

Fast simulation of wavelength continuous WDM networks

Lachlan L. H. Andrew

Abstract—This paper considers the estimation of blocking probabilities of circuit-switched WDM networks with no wavelength converters and with fixed routing. It presents an importance sampling simulation technique for determining whether or not such a network meets a specific grade of service requirement, in the sense of all routes having blocking below a given threshold. It is especially efficient for networks with high grades of service, which take a long time to simulate using conventional methods.

Index Terms—Quasi-regeneration, M/M/k/k, variance reduction.

I. INTRODUCTION

IN order to dimension networks efficiently, it is important to be able to predict the performance of a given topology. This paper will focus on the blocking probability as the performance measure. Much research has been carried out into evaluating this in the case of product-form networks [1]–[4], in which state probabilities are known in closed form. However, in many cases, there are no closed form expressions for this probability. An important example is wavelength division multiplex (WDM) networks without wavelength converters, i.e., those in which the same wavelength must be used on every link in a path. In this case, blocking occurs when a connection is requested between two nodes, but there is no wavelength free on every link of the chosen path between those nodes.

When the blocking probability is not known explicitly, simulation is the standard method of evaluation. However, there are two common cases in which direct simulation becomes very inefficient.

The first case is when the designer is only interested in the performance of a small subset of the routes, such as the routes with the highest blocking probability. Because of interactions within the network, the whole network must be simulated, despite the fact that most of the performance information is unwanted.

The second case is when the blocking probability is very low, leading to what is known as the rare event estimation problem [5]. This case occurs when it is necessary to determine the blocking rates for a wide range of loads, such as when evaluating a new routing or wavelength assignment strategy. It also occurs in applications such as inter-processor connections for distributed computation [6], [7], where very low blocking is desired, since delays due to blocking can waste significant amounts of expensive CPU time. Such interconnections frequently use expensive non-blocking architectures [8],

but could be made more cheaply from quasi-nonblocking networks, with blocking probabilities of say $B < 10^{-5}$.

Estimating the probability of rare blocking events can be made significantly more efficient by the use of importance sampling (IS) [3], [9]–[15]. This has been applied to many areas of communications [12], but there has been little work on applying it to calculating blocking probabilities [13]–[15]. The principles of IS are outlined in Section II. This paper describes an enhancement to the IS scheme of [14], which improves the performance in both of the application scenarios described above. The scheme of [14] was originally proposed for estimating blocking probability in mobile cellular networks with dynamic channel allocation; these are equivalent to circuit switched wavelength division multiplexing (WDM) networks with fixed routing and wavelength continuity constraints. This algorithm, which is outlined in Section III, was shown to have asymptotically optimal behaviour as blocking probabilities go to zero. However, for moderate blocking probabilities, it did not show any benefit over simpler schemes when finding the average blocking probability over the entire network.

The key contribution of this paper is to show, in Sections IV and V, that the algorithm of [14] actually performs very well, even for moderate blocking probabilities, when not all routes need to be simulated with equal precision. This is typically the case in practice, where only the routes with the highest blocking probability or the strictest quality of service (QoS) requirements need to be evaluated. The blocking probability of individual routes is of greater interest than the average network blocking probability; a network of 50 nodes could in principle have an overall blocking as low as 0.1%, and yet have 100% blocking on one route. Thus, this algorithm is a practical tool for investigating the impact of link capacities, wavelength allocation strategies and routing configurations in WDM networks with static routing.

In the networks considered in this paper, nodes are connected by optical links capable of carrying Λ wavelengths; for clarity of exposition, Λ will be assumed equal for all links, although the algorithm does not require this. When a call is established between two nodes, it will select and use a wavelength which is free on all links on the (static) route between the nodes. If all wavelengths are used on at least one link in the route, the call is blocked. It will be further assumed that calls arrive to routes (that is, to origin/destination pairs) according to a Poisson process, and that holding times are exponentially distributed.

After demonstrating the effectiveness of the proposed simulation algorithm on these networks, the important issue of determining when to terminate the simulation will be addressed in Section VI.

Lachlan Andrew is with the ARC Special Research Centre for Ultra-Broadband Information Networks, Department of Electrical and Electronic Engineering, The University of Melbourne, Victoria 3010, Australia. Email: l.andrew@ee.mu.oz.au. This work was supported by the Australian Research Council.

II. IMPORTANCE SAMPLING

Importance sampling [11] is a technique to improve the accuracy of estimates of stochastic events. Rather than simulating the real stochastic system, a system is simulated in which the “important” events occur more frequently. This must be done in a controlled way, so that the bias introduced can be exactly calculated, and then removed. Consider the simple one dimensional case of estimating the probability $\mathbf{P}(X < \theta)$ for a random variable X with probability density function (pdf) $f(\cdot)$. This can be done by repeatedly generating samples of X and counting the number below θ , that is, by calculating $\mathbf{E}[\mathbf{1}_{\{X < \theta\}}]$, where $\mathbf{1}_{\{A\}}$ is 1 if A is true, and 0 otherwise. However, if $\mathbf{P}(X < \theta)$ is very small, many samples will be required. It is more efficient to generate samples with a pdf $f^*(\cdot)$ with $\mathbf{P}^*(X < \theta) > \mathbf{P}(X < \theta)$. (Here \mathbf{P}^* and \mathbf{E}^* denote probability and expectation with respect to the distribution f^* .) How can $\mathbf{P}(X < \theta)$ be estimated from these distorted simulations? If, instead of incrementing a count by 1 each time $x < \theta$, the count is incremented by $f(x)/f^*(x)$, it is estimating

$$\int_{-\infty}^{\theta} \frac{f(x)}{f^*(x)} f^*(x) dx = \int_{-\infty}^{\theta} f(x) dx = \mathbf{P}(X < \theta).$$

The function $L(x) = f(x)/f^*(x)$ is called the Radon-Nikodym derivative, or likelihood ratio. Importance sampling is unbiased, as long as the likelihood function can be evaluated exactly. By choosing $L(x)$ appropriately, the variance of an estimate of $\mathbf{E}[X]$ can be reduced significantly. Although IS is often called “fast simulation”, it is important to note that it works by reducing the variance of a stochastic estimate, rather than allowing that estimate to be evaluated in less time. Reducing the variance reduces the number of samples which must be averaged to obtain a reliable estimate, which provides the speed improvement.

III. ALGORITHM

The fast simulation algorithm of [14] will now be briefly described. For a justification of the algorithm, and its theoretical properties, see [14]. The algorithm focuses on routes and collections of routes, rather than on individual links. The state information maintained about the network is the set of wavelengths currently used on each route, and the topology is represented as sets of routes which share links.

Let $\mathbf{n}(t)$ be a vector of the number of connections on each route at time t , and n be the total number of connections. Let the set of routes which use link i be called the i th clique, c_i , and let the j th cluster, \mathcal{C}_j , be the union of all cliques containing route j , i.e., those routes intersecting route j . Arrivals to route j are assumed Poisson with rate λ_j , and holding times are assumed independent with a negative-exponential distribution of mean $1/\mu$.

The simulation is based on a “backbone” Markov chain simulation, known in the simulation literature as the standard clock technique [16]. When the total occupancy of the current state is n , the time until the next event is an exponential random variable with parameter

$$\Lambda_n = \sum_j \lambda_j + n\mu.$$

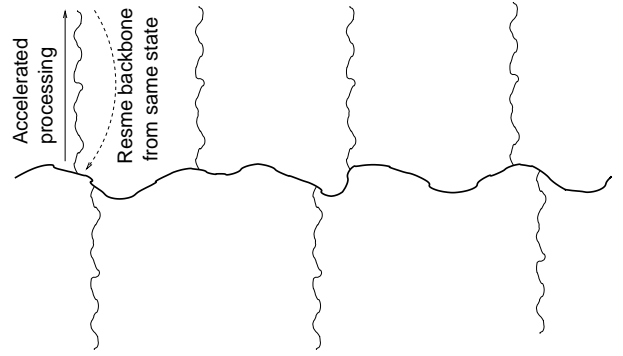


Fig. 1. “Backbone and ribs” simulation framework.

The event is declared an arrival to route j with probability λ_j/Λ_n , or a call departure from route j with probability $n_j\mu/\Lambda_n$, where n_j is the number of calls at route j , so that $n = \sum_j n_j$. If the event is a departure, then a randomly chosen wavelength on route j is released. If it is an arrival then blocking may occur. If a wavelength can be found which is free on every link of route j , then the call is accepted, and one such wavelength is marked as busy. The choice of wavelength reflects wavelength allocation strategy of the actual network. If no such wavelength exists, the call is blocked.

For the description of the IS scheme, let i be a route whose blocking probability is to be determined, and m_i denote the number of current connections in the i th cluster. Let $\lambda = \sum_{j \in \mathcal{C}_i} \lambda_j$ be the sum of the arrival rates of all routes in cluster \mathcal{C}_i . Finally, let θ_i be an occupancy threshold that specifies the number of connections in cluster i which should trigger importance sampling. The “backbone” simulation described above is paused after every arrival which causes m_i to increase to θ_i . At these instants, an importance-sampled “rib” is started by creating a copy of the current network state, as illustrated in Figure 1. (Note that a single arrival in the backbone can trigger ribs for multiple routes.) The simulation of the rib is split into two phases: an IS phase, and a recovery phase.

During the IS phase, arrivals and departures to routes not in cluster \mathcal{C}_i occur as normal. However, the probability of an arrival to route $j \in \mathcal{C}_i$ is increased by a factor of $\lambda^*(m_i)/\lambda$, and the probability of a departure from route $j \in \mathcal{C}_i$ is scaled by a factor of $\mu^*(m_i)/\mu$, where $\lambda^*(\cdot)$ and $\mu^*(\cdot)$ satisfy

$$\mu^*(m) = \frac{\lambda\mu}{\lambda^*(m-1)} \quad (1a)$$

$$\lambda^*(m) = \lambda + m(\mu - \mu^*(m)), \quad (1b)$$

starting from $\mu^*(\theta_i) = 0$. This increases the probability of the cluster occupancy increasing, increasing the probability of blocking.

To overcome the bias introduced by the importance sampling, the Radon-Nikodym derivative, L , must be evaluated. At the start of the rib, L is initialised to 1. For every arrival, it is multiplied by $\lambda/\lambda^*(m_i)$, and for every departure it is multiplied by $\mu/\mu^*(m_i)$.

Once the system enters a “blocking” state for route i (one in which an extra arrival to route i would be blocked), the IS phase is terminated and the recovery phase begins. During the

recovery phase, the true arrival and departure rates, λ_j/Λ_n and $n_j\mu/\Lambda_n$, are used for all routes in the network. The recovery phase ends once the cluster occupancy, m_i , drops below θ_i . At the end of the recovery phase, the simulation of the backbone continues where it left off at the start of the rib.

Throughout the simulation of the rib, the total amount of time, X_i , spent in blocking states for route i is recorded. By construction, the mean value $\mathbf{E}^*[LX_i]$ is then equal to $\mathbf{E}[X_i]$, the expected amount of time that would have been spent in blocking states had no importance sampling been applied. That is to say that $\mathbf{E}[X_i]$ can be estimated by

$$\frac{1}{S} \sum_{j=1}^S L(j)X_{ij}, \quad (2)$$

where X_{ij} denotes the time spent in blocking states for route i while simulating the j th rib for route i , $L(j)$ denotes the corresponding Radon-Nikodym derivative, and S is the total number of ribs for route i .

The times (along the backbone) between the starts of ribs for route i are *quasi-regeneration cycles* [17], [10], or “A-cycles”. Like the more familiar true regeneration cycles, the behaviour of the process within quasi-regeneration cycles is identically distributed. In particular, if T_i is the duration of a quasi-regeneration cycle for route i , then the proportion of time spent in blocking states is equal to [18]

$$B_i = \frac{\mathbf{E}[X_i]}{\mathbf{E}[T_i]}. \quad (3)$$

Since Poisson arrivals see time averages, B_i is the blocking probability for route i .

The time required to simulate a rib can be reduced by artificially reducing both the arrival and departure rates of routes not in the cluster of interest. This induces a slight bias [19], but for simulations of short duration, this may be more than offset by the reduction of variance resulting from the increased number of iterations that are made possible.

As an aside, note that the final value of L for the rib is

$$L = \prod_{k=\theta_i}^{M_i} \left(\frac{\lambda}{\lambda + k(\mu - \mu^*(k))} \right)^{1+b(k)}, \quad (4)$$

where M_i is the occupancy of the cluster at the end of the IS phase, $b(k) \geq 0$ is the number of calls blocked in cells within cluster \mathcal{C}_i when its occupancy was k . Most of the variance of L , and hence of X_i , comes from the variation of M_i , the cluster occupancy when blocking first occurs. In particular, by accelerating all cells in the cluster, rather than only those in the clique which eventually causes blocking, the algorithm generally causes M_i to be higher than the true average cluster occupancy at blocking, leading to over-acceleration.

A single backbone simulation can be shared by all routes, but separate ribs are required by each route. Combined with the fact that ribs often contain many more events than the corresponding cycle in the backbone, this can lead to excessive simulation time if the blocking probability of every route in a large network is to be determined by importance sampling. There are two common cases in which importance sampling

need only be applied to a subset of the routes. The first, which is discussed in detail in the following section, is when the goal is to identify the routes with the highest blocking probability and find the blocking probability of those routes. The second case is when some routes require a particularly high grade of service, while others have more lenient requirements. For example, this may occur if some routes are used to carry alarm signals while others are not. In this case, IS can be applied to the routes with very low blocking probabilities, while the higher blocking probabilities of the other routes can be obtained from the backbone using traditional simulation techniques. Such techniques include simply counting the proportion of blocked arrivals, or recording the proportion of time spent in blocking states.

IV. ROUTES WITH HIGH BLOCKING

A network designer is usually primarily interested in those routes which experience the highest blocking, since those are the ones which may need upgrading, or may infringe grade-of-service contracts. However, trying to simulate only the routes with high blocking leads to the infinite regress of needing know the result of the simulation to determine its input.

Routes with high blocking can be identified during the warmup phase, necessary for the backbone to reach equilibrium. The key is that, even if blocking states are not reached, routes with high blocking probability will on average have fewer wavelengths available when they establish a connection. Thus the routes with the highest blocking will also have large values of

$$\tilde{B}_i = \sum_k 2^{w(i,k)},$$

where $w(i,k)$ is the number of wavelengths unavailable on route i the k th time it attempts to establish a call, and the summation is over all call attempts on route i .

Because of the imperfect correlation between \tilde{B}_i and blocking, it is prudent to estimate the blocking for several routes in addition to the one with the largest \tilde{B}_i . In this study, ribs were simulated for the four routes with the largest \tilde{B}_i s after the warmup. In addition, \tilde{B}_i was continually estimated along the simulation backbone, and checked intermittently. Routes which were in the top four at any inspection point were also simulated for the full quota of A-cycles.

V. PERFORMANCE EVALUATION

A. Performance criterion

Fast simulation ultimately aims to minimise the simulation time required to evaluate a quantity with sufficient accuracy. To achieve this, both the time taken to generate an individual estimate and the error of that estimate must be considered. In particular, the statistical variance can be reduced k -fold by averaging k estimates from independent simulations. This tradeoff is quantified by the *efficiency* of an unbiased estimator, \hat{X} , defined as

$$\mathcal{E} = \frac{1}{\text{Var}[\hat{X}] \text{CPU}[\hat{X}]},$$

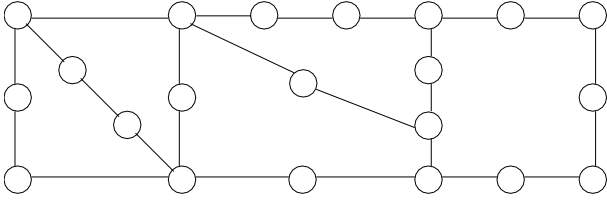


Fig. 2. ARPA2 network topology.

where $\text{CPU}[\hat{X}]$ is the time taken to generate the estimate \hat{X} . When the percentage error is of more interest than the absolute error, the appropriate measure is the relative efficiency:

$$\mathcal{E}_r = \frac{\hat{X}^2}{\text{Var}[\hat{X}]\text{CPU}[\hat{X}]}.$$

This study will use an equivalent but more intuitive measure: the effective time, \mathcal{T}_r , defined as the time that it takes to run enough iterations to achieve 10% relative error ($\text{Var}[\hat{X}] = (0.1\hat{X})^2$). This is related to efficiency by

$$\mathcal{T}_r = \frac{100}{\mathcal{E}_r}.$$

In order to evaluate the importance sampling algorithm, it will be compared with a simple benchmark algorithm which does not use A -cycles, and instead simply simulates the network, and records the time spent in blocking states.

The efficiency of the benchmark relies on being able to determine quickly whether or not the network is in a blocking state for a given route, since this must be done for each route in the network at each call arrival or departure (not just the route of the call in question). The simulator keeps track of which wavelengths are known to be free on a given route, and only when none are known to be free does it perform the costly task of finding which wavelengths are currently free on each link in the route. Each time a connection is set up, it is marked as “not known to be free” on all routes intersecting that connection, i.e., for the cluster. When a departure occurs, the wavelength cannot simply be marked as “free” for the whole cluster, since it may be blocked by other intersecting routes, and so departures are ignored. This heuristic is very efficient for first-fit wavelength assignment, because the last-choice wavelength is usually known to be free for all routes.

B. Numerical results

The algorithm described above was tested on the standard ARPA2 topology, shown in Figure 2. This is a sparse irregular network with 21 nodes and 26 links. The load on each origin/destination pair was equal, and the capacity of each link was equal. Shortest path routing was used, which took no account of the load placed on each link. This led to very unequal blocking probabilities on different links.

The method of batch means [20] was used to determine the variance of the blocking estimates, with 1000 batches of 1000 ribs per route. As in [14], ribs were only simulated for every 10th A -cycle on the backbone, to reduce correlation in the results.

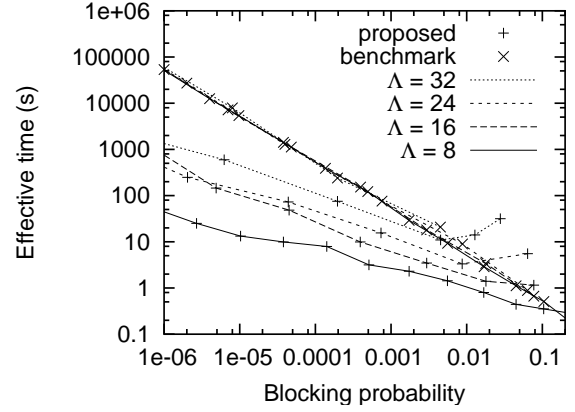


Fig. 3. Effective time vs blocking probability for ARPA2 topology, with 8 to 32 wavelengths per link.

The effective time, \mathcal{T}_r of both the importance sampling algorithm and the benchmark is shown in Figure 3, for $\Lambda = 8, 16, 24$ and 32 wavelengths per link, and a range of loads. The speed-up factor for a given blocking probability is measured by the vertical distance between the corresponding curves. These results confirm the finding of [14] that the effective time of the proposed approach increases much less rapidly as the blocking probability decreases than that of the benchmark, due to the asymptotic optimality of the underlying importance sampling.

As expected, the effective time of the benchmark algorithm depends only on the blocking probability, and not the number of wavelengths. However, the effective time of the proposed method increases as the number of wavelengths on each link increases. This is because of the “over-acceleration” effect described in Section III.

Further tests were carried out on random topologies of varying size and sparsity, generated using the Georgia Tech Internet Topology Modeling software [21]. Figure 4 shows the effective times for networks of 20, 30, 40 and 50 nodes, each with an average degree of three links per node, with $\Lambda = 8$ wavelengths. As the size of the networks increases, both the proposed algorithm and the benchmark become slower. For the increase by a factor of 2.5 in the number of nodes, the number of routes increases by a factor of 6.25. Moreover, the expected length of the routes also increases. Thus the observed increase by a factor of around 10^1 is to be expected. Importantly, both algorithms scale by a similar factor, and the proposed algorithm retains its advantage.

Figure 5 shows the effective times for 30 node networks with mean degrees of 4.4, 5.5, 6.8, 7.6 and 8.1 links per node. The degree of connectivity does not appear to have a significant impact on the performance. In each case, the proposed importance sampling algorithm is substantially faster than the benchmark algorithm.

VI. STOPPING CRITERION

A simulation should be run long enough that the statistical variance of the result is sufficiently small. In general, it is difficult to know the true variance of the simulation result. In the same way that the sample variance of Bernoulli trials is

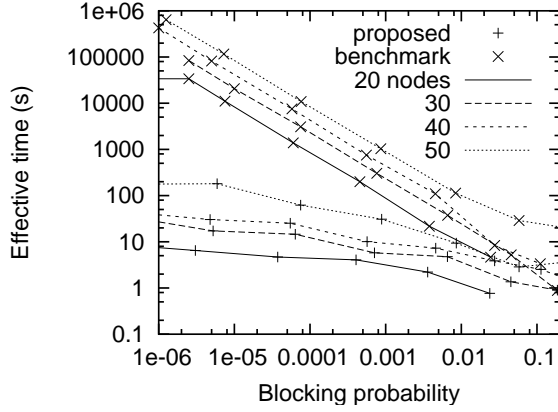


Fig. 4. Effective time vs blocking probability for networks with $\Lambda = 8$, three links per node, and between 20 and 50 nodes.

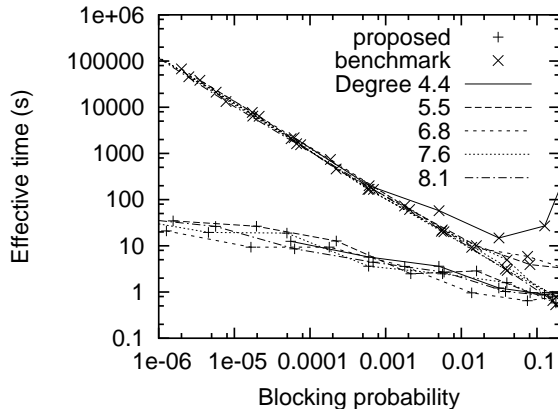


Fig. 5. Effective time vs blocking probability for 30 node networks, with $\Lambda = 8$, and average numbers of links per node between 4.4 and 8.1.

0 when no “successes” have occurred, so the sample variance of IS can be wildly different from the true variance when too few samples are taken. However, this is much harder to detect, because the sample variance is not exactly zero.

A suitable heuristic criterion to determine when IS results are sufficiently accurate is:

- 1) The sample standard deviation is sufficiently low **and either**
- 2) The largest theoretically possible Radon-Nikodym derivative, L , has been observed **or**
- 3) The impact on the estimated blocking caused by the largest observed value(s) of L is smaller than that caused by smaller values.

Note that using the sample variance as the stopping criterion introduces a slight negative bias in the estimated width of the confidence interval (although not of the estimated blocking probability). However, this effect is negligible if sufficient samples are taken.

The remainder of this section will present a heuristic justification of this criterion. To illustrate the problem, consider the simulation of a 9-node ring with $\Lambda = 25$ wavelengths per link and $\lambda = 0.7$ Erlangs per route. Simulations were performed to estimate the blocking probability of the nine four-hop paths. One set of simulations took 20 sample points

180 samples		9,000,000 samples	
mean	std. dev.	mean	std.dev
1.309e-07	2.873e-07	1.082e-06	4.904e-08
1.461e-07	2.293e-07	1.118e-06	5.350e-08
1.171e-08	4.367e-08	1.130e-06	5.833e-08
2.636e-06	1.055e-05	1.082e-06	5.001e-08
2.069e-07	7.574e-07	1.090e-06	5.123e-08
2.214e-06	4.594e-06	1.073e-06	4.397e-08

TABLE I

ESTIMATED MEAN AND ERROR IN BLOCKING ESTIMATES

for each path (180 samples in total), and the other took 1,000,000 (9,000,000 samples in total). Table I shows the estimated blocking probabilities and the estimated error (standard deviation) for six different random seeds. Neglecting correlations between the samples, the true standard deviation of the estimators will differ by a factor of $\sqrt{50000} \approx 200$, however the estimated standard deviation only drops by around one order of magnitude. This indicates that the small sample often produces an under estimate of the variance and cannot be relied upon as a stopping criterion. Moreover, in three of the six cases, the estimated 2σ confidence interval for the 180 sample case does not include the true blocking probability. This effect occurs primarily when the number of samples is less than the ratio of the maximum value of L to the true blocking probability.

To aid the explanation for this, let the possible values of L be divided into buckets whose widths form a geometric sequence, with the j th bucket centred around L_j . Let $(LY)_j$ denote the sum of those terms in (2) for which $L(j)$ is in the j th bucket, and let $Y_j = (LY)_j / L_j$.

A short simulation only estimates the probabilities of events which are most likely under the changed measure. This is seen in Figure 6, which plots Y_j against L_j for the two simulations described above. Clearly the short simulation only estimates the peak of the curve. (In this case, the different values of Radon-Nikodym derivative, L , correspond primarily to the occupancy of the cluster at the time acceleration stops, i.e., when the route of interest is first blocked. The “grass” at the bottom of this graph corresponds to A -cycles in which blocking occurs on one of the routes in the cluster but not on the route of interest, yielding non-zero $b(k)$ in (4), and a smaller value of L . However, this case contributes little to the blocking.)

The results look very different when the “correction factor” for importance sampling, L , is included. Figure 7 shows a bucket plot of the individual terms, $(LY)_j$, of the sum (2). This is similar to the plots used in [4], and the blocking estimate is the sum of the buckets. This clearly shows that, although the most *common* cases are well represented, those which contribute most to blocking are not well represented. In this case, the estimated blocking probability will be well below the true value. More importantly, the sample variance will also be much lower than the ensemble variance, resulting in an inaccurate confidence interval.

In order to determine how long the simulation must be run, consider the following heuristic argument. Simulation must be

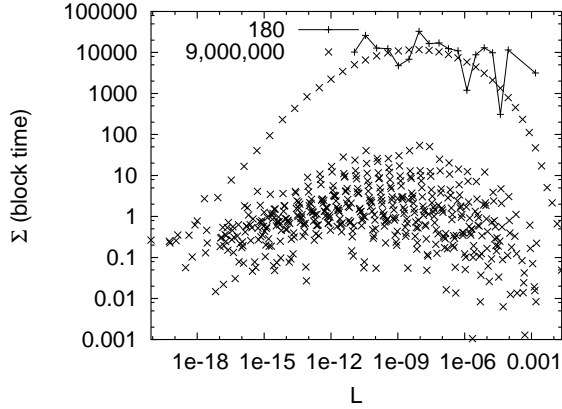


Fig. 6. Total blocking time in A -cycles, Y , against the acceleration applied to the A -cycle, L , on a logarithmic scale. Points represent the totals for all L values in a range. These buckets have equal size on the logarithmic scale shown here. For clarity, the points corresponding to the short simulation are connected by lines.

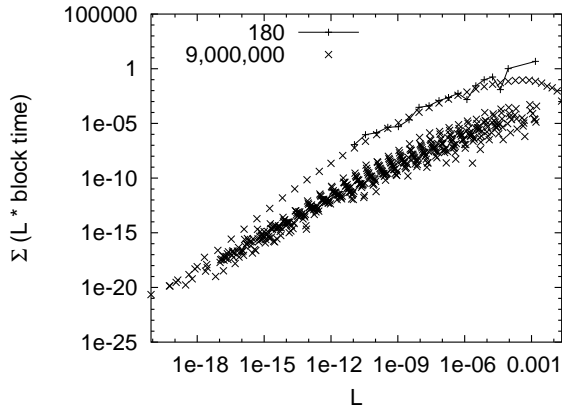


Fig. 7. Total contribution, (LY) , made by A -cycles against the acceleration applied to the A -cycle, L , on a logarithmic scale. Points represent the totals for all L values in a range. These buckets have equal size on the logarithmic scale shown here.

long enough to sample from all “important” events, those with large values of Y_j . Events which are not accelerated much will probably not occur at all in short simulations. If such events have a small impact, they can be ignored. Assume that under-accelerated events (those with very large L_j) are generally of little importance, and the importance of each under-accelerated event decreases as the acceleration given to it decreases. If this importance, Y_j , decreases fast enough for the product $(LY)_j$ also to decrease, then the rarest events can be ignored. This corresponds to condition (2) of the proposed stopping criterion.

Under ideal IS, most samples will be accelerated by the same amount, and there will be no samples with significantly higher values of L . An example of this is shown in Figure 8, for $\Lambda = 5$ wavelengths and $\lambda = 0.02$ Erlangs per route. Here $(LY)_j$ does not decrease for large L_j , even when very many samples are taken. How can this be distinguished from the case of insufficient samples described above? The key is that the maximum possible value of L is known: it occurs when all events in the accelerated portion of the A -cycle reduce the number of channels available to the route of interest ($M_i =$

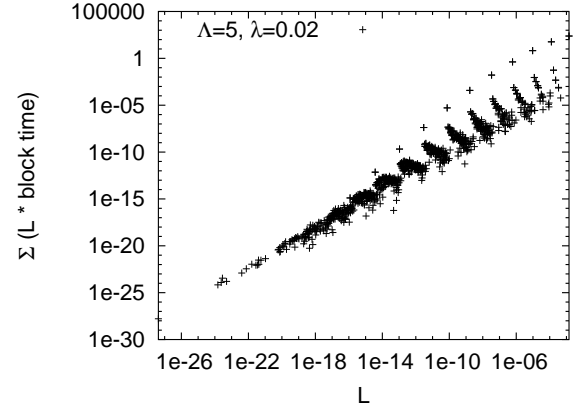


Fig. 8. Total contribution made by A -cycles against the acceleration applied to the A -cycle, L , on a logarithmic scale. Points represent the totals for all L values in a range. These buckets have equal size on the logarithmic scale shown here.

Λ , $b(s) = 0$ in equation 4). In this case, the maximum is $L \approx 0.00138062$, corresponding to the peak of the graph. This means that there can be no events which are still rare under the changed measure but which have L values larger than those observed, which contribute to the ensemble variance but not to the sample variance. The estimated standard deviation is thus reliable. This corresponds to condition (3) of the proposed stopping criterion.

VII. CONCLUSION

Importance sampling can improve the speed of estimating blocking probabilities in circuit switched networks with fixed routing by several orders of magnitude, with the greatest benefits at low blocking probabilities. If the blocking probability of only a small subset of the routes needs to be found, then IS requires less time (for a given accuracy) than direct simulation, even for moderate to high blocking probabilities. The IS approach here is applicable to any wavelength assignment technique, and is well suited to evaluating different static routing arrangements, or different wavelength assignment strategies.

REFERENCES

- [1] G. L. Choudhury, K. K. Leung, and W. Whitt, “An inversion algorithm to compute blocking probabilities in loss networks with state-dependent rates,” *IEEE/ACM Trans. Networking*, vol. 3, pp. 585–601, Oct. 1995.
- [2] J. L. Coleman, W. Henderson, and P. G. Taylor, “Product form equilibrium distributions and a convolution algorithm for stochastic Petri nets,” *Perf. Eval.*, vol. 26, pp. 159–180, Sept. 1996.
- [3] M. Falkner, M. Devetsikiotis, and I. Lambadaris, “Fast simulation of networks of queues with effective and decoupling bandwidths,” *ACM Trans. Model. Comput. Simul.*, vol. 9, pp. 45–58, Jan. 1999.
- [4] L. L. H. Andrew and F. J. Vázquez-Abad, “Filtered Gibbs sampler for estimating blocking in product form networks,” in *Proc. IASTED Wireless and Optical Communications*, Banff, Canada, July, 2002, pp. 527–532.
- [5] V. V. Kalashnikov, *Geometric sums, bounds for rare events with applications : risk analysis, reliability, queueing*. Boston, MA: Kluwer Academic, 1997.
- [6] B. Webb and A. Louri, “A class of highly scalable optical crossbar-connected interconnection networks (SOCNs) for parallel computing systems,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 11, pp. 444–458, May 2000.

- [7] G. Zhou, Y. Zhang, and W. Liu, "Optical fiber interconnection for the scalable parallel computing system," *Proc. IEEE*, vol. 88, pp. 856–863, June 2000.
- [8] Y. Yang and J. Wang, "Wide-sense nonblocking Clos networks under packing strategy," *IEEE Trans. Computers*, vol. 48, pp. 265–284, Mar. 1999.
- [9] P. Heidelberger, "Fast simulation of rare events in queueing and reliability models," *ACM Trans. Model. Comput. Simul.*, vol. 5, pp. 43–85, Jan. 1995.
- [10] C.-S. Chang, P. Heidelberger, and P. Shahabuddin, "Fast simulation of packet loss rates in a shared buffer communications switch," *ACM Trans. Model. Comput. Simul.*, vol. 5, no. 4, pp. 306–325, 1995.
- [11] G. S. Fishman, *Monte Carlo: Concepts, Algorithms, and Applications*. New York, NY: Springer-Verlag, 1996.
- [12] J. K. Townsend, Z. Haraszti, J. A. Freebersyser, and M. Devetsikiotis, "Simulation of rare events in communications networks," *IEEE Communications Magazine*, vol. 36, pp. 36–41, Aug. 1998.
- [13] M. Mandjes, "Fast simulation of blocking probabilities in loss networks," *European Journal of Operations Research*, vol. 101, no. , pp. 393–405, 1997.
- [14] F. J. Vázquez-Abad, L. Andrew, and D. Everitt, "Estimation of blocking probabilities in cellular networks with dynamic channel assignment," *ACM Trans. Model. Comput. Simul.*, vol. 12, pp. 54–81, 2002.
- [15] E. Hyytia and J. Virtamo, "Adaptive importance sampling in routing and wavelength assignment problem," *Eur. Trans. Telecommun.*, vol. 13, pp. 331–339, Jul.–Aug. 2002.
- [16] P. Vakili, "Using a standard clock technique for efficient simulation," *Operations Research Letters*, vol. 10, pp. 445–452, 1991.
- [17] A. Gaivoronski and E. Messina, "Optimization of stationary behavior of general stochastic discrete event dynamic systems," in *Proceedings of International Workshop on Discrete Event Systems, WODES '96*, London, UK, Institute of Electrical Engineers, 1996, pp. 238–243.
- [18] L. Breiman, *Probability. Classics in Applied Mathematics*, Philadelphia, PA: SIAM, 1992.
- [19] L. L. H. Andrew, "Wavelength continuous WDM networks with non-uniform service rates," in *Proc. IASTED Wireless and Optical Commun. Conf.*, Banff, Canada, 2002, pp. 356–359.
- [20] C. Alexopoulos and A. Seila, "Output data analysis," in *Handbook of Simulation*, J. Banks, Ed., New York, NY: John Wiley and Sons, 1998, ch. 7, pp. 225–272.
- [21] E. W. Zegura *et al.*, "GT-ITM," Last updated 26 July, 2000. <http://www.cc.gatech.edu/projects/gtitm>.

PLACE
PHOTO
HERE

Lachlan Andrew (M '97) received his B.Sc. in computer science in 1992, B.E. in electrical engineering in 1993 and Ph.D. in engineering in 1996, all from the University of Melbourne, Australia, where he is a currently research fellow. His research interests include performance analysis and resource allocation in optical and wireless communication networks. Specific interests are optical burst switching, routing techniques, flow control and cellular and ad hoc networks. He is a member of the IEEE and the IEE.