The Korean Communications
in Statistics Vol. 7, No. 2, 2000
pp. 499-511

Fast Simulation for Excessive Backlogs in Tandem Networks[1)]

Jiyeon Lee[2)]

Abstract

We consider a stable tandem network which consists of two M/M/1 nodes and study the probability that the total backlog exceeds a large level N. Since the excessive backlog is a rare event, it is difficult to estimate this probability efficiently by using the crude Monte Carlo simulation. Instead we perform the h-transform proposed by McDonald(1999) to obtain the twisted network, in which the node with the larger load is overloaded. Then we use it to run the fast simulation.

Keywords : Tandem network, h-transform, Fast simulation, Overload probability

1. Introduction

We consider two M/M/1 nodes with infinite buffers in tandem with service rates μ_1 and μ_2, respectively. We assume, for stability of the network, that the arrival rate λ from outside satisfies $\lambda < \mu_1$ and $\lambda < \mu_2$. In this paper, we will refer to such a system by a (λ, μ_1, μ_2) -network. Fig. 1 depicts a (λ, μ_1, μ_2)-network.

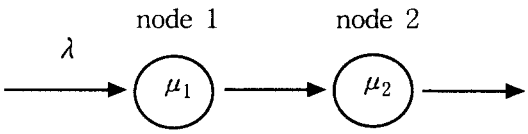

Fig. 1. A (λ, μ_1, μ_2)-network.

A (λ, μ_1, μ_2)-network can be described as a Markov jump process $M := \{M(t),\ t \geq 0\}$ on $S := \mathrm{N}^2$, where N denotes the non-negative integers. Let $\vec{x} = (x_1, x_2) \in S$ denote the

1) This work was supported by Korea Research Foundation Grant (KRF-99-103816).
2) Assistant Professor, Department of Statistics, Yeungnam University, Kyongsan 712-749. Korea, E-mail : leejy@yu.ac.kr

state of $M(t)$ where x_i represents the queue size (that is the total number of customers waiting or being served) at node i, $i=1$, 2. The generator G of M is given as an operator on a bounded function f on S;

$$\begin{aligned} Gf(\vec{x}) &= \lambda\{f(x_1+1,x_2)-f(x_1,x_2)\} \\ &\quad +\mu_1(x_1)\{f(x_1-1,x_2+1)-f(x_1,x_2)\} \\ &\quad +\mu_2(x_2)\{f(x_1,x_2-1)-f(x_1,x_2)\}, \end{aligned}$$

where $\mu_1(x_1)=\mu_1$ if $x_1>0$ and 0 otherwise, and $\mu_2(x_2)$ is defined analogously. The stationary distribution π of M is given by

$$\pi(\vec{x})=(1-\rho_1)\rho_1^{x_1}(1-\rho_2)\rho_2^{x_2} \tag{1}$$

under the condition that both loads $\rho_i=\lambda/\mu_i$, $i=1,2$ are smaller than 1. The equation (1) implies that, in the steady state, the queue sizes at the different nodes are independent. Furthermore, the queue size at node i has the stationary measure of a birth and death process with birth rate λ and death rate μ_i, $i=1,2$ (Walrand(1988)).

The event rate of the (λ, μ_1, μ_2)-network is $\lambda+\mu_1+\mu_2$. Without loss of generality we assume that $\lambda+\mu_1+\mu_2=1$ (otherwise, we can rescale time). If we regard G as the discrete generator of a Markov chain W on S, then the (λ, μ_1, μ_2)-network is precisely the uniformization of this chain. Consequently π is also the stationary distribution of W. We assume that the kernel K is associated with the Markov chain W with the stationary distribution π.

Let E denote the set of the realizations of W that reach the region $F:=\{\vec{x}=(x_1,x_2)\in S \mid x_1+x_2\geq N\}$ before hitting the state $\vec{0}=(0,0)$. We are interested in estimating, for large N, the probability $\alpha:=P_{\vec{0}}\{E\}$; the probability that W reaches F before returning to the state $\vec{0}$ given that W starts from the state $\vec{0}$. Here we can find α by the first step method. For this let $P(x_1,x_2)$ for $0\leq x_1$, $x_2\leq N$ denote the probability that W hits F before returning to the state $\vec{0}$ given that it starts from the state (x_1,x_2). Clearly, $P(0,0)=0$, $P(x_1,N-x_1)=1$ for $0\leq x_1\leq N$, and $P(1,0)=\alpha$. The first step equations give

$$\begin{aligned}
(\lambda+\mu_1)P(x_1,0) &= \lambda P(x_1+1,0)+\mu_1 P(x_1-1,1), \quad 1\le x_1\le N-1 \\
(\lambda+\mu_2)P(0,x_2) &= \lambda P(1,x_2)+\mu_2 P(0,x_2-1), \quad 1\le x_2\le N-1 \\
P(x_1,x_2) &= \lambda P(x_1+1,x_2)+\mu_1 P(x_1-1,x_2+1)+\mu_2 P(x_1,x_2-1), \\
&\qquad 1\le x_1\le N-1,\ 1\le x_2\le N-x_1-1.
\end{aligned} \tag{2}$$

However the above equations (2) may not be solved analytically because the order of the characteristic equation becomes large. Therefore the simulation is often recommended to estimate the probability α. For this stable network, the events of reaching the large total backlog are very infrequent. Hence direct Monte Carlo simulations are very slow and take up a lot of computing time. Besides, there is also the difficulty of implementing a pseudo-random generator that can function effectively during very long simulations (Schwartz and Weiss(1993), Heidelberger(1995)).

In this paper, we first perform the h-transform proposed by McDonald(1999) to find a harmonic function and its associated twisted network which represent the most likely overload behaviour of the original network. Those, then, make us to be able to obtain the estimator of the probability α for the fast simulation because the event corresponding to the excessive backlog for the twisted network has high probability. This twisted network is identical to the network which is derived by an exponential change of measure for importance-sampling estimator in Parekh and Walrand(1989) and Frater et al.(1991). They used the heuristic arguments associated with large deviation principles in Varadhan(1984) to find the change of measure while we present the analytic method which gives the twisted network. In addition, our method can be applied to the case that both nodes have even the same load.

In section 2, we introduce the h-transform method of McDonald(1999) for general queueing networks. We use it to find twisted networks for one node's overload for tandem network in subsection 2.1. In subsection 2.2, we extend the original tandem network by adding a fictitious node which records each arrival and departure of the network. Then we perform the h-transform to obtain the harmonic functions for the total backlog's overload. We notice that these harmonic functions and associated twisted networks are identical with those obtained in subsection 2.1. Section 3 shows how to use the twisted networks and the harmonic functions derived in section 2 in the fast simulation that estimates the probability α.

2. The h-transform method

We consider a stable queueing network with $r+m$ nodes, which may be represented by a

Markov jump process having a generator G in the positive orthant $S=\mathrm{N}^{r+m}$. We also assume, without loss of generality, that the event rate of this jump process is 1. If we regard G as the discrete generator of a Markov chain W on S, then the network is precisely the uniformization of this chain. We observe overload behaviour of other nodes when a chosen node becomes overloaded; that is when one coordinate of the chain W, on which we relabel the first coordinate, exceeds a given level. When this node is overloaded, other nodes may remain stable even though they are subject to higher loads. The coordinates corresponding to these super stable nodes are renumbered to $r+1$ through $r+m$. Unfortunately when the chosen node is overloaded it may drive other nodes into overload. We assume these nodes correspond to coordinates 2 through r. That is, for $\vec{x}=(x_1, x_2, \cdots, x_r, x_{r+1}, \cdots, x_{r+m})$, x_1 denotes the queue size of the chosen overloaded node; $x_2, \cdots, x_r$ are those of nodes which change to be unstable; $x_{r+1}, \cdots, x_{r+m}$ are those of nodes which remain stable. We look for a harmonic function $h(\vec{x})$ which transforms the original network into the twisted network. In the twisted network, nodes 1 through r become transient while others remain recurrent.

If $\beta=\{i_1, i_2, \cdots, i_d\}$, we say $\vec{x}$ is on the boundary S_β if $x_i=0$ for $i\in\beta$ but $x_i>0$ for $i\notin\beta$. Denote the interior of the orthant by int(S) if $x_i>0$ for all i. To find h we remove the boundary $\Delta :=\bigcup_{k=1}^{r} S_{\{k\}}$ for the first r coordinates. Define $S^\infty=\mathbf{Z}^r\times \mathrm{N}^m$, where $\mathbf{Z}$ denote the integers. If $\beta\subseteq\{r+1,\cdots, r+m\}$ and $x_i=0$ for $i\in\beta$ but $x_i>0$ for $i\subseteq\{r+1, \ldots, r+m\}\setminus\beta$, then $\vec{x}\in S_\beta^\infty$.

Define $\mathrm{int}(S^\infty) := \{\vec{x}\in S^\infty : x_i>0,\ i=r+1,\cdots, r+m\}$. We decompose $\vec{x}\in S$ as

$$\vec{x}=(\tilde{x}, \hat{x}) \text{ where } \tilde{x}\in \mathrm{N}^r,\ \hat{x}\in \mathrm{N}^m.$$

On S^∞ we assume that transitions for a chain W^∞ are given by a probability transition kernel K^∞ of the form

$$K^\infty(\vec{x}, \vec{y})= \hat{K}^\infty(\hat{x}, \hat{y})\, g(\tilde{y}-\tilde{x} \mid \hat{x}, \hat{y}),$$

where $\hat{K}^\infty(\hat{x}, \hat{y})$ is a transition kernel from N^m to N^m and $g(\cdot \mid \hat{x}, \hat{y})$ is a probability mass function given each pair $(\hat{x}, \hat{y})$. We also assume that the probability transition kernel $K(\vec{x}, \vec{y})$ of W agrees with $K^\infty(\vec{x}, \vec{y})$ when $\vec{x}\in S\setminus\Delta$. Consequently the chain W behaves like W^∞ outside the boundary Δ.

McDonald(1999) shows that in great generality one can construct the function $h(\vec{x}) := a_1^{x_1} a_2^{x_2}\cdots a_{r+m}^{x_{r+m}}$ for $\vec{x}\in S^\infty$ such that h is harmonic for the kernel K^∞, that is,

$$h(\vec{x}) = \sum_{y \in S^\infty} K^\infty(\vec{x}, \vec{y}) h(\vec{y}), \tag{3}$$

where $\boldsymbol{a} := (a_1, \cdots, a_{r+m})$ is a vector of positive constants such that $a_2 = \cdots = a_r = 1$. We assume the existence of h such that G^∞_{tw} defined below is a discrete generator. For a bounded function f on S^∞,

$$\begin{aligned} G^\infty_{\mathrm{tw}} f(\vec{x}) &:= \frac{1}{h(\vec{x})} G^\infty(h \cdot f)(\vec{x}) \\ &= \sum_{y \in S^\infty} [f(\vec{y}) - f(\vec{x})] \frac{h(\vec{y})}{h(\vec{x})} K^\infty(\vec{x}, \vec{y}), \end{aligned}$$

where G^∞ is the generator of the chain W^∞. We call G^∞_{tw} the generator of the twisted chain W^∞_{tw}. We also denote the kernel by K^∞_{tw}. Hence

$$K^\infty_{\mathrm{tw}}(\vec{x}, \vec{y}) = \frac{h(\vec{y})}{h(\vec{x})} K^\infty(\vec{x}, \vec{y}). \tag{4}$$

Of course, the solution h of the equation (3) must produce the twisted chain $W^\infty_{\mathrm{tw}} = (\widetilde{W}_{\mathrm{tw}}{}^\infty, \widehat{W}_{\mathrm{tw}}{}^\infty)$ such that $\widetilde{W}_{\mathrm{tw}}{}^\infty$ drifts to plus infinity while $\widehat{W}_{\mathrm{tw}}{}^\infty$ must be a stable Markov chain. If this fails, then we must try again by twisting another set of coordinates; that is, we must redefine the super stable nodes.

Define

$$q(\vec{x}) := \sum_{y \in S} K(\vec{x}, \vec{y}) h(\vec{y}) / h(\vec{x}) \text{ for } \vec{x} \in S \tag{5}$$

and define $K_{\mathrm{tw}}(\vec{x}, \vec{y}) = K(\vec{x}, \vec{y}) (h(\vec{y})/h(\vec{x})) q(\vec{x})^{-1}$ for $\vec{x}, \vec{y} \in S$. Then the twisted chain W_{tw} with the kernel K_{tw} has the same state space S as the original chain W does. Notice that the kernel K_{tw} agrees with the kernel K^∞_{tw} except on the boundary Δ.

2.1 The h-transform for one node's overload in tandem network

(i) For the case of $\rho_1 > \rho_2$

If the load of node 1 is larger than that of node 2, then node 1 becomes overloaded first. Take $\Delta = \{\vec{x} = (x_1, x_2) \in S; x_1 = 0\}$ and $S^\infty = \mathbf{Z} \times \mathrm{N}$. To calculate constants a_1 and a_2

for the harmonic function $h(\vec{x})=a_1^{x_1}a_2^{x_2}$, we derive one constraint in the interior, int (S^{∞}), from the equation (3);

$$\lambda a_1+\mu_1 a_1^{-1}a_2+\mu_2 a_2^{-1}=1.$$

The constraint on $S^{\infty}_{\{2\}}$ is

$$\lambda a_1+\mu_1 a_1^{-1}a_2=\lambda+\mu_1.$$

Subtracting the later constraint from the first yields $\mu_2 a_2^{-1}=\mu_2$. Consequently $a_2=1$. Substituting it into the first constraint gives $a_1=\mu_1/\lambda=1/\rho_1$. Therefore the harmonic function is $h(\vec{x})=\rho_1^{-x_1}$. The twisted kernel K^{∞}_{tw}, based on the equation (4), is given by

$$\begin{aligned}
K^{\infty}_{\text{tw}}((x_1,x_2),(x_1+1,x_2)) &= \frac{h(x_1+1,x_2)}{h(x_1,x_2)}K^{\infty}((x_1,x_2),(x_1+1,x_2))\\
&=\frac{\mu_1}{\lambda}\cdot\lambda=\mu_1\\
K^{\infty}_{\text{tw}}((x_1,x_2),(x_1-1,x_2+1)) &= \frac{h(x_1-1,x_2+1)}{h(x_1,x_2)}K^{\infty}((x_1,x_2),(x_1-1,x_2+1))\\
&=\frac{\lambda}{\mu_1}\cdot\mu_1=\lambda\\
K^{\infty}_{\text{tw}}((x_1,x_2),(x_1,x_2-1)) &= \frac{h(x_1,x_2-1)}{h(x_1,x_2)}K^{\infty}((x_1,x_2),(x_1,x_2-1))\\
&=1\cdot\mu_2=\mu_2
\end{aligned}$$

for $(x_1,x_2)\in \text{int}(S^{\infty})$. On $S^{\infty}_{\{2\}}$ we have

$$\begin{aligned}
K^{\infty}_{\text{tw}}((x_1,0),(x_1+1,0)) &= \frac{\mu_1}{\lambda+\mu_1}\\
K^{\infty}_{\text{tw}}((x_1,0),(x_1-1,1)) &= \frac{\lambda}{\lambda+\mu_2}.
\end{aligned}$$

On the other hand, it follows from the equation (5) that

$$q(\vec{x})=\begin{cases} 1 & \text{for } \vec{x}\in S\setminus\Delta\\ \dfrac{\mu_1+\mu_2}{\lambda+\mu_2} & \text{for } \vec{x}\in\Delta\setminus\{\vec{0}\}\\ \dfrac{\mu_1}{\lambda} & \text{for } \vec{x}=\vec{0}.\end{cases}$$

Thus the twisted kernel K_{tw} is given by, for $\vec{x}\in S\setminus\Delta$

$$K_{tw}(\vec{x},\vec{y}) = K^{\infty}_{tw}(\vec{x},\vec{y}),$$

and for $x_2 > 0$

$$K_{tw}((0,x_2),(1,x_2)) = \frac{\mu_1}{\mu_1+\mu_2}$$
$$K_{tw}((0,x_2),(0,x_2-1)) = \frac{\mu_2}{\mu_1+\mu_2},$$

and $K_{tw}((0,0),(1,0))=1$.

Hence the twisted chain W_{tw} with the kernel K_{tw} corresponds to the (μ_1, λ, μ_2)-network, which is obtained by interchanging the arrival rate λ and the service rate μ_1 of node 1 for the original (λ, μ_1, μ_2)-network. Notice that node 1 of the (μ_1, λ, μ_2)-network has the load $\mu_1/\lambda = 1/\rho_1$, larger than 1, which implies that this node is overloaded. On the other hand, node 2 remains stable since its load μ_1/μ_2 is smaller than 1.

(ii) For the case of $\rho_1 < \rho_2$

In this case, we consider $\Delta = \{\vec{x}=(x_1,x_2)\in S; x_2=0\}$ and $S^{\infty} = \mathbf{N}\times\mathbf{Z}$. As described earlier the constraint in the interior, int(S^{∞}), for the harmonic function $h(\vec{x}) = a_1^{x_1} a_2^{x_2}$ is

$$\lambda a_1 + \mu_1 a_1^{-1} a_2 + \mu_2 a_2^{-1} = 1$$

and the constraint on $S^{\infty}_{\{1\}}$ is

$$\lambda a_1 + \mu_2 a_2^{-1} = \lambda + \mu_2.$$

The solutions of the above equations are $a_1 = a_2 = \mu_2/\lambda = 1/\rho_2$ and the harmonic function is $h(\vec{x}) = \rho_2^{-(x_1+x_2)}$. Therefore the kernel K_{tw} is given by, for $(x_1,x_2)\in$ int(S)

$$K_{tw}((x_1,x_2),(x_1+1,x_2)) = \mu_2$$
$$K_{tw}((x_1,x_2),(x_1-1,x_2+1)) = \mu_1$$
$$K_{tw}((x_1,x_2),(x_1,x_2-1)) = \lambda,$$

on $S_{\{1\}}$

$$K_{tw}((0,x_2),(1,x_2)) = \frac{\mu_2}{\lambda+\mu_2}$$
$$K_{tw}((0,x_2),(0,x_2-1)) = \frac{\lambda}{\lambda+\mu_2},$$

on $S_{\{2\}}$

$$K_{tw}((x_1,0),(x_1+1,0)) = \frac{\mu_2}{\mu_1+\mu_2}$$

$$K_{tw}((x_1,0),(x_1-1,1)) = \frac{\mu_1}{\mu_1+\mu_2},$$

and $K_{tw}((0,0),(1,0))=1$. We notice that the (μ_2, μ_1, λ)-network corresponding to this twisted chain W_{tw} has overloaded node 2 and stable node 1.

(iii) For the case of $\rho_1=\rho_2$

If $\rho_1=\rho_2$, then both nodes become overloaded simultaneously. So take $\triangle=\{\vec{x}=(x_1,x_2)\in S; x_1=0 \text{ or } x_2=0\}$ and $S^{\infty}=\mathbf{Z}\times\mathbf{Z}$. Then we can derive one constraint for the harmonic function $h(\vec{x})=a^{x_1+x_2}$ in the interior, int(S^{∞}), such that

$$\lambda a+\mu_1+\mu_2 a^{-1}=1,$$

which implies that $h(\vec{x})=\rho_1^{-(x_1+x_2)}=\rho_2^{-(x_1+x_2)}$. Recall from the above case (ii) that the (μ_2, μ_1, λ)-network is obtained as the twisted network by using the harmonic function $h(\vec{x})=\rho_2^{-(x_1+x_2)}$. But, in this case, both nodes are overloaded because both loads are larger than or equal to 1.

2.2 The h-transform for the total backlog's overload in tandem network

We consider the behaviour of the queues at both nodes when the total backlog becomes overloaded. To make the total backlog overloaded we add a fictitious node to the network which records each arrival and departure of the network. Let ${}^{+}\vec{x}:=(x_1+x_2,\vec{x})$ and ${}^{+}\vec{y}:=(y_1+y_2,\vec{y})$ where $\vec{x}=(x_1,x_2)$, $\vec{y}=(y_1,y_2)\in S$. Then ${}^{+}W$ with the state vector ${}^{+}\vec{x}$ on ${}^{+}\vec{S}$ denotes an augmented chain where the first coordinate of ${}^{+}W$ tracks the number of customers at both nodes. Then the kernel of ${}^{+}W^{\infty}$ is given by

$${}^{+}K^{\infty}({}^{+}\vec{x}, {}^{+}\vec{y}) = \chi\{x_1+x_2={}^{+}x_1, y_1+y_2={}^{+}y_1\}K^{\infty}(\vec{x},\vec{y}),$$

where $\chi\{A\}$ denotes the indicator of an event A.

(i) For the case of $\rho_1>\rho_2$

The twist constants are defined as

$$\boldsymbol{a}:=(\rho_1^{-1},1,\rho_1).$$

Then the harmonic function $h({}^{+}\vec{x})$ is given by

$$h({}^{+}\vec{x}) = \rho_1^{-(x_1+x_2)} 1^{x_1} \rho_1^{x_2}$$
$$= \rho_1^{-x_1}.$$

So the (μ_1, λ, μ_2)-network is obtained as the twisted network in subsection 2.1.

Notice that the twisted kernel ${}^{+}K_{\mathrm{tw}}^{\infty}$ of ${}^{+}W_{\mathrm{tw}}^{\infty}$ is given by

$$\begin{aligned}{}^{+}K_{\mathrm{tw}}^{\infty}({}^{+}\vec{x}, {}^{+}\vec{y}) &= \chi\{x_1+x_2={}^{+}x_1, y_1+y_2={}^{+}y_1\} K^{\infty}(\vec{x}, \vec{y}) \frac{h({}^{+}\vec{y})}{h({}^{+}\vec{x})} \\ &= \chi\{x_1+x_2={}^{+}x_1, y_1+y_2={}^{+}y_1\} K^{\infty}(\vec{x}, \vec{y})\, \rho_1^{x_1-y_1}.\end{aligned}$$

By construction ${}^{+}K_{\mathrm{tw}}^{\infty}$ is the kernel on ${}^{+}S^{\infty}$ obtained by deleting the boundaries for the first two coordinates of ${}^{+}W$ which are transient. Hence when the total backlog is overloaded, node 1 is also overloaded but node 2 with the smaller load remains stable in the (μ_1, λ, μ_2)-network.

(ii) For the case of $\rho_1 < \rho_2$

The twist constants are defined as

$$\boldsymbol{a} := (\rho_2^{-1}, 1, 1).$$

Then the harmonic function is $h({}^{+}\vec{x}) = \rho_2^{-(x_1+x_2)}$ and the twisted network is the (μ_2, μ_1, λ)-network. In this twisted network, when the total backlog becomes overloaded node 1 remains stable but node 2 is subject to overload.

(iii) For the case of $\rho_1 = \rho_2$

In this case, if the total backlog becomes overloaded, then both nodes may be overloaded. So letting the twisted constants

$$\boldsymbol{a} := (\rho_1 = \rho_2, 1, 1)$$

we have the harmonic function $h({}^{+}\vec{x}) = \rho_1^{-(x_1+x_2)} = \rho_2^{-(x_1+x_2)}$, which is the same as one for the case $\rho_1 < \rho_2$. Of course, the same twisted network is obtained.

3. Fast simulation for the excessive backlog

We define a cycle as the duration starting from the state $\vec{0}$ and ending at the instant the chain W, for the first time, returns to the state $\vec{0}$. Let Ω be the sample space for the cycle trajectories. The elements of Ω can take the one of two forms. If W never reaches the region F during the cycle given by $\omega \in \Omega$, then ω is of the form $\{\vec{\omega}_0, \vec{\omega}_1, \cdots, \vec{\omega}_q\}$ where $\vec{\omega}_0 = \vec{0}$, $\vec{\omega}_1 = (1,0)$, $\vec{\omega}_k = (\omega_{1k}, \omega_{2k})$ for $k \geq 1$, and q is the time to return to the state $\vec{0}$. If W reaches the region F during the cycle, then ω is of the form $\{\vec{\omega}_0, \vec{\omega}_1, \cdots, \vec{\omega}_p, \vec{\omega}_{p+1}, \cdots, \vec{\omega}_q\}$ where p is the smallest integer such that $\omega_{1p} + \omega_{2p} \geq N$. Note that the former form is simply a special case of the latter form, so we will only consider cycles of the latter.

Let us define

$$V_k := \chi\{ W \text{ reaches } F \text{ in the } k\text{th cycle}\}.$$

Notice that V_k's are i. i. d. random variables having $\alpha = E[V_k]$. For the direct Monte Carlo simulation,

$$\alpha_n := \frac{V_1 + V_2 + \cdots + V_n}{n}$$

is an unbiased and convergent estimator of α for n simulated cycles.

To obtain another estimator of α for the fast simulation, we consider two cycles, namely an original cycle and a twisted cycle, both starting from the state $\vec{0}$ and moving to the state (1,0) with probability 1. The original (twisted) cycle uses the original (twisted) transition kernel K (K_{tw}) from the state (1,0) to the region F, and then returns to the state $\vec{0}$ by using the original transition kernel K. Then the original cycle generated by the original kernel K has the probability function

$$\nu(\omega) := \pi(\vec{\omega}_0) \prod_{k=1}^{p} K(\vec{\omega}_{k-1}, \vec{\omega}_k) \prod_{k=p+1}^{q} K(\vec{\omega}_{k-1}, \vec{\omega}_k),$$

where $K(\vec{\omega}_0, \vec{\omega}_1) = 1$. If the cycle is generated as the twisted cycle, then similarly it has the probability function

$$\nu_{\text{tw}}(\omega) := \pi(\vec{\omega}_0) \prod_{k=1}^{p} K_{\text{tw}}(\vec{\omega}_{k-1}, \vec{\omega}_k) \prod_{k=p+1}^{q} K(\vec{\omega}_{k-1}, \vec{\omega}_k).$$

Notice that $K^{\infty} = K$ and $K^{\infty}_{\text{tw}} = K_{\text{tw}}$ on $S \setminus \Delta$. If $\vec{\omega}_k \in \Delta \setminus \{\vec{0}\}$ and $\vec{\omega}_{k+1} \in S$, then we have that

$$
\begin{aligned}
K^{\infty}(\vec{\omega}_k, \vec{\omega}_{k+1}) &= A\,K(\vec{\omega}_k, \vec{\omega}_{k+1}) \\
K^{\infty}_{\mathrm{tw}}(\vec{\omega}_k, \vec{\omega}_{k+1}) &= B\,K_{\mathrm{tw}}(\vec{\omega}_k, \vec{\omega}_{k+1}),
\end{aligned}
\tag{6}
$$

where

$$
A=\begin{cases} \lambda+\mu_2 & \text{for } \rho_1\geq\rho_2 \text{ and } \omega_{1k}=0 \\ \lambda+\mu_1 & \text{for } \rho_1\leq\rho_2 \text{ and } \omega_{2k}=0 \end{cases}
$$

and

$$
B=\begin{cases} \mu_1+\mu_2 & \text{for } \rho_1\neq\rho_2 \text{ or } \omega_{2k}=0 \\ \lambda+\mu_2 & \text{for } \rho_1=\rho_2 \text{ and } \omega_{1k}=0. \end{cases}
$$

Let $N_\triangle(\omega)$ be the number of $k\in\{1,2,\cdots,p\}$ such that $\vec{\omega}_k\in\triangle$. From (6) and $K^{\infty}_{\mathrm{tw}}(\vec{\omega}_k, \vec{\omega}_{k+1})=K^{\infty}(\vec{\omega}_k, \vec{\omega}_{k+1})h(\vec{\omega}_{k+1})/h(\vec{\omega}_k)$, the probability function $\nu(\omega)$ can be rewritten as

$$
\begin{aligned}
\nu(\omega) &= \left(\frac{B}{A}\right)^{N_\triangle(\omega)} \pi(\vec{\omega}_0) \prod_{k=1}^{p} K_{\mathrm{tw}}(\vec{\omega}_{k-1}, \vec{\omega}_k)\frac{h(\vec{\omega}_{k-1})}{h(\vec{\omega}_k)} \prod_{k=p+1}^{q} K(\vec{\omega}_{k-1}, \vec{\omega}_k) \\
&= \left(\frac{B}{A}\right)^{N_\triangle(\omega)} \frac{h(\vec{\omega}_1)}{h(\vec{\omega}_p)} \pi(\vec{\omega}_0) \prod_{k=1}^{p} K_{\mathrm{tw}}(\vec{\omega}_{k-1}, \vec{\omega}_k) \prod_{k=p+1}^{q} K(\vec{\omega}_{k-1}, \vec{\omega}_k) \\
&= L(\omega)\,\nu_{\mathrm{tw}}(\omega),
\end{aligned}
\tag{7}
$$

where

$$
\begin{aligned}
L(\omega) &= \left(\frac{B}{A}\right)^{N_\triangle(\omega)} \frac{h(\vec{\omega}_1)}{h(\vec{\omega}_p)} \\
&= \begin{cases} \left(\dfrac{\mu_1+\mu_2}{\lambda+\mu_2}\right)^{N_{\triangle_1}(\omega)} \rho_1^{\omega_{1p}-1} & \text{for } \rho_1>\rho_2 \\ \left(\dfrac{\mu_1+\mu_2}{\lambda+\mu_1}\right)^{N_{\triangle_2}(\omega)} \rho_2^{N-1} & \text{for } \rho_1\leq\rho_2, \end{cases}
\end{aligned}
$$

where $\triangle_i=\{\vec{x}\in S \mid x_i=0\}$, $i=1,2$.

Denoting $V(\omega):=\chi\{W$ reaches F in the cycle $\omega\}$ and using (7), we get

$$
\begin{aligned}
\alpha &= \sum_{\omega\in\Omega} V(\omega)\,\nu(\omega) \\
&= \sum_{\omega\in\Omega} V(\omega)\,L(\omega)\,\nu_{\mathrm{tw}}(\omega).
\end{aligned}
$$

Let

$$
\begin{aligned}
U_k &:= \chi\{W_{\mathrm{tw}} \text{ reaches } F \text{ in the } k\text{th cycle}\} \\
&= \begin{cases} \chi\{(\mu_1, \lambda, \mu_2)-\text{network reaches } F \text{ in the } k\text{th cycle}\} & \text{for } \rho_1>\rho_2 \\ \chi\{(\mu_2, \mu_1, \lambda)-\text{network reaches } F \text{ in the } k\text{th cycle}\} & \text{for } \rho_1\leq\rho_2 \end{cases}
\end{aligned}
$$

and

$$\mathcal{L}_k := \begin{cases} \left(\dfrac{\mu_1+\mu_2}{\lambda+\mu_2}\right)^{N^k_{\triangle_1}} \rho_1^{N_k-1} & \text{for } \rho_1 > \rho_2 \\ \left(\dfrac{\mu_1+\mu_2}{\lambda+\mu_1}\right)^{N^k_{\triangle_2}} \rho_2^{N-1} & \text{for } \rho_1 \leq \rho_2, \end{cases}$$

where N_k denotes the queue size of node 1 when $\mho_k=1$, $N^k_{\triangle_1}$ denotes the number of visits to the boundary $\triangle_1$ during the kth cycle for (μ_1, λ, μ_2)-network, and $N^k_{\triangle_2}$ denotes the number of visits to the boundary $\triangle_2$ during the kth cycle for (μ_2, μ_1, λ)-network, respectively. Then for n simulated twisted cycles

$$\alpha_n^* := \frac{\mho_1\mathcal{L}_1 + \mho_2\mathcal{L}_2 + \cdots + \mho_n\mathcal{L}_n}{n}$$

is an unbiased and consistent estimator of the probability α for the fast simulation.

Table 1 shows the results of some experiments. In Table 2, there are the empirical means, standard deviations and coefficients of variation of the estimates obtained by the twisted network for the same examples as in Table 1. The first step equations were solved by using the IMSL routine, LEQT2F(Parekh and Walrand(1989)). Notice that the convergence under the fast simulation seems to be more rapid than that under the direct simulation.

Table 1. Simulations for tandem networks

Method	Direct Simulation		Fast Simulation	
Example I (0.05, 0.1, 0.85)-network N=15 $\alpha=3.459\times10^{-5}$				
n	10000	20000	100	500
$\alpha_n(\alpha_n^*)$	0.0	0.0	3.7122×10^{-5}	3.418×10^{-5}
Example II (0.1, 0.5, 0.4)-network N=13 $\alpha=2.104\times10^{-7}$				
n	20000	30000	700	1000
$\alpha_n(\alpha_n^*)$	0.0	0.0	2.3998×10^{-7}	1.996×10^{-7}
Example III (0.2, 0.4, 0.4)-network N=20 $\alpha=1.812\times10^{-5}$				
n	10000	20000	100	500
$\alpha_n(\alpha_n^*)$	0.0	0.0	1.464×10^{-5}	1.781×10^{-5}

Table 2. Empirical standard deviations for tandem networks

Example I (0.05, 0.1, 0.85)-network N=15 $a=3.459\times10^{-5}$ # of experiments=20		
n	100	500
Empirical Mean	3.818×10^{-5}	3.487×10^{-5}
Empirical Std. Dev.	5.428×10^{-7}	2.081×10^{-7}
C.V. ×100%	1.383	0.5968
Example II (0.1, 0.5, 0.4)-network N=13 $a=2.104\times10^{-7}$ # of experiments=20		
n	700	1000
Empirical Mean	2.431×10^{-7}	2.007×10^{-7}
Empirical Std. Dev.	3.423×10^{-8}	1.146×10^{-8}
C.V. ×100%	14.081	5.710
Example III (0.2, 0.4, 0.4)-network N=20 $a=1.812\times10^{-5}$ # of experiments=20		
n	100	500
Empirical Mean	1.574×10^{-5}	1.769×10^{-5}
Empirical Std. Dev.	2.645×10^{-6}	6.998×10^{-7}
C.V. ×100%	16.804	3.956

References

[1] Frater, M. R., Lennon, T. M. and Anderson, B. D. O. (1991). Optimally Efficient Estimation of the Statistics of Rare Events in Queueing Networks. *IEEE Transactions on Automatic Control*, Vol. 36, 1395-1405.

[2] Heidelberger, P.(1995). Fast Simulation of Rare Events in Queueing and Reliability Models. *ACM Transactions on Modeling and Computer Simulation*, Vol. 5, 43-85.

[3] McDonald, D. (1999). Asymptotics of First Passage Times for Random Walk in an Orthant. *Annals of Applied Probability*, Vol. 9, 110-145.

[4] Parekh, S. and Walrand, J. (1989). A Quick Simulation Method for Excessive Backlogs in Networks of Queues. *IEEE Transactions on Automatic Control*, Vol. 34, 54-66.

[5] Schwartz, A. and Weiss, A. (1993). Induced Rare Events: Analysis via Large Deviations and Time Reversal. *Advances in Applied Probability*, Vol. 25, 667-689.

[6] Varadhan, S. R. S. (1984). *Large Deviations and Applications*, SIAM.

[7] Walrand, J. (1988). *An Introduction to Queueing Network*, Prentice Hall International, Inc.