

University of Wollongong  
**Research Online**

---

Faculty of Informatics - Papers (Archive)

Faculty of Engineering and Information  
Sciences

---

November 2001

## An Optimal Importance Sampling Method for a Transient Markov System

H. Qi

*University of Wollongong*

Y. C. Wei

*Taiwan Defence Ministerial Office, Taipei*

L. Wei

*University of Wollongong*

Follow this and additional works at: <https://ro.uow.edu.au/infopapers>



Part of the [Physical Sciences and Mathematics Commons](#)

---

### Recommended Citation

Qi, H.; Wei, Y. C.; and Wei, L.: An Optimal Importance Sampling Method for a Transient Markov System 2001.

<https://ro.uow.edu.au/infopapers/75>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: [research-pubs@uow.edu.au](mailto:research-pubs@uow.edu.au)

---

# An Optimal Importance Sampling Method for a Transient Markov System

## Abstract

In this paper an optimal importance sampling (IS) method is derived for a transient markov system. Several propositions are presented. It is shown that the optimal IS method is unique, and it must converge to the standard Monte Carlo (MC) simulation method when the sample path length approaches infinity. Therefore, it is not the size of the state space of the Markov system, but the sample path length, that limits the efficiency of the IS method. Numerical results are presented to support the argument.

## Keywords

Markov processes Monte Carlo methods, importance sampling, optimisation

## Disciplines

Physical Sciences and Mathematics

## Publication Details

This paper was originally published as: Qi, H, Wei, YC & Wei, L, An Optimal Importance Sampling Method for a Transient Markov System, Global Telecommunications Conference, 2001. GLOBECOM '01. 25-29 November 2001, 2, 1152-1156. Copyright IEEE 2001.

# An Optimal Importance Sampling Method for a Transient Markov System

Honghui Qi<sup>1</sup>, Michael (Y. C.) Wei<sup>2</sup>, Lei Wei<sup>1</sup>

<sup>1</sup>School of Electrical, Computer and Telecommunications Engineering,  
University of Wollongong, NSW2522, Australia

<sup>2</sup>Taiwan Defense Ministerial Office, Taipei

*Abstract*—In this paper an optimal importance sampling (IS) method is derived for a transient Markov system. Several propositions are presented. It is shown that the optimal IS method is unique, and it must converge to the standard Monte Carlo (MC) simulation method when the sample path length approaches infinity. Therefore, it is not the size of the state space of the Markov system, but the sample path length, that limits the efficiency of the IS method. Numerical results are presented to support the argument.

## I. INTRODUCTION

Analytical computation (AC) and Monte Carlo (MC) simulation are two basic tools for determining the performance of a system or network. For a simple task the analytical method is favorable. However, for a sophisticated task a mathematical approach is often intractable, or, at least, the analysis becomes complex. In such cases, a simulation approach is often used. MC simulations can produce very accurate estimates provided that the number of simulation trials is sufficiently large. However, this condition can be severe. For example, for a 95% confidence interval of  $[2P(e)/5, 8P(e)/5]$  [1], where  $P(e)$  denotes error probability, the standard MC approach requires no less than  $10^8$  simulation trials for  $P(e) = 10^{-7}$  [2] and [3].

A fast simulation technique, known as importance sampling (IS), has been developed to reduce the number of simulation trials and has perhaps the most potential for offering substantial run-time savings in code construction.

In [4], Shanmugan and Balaban introduced the IS concept to study the BER estimation for some simple communication systems. The method proposed in [4] is to scale the original noise density to increase the probability of simulation samples taken from the importance regions. The results of [4] and [5] showed that the efficiency of the scaled IS method is limited by the state space size (which is often referred to as the memory size) of the system. In general, the larger the memory size, the less efficient the scaled IS method is. Since then, many IS methods have been proposed and studied in telecommunication fields. A summary of this work can be found in [1].

In [17], Wei proposed an optimal conditional importance sampling (OCIS) method. OCIS has recently been identified as a method often used in practice without knowledge of its connection with IS techniques. The OCIS method is optimal under the condition that only certain system parameters can be biased. Several applications can be found in [18]-[20].

Unfortunately, the OCIS method is only applicable to systems without memory. It is a difficult problem to construct an IS method for systems with a large scale state space (or large

memory). First used by Cottrell et al. [7], the theory of large deviations techniques (LDT) was developed to design an efficient IS simulation method in the communication field. This work has been developed into a coherent simulation methodology through a series of papers by Sadowsky, Bucklew, and many others, [5], [8]-[11]. In the 1990s, important contributions include error event simulation [9] and suboptimal biased distributions [12] [13]. The rich history and development of this concept can be found in survey papers [14] and [15].

In this paper, we will derive an optimal IS method for a transient Markov system. The efficiency of the method will also be studied. This paper is organized as follows. In Section II a transient Markov system is introduced and the MC and IS methods are reviewed. In Section III, an optimal IS (OIS) method for this transient Markov system is derived and several propositions are presented. The efficiency of the OIS method is then studied in Section IV. Section V concludes the paper.

## II. A TRANSIENT MARKOV SYSTEM AND REVIEW OF MC AND IS METHODS

Consider a transient Markov system with a total of  $S$  states given in Figure 1. Suppose that we are interested in estimating

$$P_L \equiv \sum_{t=1}^L f(X_t = 0 | X_0 = 1) \pi_0(1). \quad (1)$$

where  $\pi_0(i) = \pi_0(X_0 = i)$  is the initial probability of the state  $X_0 = i$  at step 0 and  $f(X_t = 0 | X_0 = i) \equiv P(X_t = 0, X_{t-1} \neq 0, \dots, X_1 \neq 0 | X_0 = i)$  is the first passage time probability, defined as the probability that the first entry to state 0 occurs at a time  $t$   $1 \leq t \leq L$  given that  $X_0 = i$ . In other words,  $f(X_t = 0 | X_0 = i)$  is the probability of hitting state 0 in  $L$  or fewer steps.  $P_L$  tends to 1 as  $L \rightarrow \infty$ .

Clearly, when  $S$  is small, we can compute  $P_L$  directly using (1). However, when  $S$  is large, say  $S = 10^6$ , it becomes very difficult to compute  $P_L$  (at least it is computationally impractical). Thus, we might be obliged to use simulation instead.

**Standard MC method:** The complexity of the MC simulation method is independent of the number of states ( $S$ ). It is a more flexible way to estimate  $P_L$ . The MC approach for estimating  $P_L$  involves the following two steps:

**Step 1** Generate  $N$  independent sample vectors (paths)  $\mathbf{Y}_1, \dots, \mathbf{Y}_N$  from  $\{X_t, t \geq 0\}$  where  $\mathbf{Y}_j = [y_{j,0}, \dots, y_{j,L}]$ ,  $y_{j,i}$  is the state at time  $i$  of the path  $j$

**Step 2** Form the estimator

$$\hat{P}_L = \frac{1}{N} \sum_{j=1}^N 1_{\mathbf{Y}}(\mathbf{Y}_j), \quad (2)$$

where  $1_{\mathbf{Y}}(\cdot)$  is the one-zero indicator function, that is  $1_{\mathbf{Y}}(\mathbf{Y}_j) = 1$  if  $y_{j,i} = 0$  for any  $i \in \{0, \dots, L\}$ , and  $1_{\mathbf{Y}}(\mathbf{Y}_j) = 0$  otherwise.

Since  $E_{\mathbf{Y}}[1_{\mathbf{Y}}(\mathbf{Y}_j)] = P_L$ , this estimator is unbiased. Assuming that the simulation trials are independent of each other, the variance of this estimator is  $\sigma_{\hat{P}_L}^2 = \frac{P_L - P_L^2}{N}$ . Thus, when  $P_L$  is the probability of a rare event, a large amount of simulation trials are required to achieve a specified precision for  $\hat{P}_L$ . The task becomes very computationally demanding. Now let us briefly review IS techniques.

**Review of importance sampling:** The objective of using IS techniques is to obtain a significant reduction in the number of simulation trials required to obtain the same estimator precision as the MC simulation. The application of IS techniques can provide solutions to problems involving rare events that would otherwise not be possible using either a pure analytical method or the MC simulation method. The key concept of IS is to bias the original transition probability matrix  $\mathbf{P}$  in such a way that the rare events occur more frequently. Using a likelihood ratio function and assuming that the chain starts in a certain state with probability one, the IS method yields a statistically unbiased estimator,

$$\tilde{P}_L = \frac{1}{N} \sum_{j=1}^N 1_{\mathbf{Y}}(\mathbf{Y}_j) W(\mathbf{Y}_j) \quad (3)$$

where

$$W(\mathbf{Y}_j) = \frac{\prod_{t=1}^L P(X_t^j | X_{t-1}^j)}{\prod_{t=1}^L Q(X_t^j | X_{t-1}^j)}, \quad (4)$$

$\mathbf{Y}_j$  is a sampled trajectory generated from a biased transition probability matrix  $\mathbf{Q}_t$  rather than  $\mathbf{P}$ , and  $W(\mathbf{Y}_j)$  is a compensatory factor, called the likelihood ratio. For independent and uncorrelated simulations, the variance of  $\tilde{P}_L$  is

$$\sigma_{\tilde{P}_L}^2 = \frac{E_Q[Z_j^2] - P_L^2}{N}, \quad (5)$$

where  $Z_j = 1_{\mathbf{Y}}(\mathbf{Y}_j) W(\mathbf{Y}_j)$ .

The optimal IS method requires  $E_Q[Z_j^2]$  (or equivalently the variance) to be minimized. In the past, the IS method for Markov systems was largely based on the large deviation technique (LDT). We need to compute the asymptotical optimal biased transition probability matrix and then simulate the performance. However, when  $S$  is large, it becomes computationally complex to compute the biased transition matrix alone. For example, if the code constraint length is 11, then, in order to compute the asymptotical optimal biased transition probability matrix using LDT, we need to invert a 1024 by 1024 matrix for each code we evaluate. This forces us to look for a better way to perform the code search. In the next section, we will show an optimal IS method and its properties.

### III. OPTIMAL IMPORTANCE SAMPLING FOR THE TRANSIENT MARKOV CHAIN

In this section, let us first go through an example. We then present several important properties and an optimal IS method for the transient Markov chain.

**Example 1:** Consider a special Markov chain given in Fig. 1. Suppose we need to estimate  $P_L = \sum_{t=1}^L f(X_t = 0 | X_0 = 1)$ , with  $L=4$ ,  $\pi_0(1) = 1$ .

Let  $q_t(i|j)$ s denote the element of  $i^{\text{th}}$  row and  $j^{\text{th}}$  column of  $\mathbf{Q}_t$  matrix. The task of obtaining OIS is reduced to finding optimal  $q_t(i|j)$ s, denoted as  $q_t^{(o)}(i|j)$ s, which minimize

$$\begin{aligned} E_Q[Z_j^2] &= \frac{(\frac{1}{2})^2 p^2}{q_1(2|1)q_2(0|2)} + \frac{(\frac{1}{2})^2 (\frac{1}{2})^2 p^2}{(1-q_1(2|1))q_2(2|3)q_3(0|2)} \\ &\quad + \frac{(\frac{1}{2})^2 (1-p)^2}{q_1(2|1)(1-q_2(1|2))} \frac{(\frac{1}{2})^2 p^2}{q_3(2|1)q_4(0|2)} \\ &\quad + \frac{(\frac{1}{2})^2}{(1-q_1(2|1))} \frac{(\frac{1}{2})^2}{(1-q_2(2|3))} \frac{(\frac{1}{2})^2 p^2}{q_3(2|3)q_4(0|2)} \end{aligned} \quad (6)$$

The values of  $q_t$ s in (6) need to be determined to minimize  $E_Q[Z_j^2]$ . Eqn.(6) seems very complicated, but the minimization procedure can start from  $q_4(0|2)$ . Since  $q_4(0|2)$  only appears in the last two terms in (6) and  $0 \leq q_4(0|2) \leq 1$ , we have  $q_4^{(o)}(0|2) = 1$ . Similarly we have  $q_3^{(o)}(2|1) = q_3^{(o)}(2|3) = 1$ . Thus, (6) is simplified as

$$\begin{aligned} \min\{E_Q[Z_j^2]\} &= \min\left\{ \frac{(\frac{1}{2})^2 p^2}{q_1(2|1)q_2(0|2)} + \frac{(\frac{1}{2})^2 (\frac{1}{2})^2 p^2}{(1-q_1(2|1))q_2(2|3)} \right. \\ &\quad \left. + \frac{(\frac{1}{2})^2 (1-p)^2 (\frac{1}{2})^2 p^2}{q_1(2|1)(1-q_2(0|2))} + \frac{(\frac{1}{2})^2 (\frac{1}{2})^2 (\frac{1}{2})^2 p^2}{(1-q_1(2|1))(1-q_2(2|3))} \right\}. \end{aligned} \quad (7)$$

Furthermore, we have

$$q_2^{(o)}(0|2) = \frac{2}{3-p}, \quad q_2^{(o)}(2|3) = \frac{2}{3}. \quad (8)$$

Finally, the minimization is reduced to

$$\min\{E_Q[Z_j^2]\} = \min\left\{ \frac{\frac{1}{16} p^2 (3-p)^2}{q_1(2|1)} + \frac{\frac{9}{64} p^2}{1-q_1(2|1)} \right\}. \quad (9)$$

Consequently, we have  $q_1^{(o)}(2|1) = \frac{6-2p}{9-2p}$  and

$$\min\{E_Q[Z_j^2]\} = \frac{1}{64} p^2 (9-2p)^2. \quad (10)$$

This example clearly shows that we can obtain  $\mathbf{Q}^{(o)}$  for the transient Markov system. Now we aim to prove such a method is true for transient Markov chains with any length and any number of states. Due to the page limitation, the proofs of these propositions will be omitted.

To conduct such a minimization procedure, we need the following lemma. The lemma describes the formula for solving the minimization problem in (7) and (9).

**Lemma 2:** For given nonnegative real  $a_1, a_2$ , we have

$$x_k^{(o)} = \arg_{x_k} \min_{0 \leq x_k \leq 1} \left\{ \frac{a_1}{x_k} + \frac{a_2}{1-x_k} \right\} = \frac{\sqrt{a_k}}{\sqrt{a_1} + \sqrt{a_2}}, \quad (11)$$

$$\min_{0 \leq x_k \leq 1} \left\{ \frac{a_1}{x_k} + \frac{a_2}{1-x_k} \right\} = \left( \sum_{j=1}^2 \sqrt{a_j} \right)^2 \quad (12)$$

for  $k = 1, 2$ . Now let us summarize the algorithm to compute  $E_Q[Z_j^2]$ .

**Algorithm 3:** For a given  $\mathbf{Q}_\tau$ ,  $\tau = 1, \dots, L$ ,  $E_Q[Z_j^2]$  can be computed by the following recursive procedure. (Suppose state 1 and state 0 are the initial and target states respectively.)

*Step 1* Set  $\tau = 0$ , a vector  $\Pi_\tau = [\Pi_\tau^{(0)}, \dots, \Pi_\tau^{(S-1)}] = [0, 1, 0, \dots, 0]$  and  $\text{SUM}=0$ .

*Step 2* Compute the transition matrix  $\Psi_\tau$  with its  $j^{\text{th}}$  row and  $i^{\text{th}}$  column equal to  $\frac{P^2(i|j)}{q_\tau(i|j)}$ .

*Step 3* Compute  $\Pi_\tau = \Pi_{\tau-1} \Psi_{\tau-1}$ .

*Step 4*  $\text{SUM}=\text{SUM}+\Pi_\tau$  (that is, the summation of  $\text{SUM}$  and the first element of  $\Pi_\tau$ ).

*Step 5* Set  $\Pi_\tau = \Pi_\tau + [-\Pi_\tau^{(0)}, 0, 0, \dots, 0]$ , (that is, setting the first element of  $\Pi_\tau$  to zero).

*Step 6* Repeat steps 2-5 until  $\tau = L$  and  $E_Q[Z_j^2] = \text{SUM}$ .

Algorithm 3 leads us to the following proposition.

**Proposition 4:** The optimal IS transition matrix  $\mathbf{Q}_\tau^{(o)}$  is only dependent on  $\mathbf{Q}_{\tau'}^{(o)}$ , for  $\tau < \tau' \leq L$  and not dependent on  $\mathbf{Q}_{\tau''}^{(o)}$ , for  $0 < \tau'' < \tau$ .

This result leads to the following method to systematically compute the optimal IS transition probability matrix  $\mathbf{Q}_\tau^{(o)}$ . For simplicity, we present the rest of the results for the case of two one-step transition branches  $J = 2$  and assume that the targeted state is state 0 and the starting state is state 1. The result can be easily generalized.

Recursively applying Algorithm 3 and Proposition 4, we can show the following algorithm for computing the optimal transition probability matrix.

**Algorithm 5:** The optimal IS transition probability matrix  $\mathbf{Q}_\tau^{(o)}$ , which minimizes  $E_Q[Z_j^2]$ , can be computed by the following procedures:

**Step 1** Set  $\tau = L$  and a  $S$ -element vector  $\Pi_\tau = [\Pi_\tau^{(0)}, \dots, \Pi_\tau^{(S-1)}]^t = [1, 0, \dots, 0]^t$  where superscript  $t$  denotes matrix transpose. Set the element at the  $k^{\text{th}}$  row and  $i^{\text{th}}$  column of transition weight matrix  $\Psi_\tau$  equal to  $\frac{P^2(k|i)}{q_\tau(k|i)}$ .

**Step 2** Compute

$$\Pi_{\tau-1} = \Psi_\tau \Pi_\tau + [1, 0, \dots, 0]^t. \quad (13)$$

Each element of  $\Pi_{\tau-1}$  is in one of the following two formats:  $\frac{a}{q_\tau(k|i)}$  or  $\frac{a}{q_\tau(k|i)} + \frac{b}{1-q_\tau(k|i)}$ . If it is the first format, then we have  $q_\tau^{(o)}(k|i) = 1$ . If it is the second format, then we can compute  $q_\tau^{(o)}(k|i)$  according to Lemma 2.

**Step 3** Substitute  $q_\tau^{(o)}(k|i)$ s into (13) and then update  $\Pi_{\tau-1}$ .

**Step 4** Set  $\tau = \tau - 1$  and repeat steps 2 and 3 until  $\tau = 1$  and  $E_Q[Z_j^2] = \Pi_0^{(1)}$ .

Algorithm 5 leads us to the following result.

**Proposition 6:** For a given transient Markov system with transition probability matrix  $\mathbf{P}$  and a length of time  $L$ , the value of  $\mathbf{Q}_\tau^{(o)}$ ,  $\tau = 1, \dots, L$  is unique.

Proposition 6 shows that there is only one optimal IS solution for the transient Markov system. Several interesting questions are then: does  $\mathbf{Q}_0^{(o)}$  converge when  $L \rightarrow \infty$ ? If so, where does  $\mathbf{Q}_0^{(o)}$  converge to? The following proposition provides the answers.

**Proposition 7:** The optimal solution of  $\mathbf{Q}_0^{(o)}$  converges to  $\mathbf{P}$  for  $L \rightarrow \infty$ , that is,

$$\lim_{L \rightarrow \infty} \mathbf{Q}_0^{(o)} = \mathbf{P} \quad (14)$$

$$\lim_{L \rightarrow \infty} \min E_Q[Z_j^2] = \lim_{L \rightarrow \infty} P_L = 1. \quad (15)$$

Proposition 7 is very interesting. It shows that, as  $L \rightarrow \infty$ , the best importance sampling method reduces to the standard Monte Carlo simulation. This has been confirmed by many previous studies (see [26]), largely based on simulation trials. We have now shown it analytically. This implies that, as the length of sample path becomes large, we have to let  $\mathbf{Q}_t^{(o)}$ s, for  $t = 0, 1, 2, \dots$ , initially converge to  $\mathbf{P}$ .

#### IV. NUMERICAL RESULTS

In this section, we will study the efficiency of optimal IS method. Consider the system given in Fig.1 with  $p = 0.0001$  and  $L = 9$ . In Fig.2 we present the OIS results. At each branch we label two probabilities: the top one corresponds to the original probability and the bottom one corresponds to  $q_\tau^{(o)}(i|j)$ . For example, at  $t = 3$ , the original transition probability from state 2 to state 0 is 0.0001 and the optimal biased transition probability is 0.359586. Similarly the original and biased probabilities from state 1 to state 2 are 0.5 and 0.5778 respectively and the probabilities from state 3 to 2 are 0.5 and 0.581 respectively. For this example we can compute the speed-up factor. Denote the speed-up factor by

$$S_f = \frac{P_L - P_L^2}{E_Q[Z_j^2] - P_L^2}. \quad (16)$$

For a normalized error of 1%, we can calculate the speed-up gain for this simple case. The true value of  $P_L$  is  $2.777 \times 10^{-4}$ . In Table I, we see that the optimal IS method only needs one simulation.

TABLE I  
SPEED UP FACTOR

Method	$E_Q[Z_j^2]$	$P_L$	No. of trials
OIS	$7.7 \times 10^{-8}$	$2.7 \times 10^{-4}$	$\cong 1$
MC	$2.7 \times 10^{-4}$	$2.7 \times 10^{-4}$	$3.6 \times 10^5$

In Fig.3 we show the speed-up gain and accuracy of the OIS method for a transient Markov chain with 1024 states and  $L = 3$ . As can see from Fig.3(a) when  $p$  decreases, the speed-up factor increases dramatically. This indicates that the complexity of the standard MC simulation method increases as  $p$

decreases, but the complexity of the sub-optimal method is almost unchanged. In Fig.3(b) we see that the difference between the estimated  $P_L$  values is less than 10% of the value computed by the analytic computation method.

We also applied the method to estimate  $P_L$  for a code with  $2^{18}$  states. Since we could not obtain an accurate estimate using the MC method, we cannot compute the speed-up factor for this case. However, comparing the computer simulation time, we see a huge difference. It took 41.22 sec CPU time on a PC computer(300MHZ CPU) to obtain a  $P_L$  value with the normalized standard deviation of 30%, but, we could not obtain a single hit after 4 days using the MC method.

## V. DISCUSSION AND CONCLUSIONS

In this paper we have demonstrated that for a simple transient Markov system we can derive an optimal IS method. We further studied several propositions. We proved analytically that the optimal IS method reduces to the Monte Carlo simulation method if the length of sample path tends to infinity. The numerical results have been presented to study the efficiency of the IS methods. This OIS has wide application, for example code search [23] and queuing networks [24] [25].

## REFERENCES

- [1] G. Orsak and B. Aazhang, "Techniques for Estimating the Bit Error Rate in the Simulation of Digital Communication Systems," *IEEE Journal on Selected Areas in Communications (JSAC)*, Vol. 2, pp. 153-170, Jan. 1984
- [2] G. Orsak and B. Aazhang, "On the Theory of Importance Sampling Applied to the Analysis of Detection Systems," *IEEE Trans. on Communications*, Vol. 37, pp. 332-339, Apr. 1989.
- [3] G. Orsak and B. Aazhang, "Efficient Importance Sampling Techniques for Simulation of Multiuser Communication Systems," *IEEE Trans. on Communications*, Vol. 40, pp. 332-339, Jun. 1992.
- [4] K. S. Shanmugam and P. Balaban, "A Modified Monte-Carlo Simulation Technique for the evaluation of Error Rate in Digital Communication Systems," *IEEE Trans. on Communications*, Vol. 28, pp. 1916-1924, Nov. 1980.
- [5] P. M. Hahn and M. C. Jeruchim, "Developments in the Theory and Application of Importance Sampling," *IEEE Trans. on Communications*, Vol. 35, pp. 706-714, Jul. 1987.
- [6] M. C. Jeruchim, P. M. Hahn, K. P. Smyntek, and R. T. Ray, "An Experimental Investigation of Conventional and Efficient Importance Sampling," *IEEE Trans. on Communications*, Vol. 37, pp. 578-587, Jun. 1989.
- [7] J. A. Bucklew, "Large Deviations and Rare Events in the study of Stochastic Algorithms," *IEEE Trans on Autom. Control*, Vol. 28, pp. 907-920, Sept. 1983.
- [8] J. A. Bucklew, P. Ney, and J. S. Sadowsky, "Monte Carlo Simulation and Large Deviations Theory for Uniformly Recurrent Markov Chains," *J. Appl. Prob.*, Vol. 27, pp. 44-59, 1990.
- [9] J. S. Sadowsky, "A New Method for Veberti Decoder Simulation Using Importance Sampling," *IEEE Trans. on Communications*, Vol. 38, pp. 1341-1351, Sept. 1990
- [10] J. S. Sadowsky, J. A. Bucklew, "On Large Deviations Theory and asymptotically efficient Monte Carlo estimation," *IEEE Trans. on Inform. Theory*, Vol. 36, pp. 579-588, Mar. 1990
- [11] J. S. Sadowsky, "Large Deviations Theory and Efficient Simulation of Excessive Backlogs in A GI/GI/m Queue," *IEEE Trans. on Autom. Control*, Vol. 36, pp. 1383-1394. Dec. 1991.
- [12] G. Orsak and B. Aazhang, "Constrained Solutions in Importance Sampling via Robust Statistics," *IEEE Trans. on Inform. Theory*, Vol. 37, pp. 307-316, Feb. 1991.
- [13] G. Orsak and B. Aazhang, "A Class of Optimum Importance Sampling Strategies," *Inform. Sci.*, Vol. 84, No. 1-2, pp. 139-160, 1995.
- [14] P. J. Smith, M. Shafi, H. Gao, "Quick Simulation: A review of Importance Sampling Techniques in Communications Systems," *IEEE Journal on Selected Areas in Communications (JSAC)*, Vol. 15, pp.597-613, May. 1997.
- [15] P. Heidelberger, "Fast Simulation of Rare Event in Queuing and reality Models," *ACM Trans on Modeling and Computer Simulation*, Vol. 5, No. 1, pp. 43-85 Jan. 1995
- [16] P.Shahabuddin, P. Glasserman, and P. Heidelberger, "Splitting for Rare Events Simulation: Analysis of Simple Cases," *Proceedings of Winter Simulation Conf.*, pp. 302-308, Coronado, CA, Dec. 1996.
- [17] L. Wei, "Efficiency of Optimal Conditional Importance Sampling," *Electronics Letters*, pp.1322-1323, 3rd Aug. 1995.
- [18] L. Wei, "Estimated Bit Error Probability of DS-SSMA/MPSK with Coherent Detector on Satellite Mobile Channel," *IEEE Journal on Selected Areas in Communications (JSAC)*, Vol. 13, pp. 250-263, Feb. 1995.
- [19] L. Wei, "Estimated Bit Error Probability of DS-SSMA/MDPSK with Differential Phase Detector on Satellite Mobile Channel," *IEE Proceeding I, (communications)*, Vol. 143, No.5, pp. 297-304, Oct. 1996.
- [20] L. Wei and C. Schlegel, "Synchronization requirements for Multi-user OFDM on Satellite Mobile and Two-Path Rayleigh Channels," *IEEE Trans. on Communications*, Vol. 43, pp. 887-895, Feb./Mar./Apr. 1995.
- [21] M. Devetsikiotis and J. K. Townsend, "Statistical Optimization of Dynamic Importance Sampling Parameters for Efficient Simulation of Communication Networks," *IEEE/ACM Trans. on Networking*, Vol. 1, No.3, pp. 293-305, Jun. 1993.
- [22] Y. Wei and L. Wei, "Simple Method for Designing Fast Recovery Convolutional Codes for M-algorithm Using Large Deviation Technique," *Proceeding of IEEE Globecom Mini-conference in Communications Theory*, Nov.8-12, 1998, Sydney, pp.274-278.
- [23] Y. Wei and L. Wei, "Construction of Fast Recovery Codes Using a New Optimal Importance Sampling Method," *IEEE IEEE Trans. on Inform. Theory*, accepted provided revision.
- [24] M. Y. Wei, *Fast Simulation Methods For Communication Systems/Networks*, Ph.D. Dissertation, Dept. of Engineering, FEIT, ANU, 1999.
- [25] Y. Wei, H. Qi and L. WEI, "Optimal Importance Sampling for Queuing Networks," *IEEE/ACM Trans. on Modelling and Computer Simulation*, submitted, 2001.
- [26] S. Andradottir, *On the choice of alternative measures in importance sampling with Markov chain*, Operations research, No. 3, May-Jun. 1995
- [27] S. Parekh and J. Walrand, "A Quick Simulation Method for Excessive Backlogs in Networks of Queues," *IEEE Trans on Autom. Control*, Vol. 34, pp. 54-66, Jan. 1989

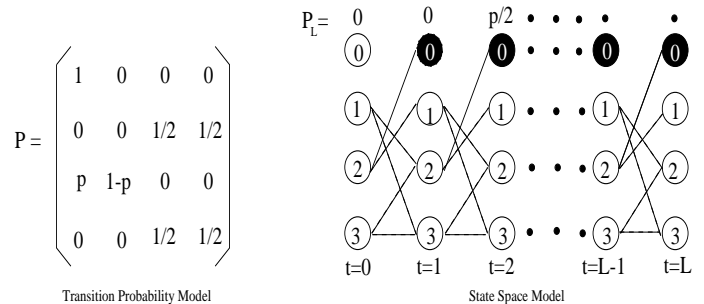


Fig. 1. A special Markov chain.

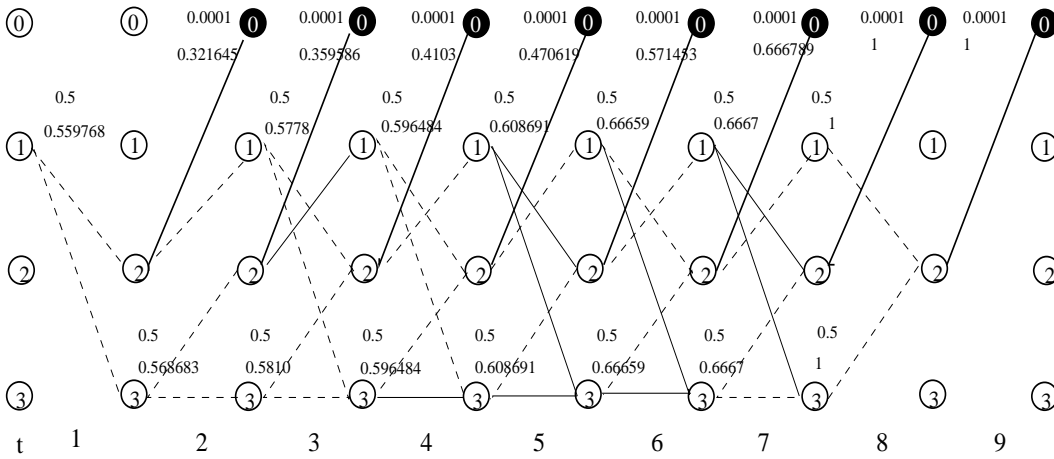
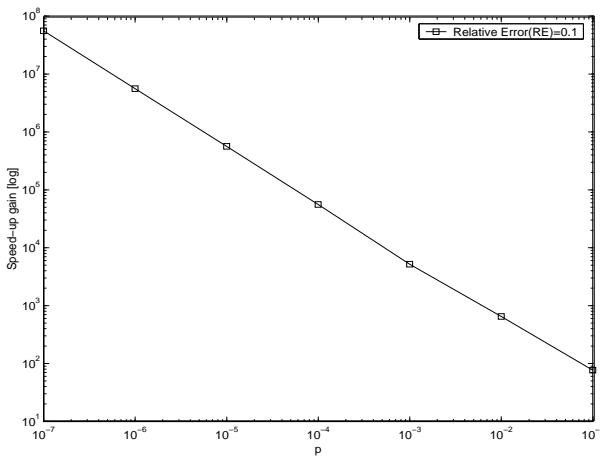


Fig. 2. A numerical example of optimal  $Q$  for  $L=9$



(a) Speed-up gain as a function of  $p$

$P_L$ $p$	by IS	by MC	by AC
$10^{-7}$	1.641576e-9	1.662916e-9	1.699883e-9
$10^{-6}$	1.640321e-8	1.682637e-8	1.699884e-8
$10^{-5}$	1.575751e-7	1.996990e-7	1.699893e-7
$10^{-4}$	1.575848e-6	1.682637e-6	1.699882e-6
$10^{-3}$	1.519360e-5	1.739400e-5	1.709141e-5
$10^{-2}$	1.640159e-4	1.561483e-4	1.709141e-4

(b) Accuracy

Fig. 3. Speed-up gain and accuracy for a Markov system with  $S = 1024$  states and  $L = 9$