic results does not hold, and that is an inefficient procedure for controlling the set effect.

Kwon and Tew [27] for combining antithetic variates and control variates, proposed three combined methods. The efficiency of each combined method was proved on a simulation model of a resource constrained activity network. In the same reference, the author stated

that the choice of combined methods depends on the degree of correlation between the control variates and the response.

L'Ecuyer and Buist [31] simulate a telephone call center model, where agents answer incoming calls and show that the variance is reduced by using a combined method of control variates and stratified sampling. They argued that proper use of common random numbers reduces the convergence rates of the variance of the performance measures across all configurations of the system.

DS, RDS and LHS are all based on a random permutation of the input numbers but both DS and RDS select their values differently from LHS.

In Saliby [54], there is a discussion suggesting that DS has a lower variance than LHS without referring to the produced bias. However, it is well known that both LHS and RS are unbiased (Mc Kay et al. [35] and Drew and Homen de Mello [16]).

Several empirical comparisons on a PERT network, an M/M/1 queue and on an inventory system (Saliby [52]) show that the estimates of the output random variables parameters produced through simulation using DS are with lower variance than those obtained by RS method. That studied also the Newsboy problem and the obtained results are exactly the theoretical values. This is a particular case, where the estimators are independent of the sequence of the input values.

Saliby and Pacheco [55] compare the efficiency of six Monte Carlo simulation sampling methods namely QMC using Halton, Sobol and Faure numeric sequences, DS, LHS and RS in two finance applications: a project risk analysis and a correlated stock portfolio. In this study, DS and LHS have shown the best results. Therefore, both DS and LHS outperformed QMC methods.

Tari and Dahmani [64] have shown that RDS is an efficient sampling in simulation studies developing methods. This reference was concerned with ensuring that the application of the RDS method is safe, correctly designed and fitted to any simulation in an economical and undemanding manner.

Tari and Dahmani [61, 62] compare the efficiency of DS and RDS on a production system of the flow shop type by discrete event simulation method and a manufacturing production system showing that RDS produces better estimates of the output random variables parameters produced through simulation than DS and RS and are with lower variance. So, it has been shown that RDS is an improvement over DS, RS, LHS and QMC methods, and it is capable of substantially reducing the cost of running simulation experiments, if properly applied. As a conclusion, we deduce that RDS outperform all other sampling methods.

7. Simulation Studies

There are many diagrams and descriptions that outline the key processes in a simulation study as there are authors, who have written about the subject. Among them are Bank et al. [6] and Robinson [46]. The outline of a simulation study that must be performed are firstly, a description of the model, that is, to be developed, secondly, the simulation model implemented on a computer, then the results of the experiment and finally the improvement in the real world obtained from implementing the solutions. In fact, simulation modeling involves both repetition and iteration. A large simulation model may take a number of hours to run and of course, many runs may be required for thorough experimentation. Because simulation is time-consuming approach, it is recommended that it be used as a means of last resort, rather than the preferred option (Pidd [45]). That said, simulation is often the only resort. Indeed, simulation studies systems whose analytical solution is difficult even impossible. Surveys of modelling practice demonstrate that simulation is one of the most commonly used modeling techniques.

The development of simulation software has been very closely allied to the development of computing. A wide range of software is now available for developing simulation models (Law and Kelton [28]). Most simulation software is not cheap and most simulation projects take few weeks to complete. Few software applications require the computing power necessary for simulation. Even Taylor et al. [65] explore the use of net conferencing during simulation studies. Beyond the simulation software, there is a range of other packages that may sometimes be useful

during a simulation study. Spreadsheets, databases, and statistical packages may be useful for working through more complex analyses.

8. Parallelism

It is obvious, from Section 7 that a simulation study is expensive in time and memory and more generally in necessary resources (Ahn and Danzig [1]). Indeed, the traditional simulation tools could be unusable when complex systems are studied. We are then confronted with the problem of the best compromise between the precision of simulation estimates and the cost of the experiment. The latter of course must be balanced against the benefits that can be gained from the use of simulation, by the same way, the precision of the estimates that are often an order of magnitude greater than the cost. Unfortunately, the limited means of available sequential computation make this achieving goal difficult. Then, parallelism proposes itself as an inevitable solution to the above problem. It can be regarded as a means of reducing the cost in time and memory of a program solving a complex problem.

It is well known that the parallel processing of a program by more than one task is independent with each task being able to execute the same or different statement at the same moment. In parallelism, some mechanisms are more efficient on quite particular architectures. Indeed, the choice of adequate mechanisms with problem specificities is necessary. A programmed application using threads is more appropriate on shared memory than on distributed memory architectures. On the other hand, if we have a large amounts of data, it is more appropriate to distribute the data on different available processors. In this case, to ensure a better data distribution between different nodes, the most appropriate architecture is distributed memory architecture. Nevertheless, both architectures can be appropriate, if the same program is processed on small amounts of data. Parallelism is expressed by several models of programming like programming by exchange of messages and communication by shared variables. It depends on the considered type of machines, those with shared memory or distributed memory, as well as on the used architecture. We can find in the literature various classifications for parallel architectures. Some examples are classification

of Flynn (Flynn [20]); Taxonomy of Skililcorn (Skililcorn [57]) and the classification of Duncan (Duncan [17]). They are all based on particular and important points of the architecture but the classification of Flynn is the most popular.

There is several work on the parallelization of Monte Carlo methods, but all kinds of variations are still present in a randomly generated sample and the simulation estimates remain affected by such sampling errors through Monte Carlo parallel algorithms already proposed, for example, by Alme et al. [2], Tuffin and Le Ny [68] and Okten and Srinivasan [41].

Given that RDS was selected as the best sampling method to improve upon the precision of the simulation estimates, its parallelization is then suggested to reduce the cost of running simulation experiments and in the same way, it keeps the assets of RDS mainly ensuring that the simulation results are efficient and can be implemented safely in the real world.

9. The Proposed Parallel Simulation

In this subsection, we propose a parallel Monte Carlo simulation using refined descriptive sampling method. The latter's algorithm given in Tari [64] can be regarded as an algorithm of dependent instructions capable to be executed just in sequential way. The only instruction able to be parallelized is the loop of the array filling up with regular numbers that generates a high communication cost given by the following formula:

$$\begin{aligned} \text{communcation cost} &= \text{latency } \textit{time} \\ &\quad + \text{overcost } \textit{time} \\ &\quad + \text{transferred } \textit{time}. \end{aligned}$$

Such as:

Latency time: Initialization time of the networks parameters.

Overcost time: Make-ready time of the message.

Transferred time: Necessary time to transfer the message.

To ovoid such communication cost that can be generated by the parallelization of RDS method, parallelism of the number of replicated runs is proposed for a future directions using RDS to generate input distributions. We have then carried out a parallel Monte Carlo simulation program running only on machines with distributed memory. It is straightforwardly a replication of the whole RDS program on various processors. Given the regular number generation of RDS, its parallelization is easy and simplified regarded to the load balancing. The latter is ensured since the same RDS program runs on all machines. Accordingly, the Single Program Multiple Data (SPMD) model has been selected for a better programming related to the replicated runs and the Message Passing Interface library (MPI Forum [69]) has been suitably chosen to be used with such model. The use of SPMD model will provide an optimal efficiency of the proposed RDS parallel algorithm. The chosen library supports well the language $C/C++$ and FORTRAN and allows the data exchange between the processors. The selected library fits well our requirements; nevertheless, we can find other libraries like Parallel Virtual Machine (PVM) (Gest et al. [21]), Portable Programs for Parallel Processors (Butler and Lusk [10]) and Open Multi Processing (Chergui and Lavallée [12]).

Let us suppose $Q+1$ processors taking part in the simulation experiments. We appoint one processor as a master and the remaining Q by the slaves Ñ processors. We know that the processing time of a parallel program is obtained by the following formula:

$$T_{parallel} = Computation\ time + communication\ cost,$$

where

computation time is the processing time.

In order to carry out a simulation experiment, the master distributes the program of each developing simulation model between the number of slaves processors, such as each processor carries out the same copy of such program. Indeed, in parallelism, the program distribution is done in

turn between the various processors but in the proposed parallelization, the master distributes to each processor the program once together with the number of replicated runs to carry out. The reception of simulation results by the master is done in the same manner. This sort of distribution reduces the communication cost.

We know that the load balancing is one of the most significant concepts in the process of parallelism. Given a simulation experiment of N simulation runs, the load balancing can be introduced if N is greater than Q. We can pass around the load balancing, if the considered developing simulation model is regular that generates a number of runs multiple of a slave processor number. Consequently, instead of carrying out N replicated runs as in the case of a sequential program, each slave processor carries out just N/Q runs. At the end of the simulation, each processor computes and sends its simulation results to the master that computes, in its turn, the final results for each parameter under study. Consequently, these N/Q runs are regarded as one simulation experiment and we can say that the developed simulation model is regular and the load balancing is then not taken into account. This sort of parallelization is called distribution of computation (Lin [32]). In this manner, we reduce the time of parallel running simulation experiments by reducing the computation time.

To ensure the independence of the generated prime numbers related to the sampling method itself, we suppose independent prime numbers generators on each processor. We first generate an unspecified integer number, then, we multiply it by 2 and we add 1 in order to make the generated number, an odd number, finally, we test if the odd number is a prime number by successive division on odd numbers, starting from three with a step of two. If it is a prime number we use it, otherwise, we go back to generate another unspecified integer number. In this manner, we reduce the time allocated to the generation of prime numbers by 75%, we then reduce the time of sequential running simulation experiments and by the way, we reduce again the parallel time by reducing the computation time.

In the proposed parallelization, the number of messages is reduced from two N to two Q, then the influence of the communication cost is

highly controlled and therefore, it is regarded as a wasted time compared to a computation time. Then,

$$T_{parallel} = computation\ time.$$

Furthermore, it is well known that the parallel time of running experiments is inversely proportional (a linear function) to the number of available processors. Then, for a given number of runs, whatever the number of processors is, the parallel simulation results are similar to those obtained sequentially, but the parallel time of running simulation experiments is different and for a given sequential time, it decreases when Q increases according to the following formula

$$T_{parallel} = \frac{T_{sequential}}{Q}.$$

10. Conclusions and Remarks

The use of random sampling, the study of the variability and variance reduction techniques are described above to reduce the sampling errors in a simulation study along with some sampling methods that can be also regarded as variance reduction techniques. These sampling methods include ranked set sampling, systematic sampling, stratified sampling, Latin hypercube sampling, L ranked set sampling, Importance sampling, quasi Monte Carlo methods, descriptive sampling and refined descriptive sampling. A comparison between all these sampling methods was given generating the best sampling. The principal interest of the proposed MC parallelism using the best sampling RDS procedure is its running speed while preserving the assets of the sampling method, its efficiency and safety. Since this method works by replication, then it was naturally parallelizable. Given the scalability advantage of the SPMD model, the proposed parallelization is more appropriate with a big number of replicated runs and it is related to both, the algorithm of RDS procedure and the resolution algorithm of a complex problem that can be seen in a future work.

References

[1] J. S. Ahn and P. Danzig, Packet: network simulation: Speedup and accuracy versus timing granularity, IEEE/ACM Transactions on Networking 4(5) (1996), 743-757.

[2] H. J. Alme, G. Rodrigue and G. Zimmerman, Domain decomposition for parallel Laser-Tissue Models with Monte Carlo Transport, Niederreiter, H. Spanier, J. eds., Monte Carlo and Quasi Monte Carlo Methods, pp. 86-97, Springer, 1998.

[3] A. D. Al-Nasser, *L*-Ranked set sampling: A generation procedure for robust visual sampling, Communications in Statistics: Simulation and Computation 36(1) (2007), 33-43.

[4] E. C. Anderson, Monte Carlo Methods for Inference in Population Genetic Models, 2001.

[5] A. N. Avramidis and P. L'Ecuyer, Efficient Monte Carlo and quasi-Monte Carlo option pricing under the variance-gamma model, Management Science 52(12) (2006), 1930-1944.

[6] J. Banks, J. S. Carson, B. L. Nelson and D. M. Nicol, Discrete- Event System Simulation (3rd edition), Upper Saddle River, New Jersey, Prentice Hall, 2001.

[7] V. D. Barnett, Random Negative Exponential Deviates 2-Pair-Wise Correlated Sets of 10,000 Observations: With Facilities for the Generation of Random χ^2 Deviates on Any Integral Number of Degrees of Freedom, Cambridge University Press, New York, 2005.

[8] P. Bratley, B. L. Fox and L. E. Schrage, A Guide to Simulation (2nd edition), Springer-Verlag, New York, 1987.

[9] M. E. Brenner, Selective sampling-a technique for reducing sample size in simulations of decision making problems, Journal of Industrial Engineering 14 (1963), 291-296.

[10] R. M. Butler and E. L. Lusk, Monitors, message and clusters: The p^4 parallel programming system, Parallel Computing 20(4) (1994), 547-564.

[11] R. C. H. Cheng, The use of antithetic control variates in computer simulations, Winter Simulation Conference, Proceedings of the 13th conference on winter simulation, IEEE press, 1981.

[12] J. Chergui and P. F. Lavallée, Open MP Parallélisation multi tâches pour machines à mémoire partagée, Institut du developpement et des ressources en Informatique scientifique, 2006.

[13] R. Davidson and J. G. Mac Kinnon, Regression-based methods for using control and antithetic variates in Monte Carlo simulation, Journal of Econometrics 54 (1992), 203-222.

[14] T. R. Dell and J. L. Clutter, Ranked set sampling theory with order statistics background, Biometrics 28 (1972), 545-555.

[15] A. Doucet, N. de Freitas and N. Gordon, Sequential Monte Carlo Methods in Practice, Springer Verlag, 2001.

[16] S. Drew and T. Homen de Mello, Some large deviations results for latin hypercube sampling, Winter Simulation Conference, M. E. Kuhl, N. M. Steiger, F. B. Armstrong and J. A. Joines, eds., 2005.

[17] R. Duncan, A survey of parallel computer architectures, Computer 23(2) (1990), 5-16.

[18] S. Ehrenfeld and S. Ben Tuvia, The efficiency of statistical simulation procedures, Technometrics 4 (1962), 257-275.

[19] G. S. Fishman, Monte-Carlo: Concepts, Algorithms and Applications, Springer-Verlag, 1997.

[20] Flynn, Very high speed computing systems, Proceeding IEEE 54(12) (1966), 1901-1909.

[21] A. Gest, A. Beguelin, J. Dongarra, W. Jiang, R. Manchek and V. Sunderam, PVM: Parallel Virtual Machine, a user guide and tutorial for network, Parallel Computing, The MIT press, 1994.

[22] L. K. Halls and T. R. Dell, Trial of ranked set sampling for forage yiels, Forest Science 12 (1966), 22-26.

[23] J. M. Hammersley and K. W. Morton, A new Monte Carlo technique: Antithetic variates, Proc. Cambridge Philos. Soc. 52 (1956), 449-475.

[24] N. Hoshino and A. Takemura, On reduction of finite sample variance by extended Latin hypercube sampling, Bernoulli 6(6) (2000), 1035-1050.

[25] M. G. Kendall and S. Babington, Randomness and random sampling numbers, J. Roy. Stat. Soc. 101 (1938), 147-166.

[26] J. P. C. Kleijnen, Design and analysis of simulations: Practical statistical techniques, Simulation 28 (1977), 81-90.

[27] C. Kwon and J. D. Tew, Strategies for combining antithetic variates and control variates in designed simulation experiments, Management science 40(8) 1994.

[28] A. M. Law and W. D. Kelton, Simulation Modeling and Analysis (3rd ed.), McGraw-Hill, New-York, 2000.

[29] P. L'Ecuyer, Quasi Monte Carlo methods for simulation, Winter Simulation Conference, S. Chick, P. J. Sanchez, D. Ferrin and D. J. Morrice, eds., 2003.

[30] P. L'Ecuyer, C. Lécot and B. Tuffin, Randomized Quasi-Monte Carlo Simulation of Markov Chains with an Ordered State Space, H. Niederreiter and D. Talay, eds., Monte Carlo and Quasi-Monte Carlo Methods 2004, Springer-Verlag, 2006.

[31] P. L'Ecuyer and E. Buist, Variance reduction in the simulation of call centers, L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol and R. M. Fujimoto, eds., Proceedings of the 2006 Winter Simulation Conference, 2006.

[32] Y. B. Lin, Parallel independent replicated simulation on networks of workstations, Proceedings of the 8th Workshop on Parallel and Distributed Simulation, ed., I. C. S. Press, 1994.

[33] W. L. Loh, On Latin hypercube sampling, The Annals of Statistics 24 (1996), 2058-2080.

[34] G. A. McIntyre, A method for unbiased selective sampling using ranked sets, Aust. J. Agri. Res. 3 (1952), 385-390.

[35] M. D. McKay, R. J. Beckman and W. J. Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics 21 (1979), 239-245.

[36] W. J. Morokoff and R. E. Caflisch, Quasi random sequences and their discrepancies, SIAM J. Sci. Comput. Dec: (1994), 1571-1599.

[37] H. A. Muttlak, Parameters estimation in a simple linear regression using rank set sampling, Biometrical Journal 37(7) (2007), 799-810.

[38] B. L. Nelson and B. W. Schmeiser, Decomposition of some well known variance reduction technique, J. Stat. Comput. Simul. 23(3) (1986), 183-209.

[39] H. Niederreiter, Random number generation and quasi Monte-Carlo methods, Philadelphia: CBMS-SIAM 63, 1992.

[40] G. Okten, A probabilistic result on the discrepancy of a hybrid-Monte Carlo sequence and applications, Monte Carlo Methods and Applications 2(4) (1996), 255-270.

[41] G. Okten and A. Srinivasan, Parallel Quasi-Monte Carlo Applications on a Heterogeneous Cluster, K. T. Fang, F. J. Hickernell and H. Niederreiter, eds., Monte Carlo and Quasi- Monte Carlo Methods, Springer-Verlag, 2002.

[42] G. Okten, B. Tuffin and V. Bugaro, A central limit theorem and improved error bounds for a hybrid-Monte Carlo sequence with applications in computational finance, Journal of Complexity 22(4) (2006), 435-458.

[43] A. B. Owen, A central limit theorem for Latin hypercube sampling, J. Roy. Statist. Soc. Ser. B 54 (1992), 541-551.

[44] A. B. Owen, Monte Carlo variance of scrambled net quadrature, SIAM J. Numer. Anal. 34 (5) (1997), 1884-1910.

[45] M. Pidd, Computer Simulation in Management Science (4th ed.), Chichester: John Wiley and Sons, 1998.

[46] R. Robinson, Simulation: The practice of Model Development and Use, Chichester: John Wiley and Sons, 2004.

[47] C. P. Robert and G. Casella, Monte Carlo Statistical Methods (2nd ed.), Springer Science+Business Media Inc., 2004.

[48] K. W. Ross, D. Tsang and J. Wang, Monte-Carlo summation and integration applied to multichain queueing networks, J. Assoc. Comput. Mach. 41(6) (1994), 1110-1135.

[49] S. M. Ross, Simulation, 2nd ed., Academic Press, New York, 1997.

[50] R. Y. Rubinstein, Simulation and the Monte Carlo Method, Wiley, New York, 1981.

[51] E. Saliby, Understanding the variability of simulation results: An empirical study, J. Oper. Res. Soc. 41(4) (1990), 319-327.

[52] E. Saliby, Descriptive Sampling: A better approach to Monte Carlo simulation, J. Oper. Res. Soc. 41(12) (1990), 1133-1142.

[53] E. Saliby and R. J. Paul, Implementing descriptive sampling in three phase discrete event simulation models, J. Oper. Res. Soc. 44 (1993), 147-160.

[54] E. Saliby, Descriptive sampling: an improvement over Latin hypercube sampling, Winter Simulation Conference (1997), 230-233.

[55] E. Saliby and F. Pacheco, An empirical evaluation of sampling methods in risk analysis simulation: Quasi Monte Carlo, Descriptive Sampling, and Latin Hypercube sampling, Winter Simulation Conference (2002), 1606-1610.

[56] R. E. Shannon, Systems Simulation the Art and Science, Prentice-Hall, New-Jersey, 1975.

[57] D. B. Skililcorn, A taxonomy for computer architectures, IEEE Computer 21(11) (1985), 46-57.

[58] I. H. Sloan and S. Joe, Lattice Methods for Multiple Integration Clarendon Press, Oxford, 1994.

[59] M. Stein, Large sample properties of simulations using Latin hypercube sampling, Technometrics 29 (1987), 143-151.

[60] T. Ken Seng and P. P. Boyle, Applications of randomized low discrepancy sequences to the valuation of complex securities, J. Econom. Dynam. Control 24(11-12) (2000), 1747-1782.

[61] M. Tari and A. Dahmani, The three phase discrete event simulation using some sampling methods, IJAMAS 3D 05 (2005), 37-48.

[62] M. Tari and A. Dahmani, The refining of descriptive sampling, IJAMAS 3M05 (2005), 41-68.

[63] M. Tari and A. Dahmani, Flowshop simulator using different sampling methods, Operational Research: An International Journal 5 (2005), 261-272.

[64] M. Tari and A. Dahmani, Refined descriptive sampling: A better alternative to Monte Carlo simulation, Simulation Modelling Practice and theory 14 (2006), 143-160.

[65] S. J. E. Taylor, V. Hlupic, S. Robinson and J. Ladbrook, GroupSim: Investigating issues in collaborative simulation modeling, Oper. Res. Soc. Simul. (2002), 11-18.

[66] K. D. Tocher, The Art of Simulation, English University Press, London, 1963.

[67] B. Tuffin, Variance reductions applied to productform multi-class queueing networks, ACM Trans. Modeling and Computer Simulation 7(4) (1997), 478-500.

[68] B. Tuffin and L. M. Le Ny, Parallélisation d'une combinaison des méthodes de Monte-Carlo et quasi Monte-Carlo et application aux réseaux de files d'attente, RAIRO Operations Research 34 (2000), 85-98.

[69] MPI Forum: April MPI (A Message Passing Interface Standard), 1994.

■