

MARKOV CHAIN SMALL-WORLD MODEL WITH ASYMMETRIC TRANSITION PROBABILITIES*

JIANHONG XU[†]

Abstract. In this paper, a Markov chain small-world model of D.J. Higham is broadened by incorporating asymmetric transition probabilities. Asymptotic results regarding the transient behavior of the extended model, as measured by its maximum mean first passage time, are established under the assumption that the size of the Markov chain is large. These results are consistent with the outcomes as obtained numerically from the model.

The focus of this study is the effect of a varying degree of asymmetry on the transient behavior which the extended model exhibits. Being a quite interesting consequence, it turns out that such behavior is largely influenced by the strength of asymmetry. This discovery may find applications in real-world networks where unbalanced interaction is present.

Key words. Asymmetry, Small-world, Ring network, Markov chain, Mean first passage time.

AMS subject classifications. 05C50, 15A51, 60J20, 60J27, 65C40.

1. Introduction. First we briefly describe the Markov chain small-world model of a ring network originally proposed by Higham [9]. For background material on the small-world phenomenon, we refer the reader to [13, 14, 15].

Consider a ring network consisting of $N + 1$ vertices, which are labeled clockwise as $0, \dots, N$. Initially, each vertex is connected to its neighboring vertices¹ by directed edges. A one-dimensional symmetric periodic random walk is then introduced on such a ring network. Specifically, starting from any vertex, the process moves to any of the neighboring vertices with equal probability $1/2$ over one time unit, thus leading to a homogeneous ergodic Markov chain with state space $S = \{0, \dots, N\}$ and transition

*Received by the editors October 6, 2008. Accepted for publication November 11, 2008. Handling Editor: Michael Neumann.

[†]Department of Mathematics, Southern Illinois University Carbondale, Carbondale, Illinois 62901, USA (jxu@math.siu.edu).

¹For any specified vertex, this term refers to the two nearest neighboring vertices on the ring network.

matrix

$$P_0 = \begin{bmatrix} 0 & p & & & p \\ p & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & p \\ p & & & p & 0 \end{bmatrix} \in \mathbb{R}^{(N+1) \times (N+1)},$$

where $p = 1/2$.

To model the small-world phenomenon on the ring network, the above Markov chain is modified by also allowing the process to jump to non-neighboring vertices with some small probability ϵ . For such a modified Markov chain, the transition matrix becomes

$$P_\epsilon = \begin{bmatrix} \epsilon & \tilde{p} & \epsilon & \cdots & \tilde{p} \\ \tilde{p} & \ddots & \ddots & \ddots & \\ \epsilon & \ddots & \ddots & \ddots & \\ \vdots & \ddots & \ddots & \ddots & \tilde{p} \\ \tilde{p} & & & \tilde{p} & \epsilon \end{bmatrix} \in \mathbb{R}^{(N+1) \times (N+1)},$$

where $0 < \epsilon \leq \frac{1}{N-1}$ and $\tilde{p} = \frac{1}{2} [1 - (N-1)\epsilon]$. We comment that the idea of adding long-range random jumps to non-neighboring vertices is in line with that of the Newman-Moore-Watts model [14], where such jumps are called shortcuts. The small-world phenomenon is known, see [14], to emerge as a result of adding a small amount of random shortcuts to the ring network, even though this framework seems simplistic as compared with real-world networks.

It is shown in [5, 9] that the Markov chain model as described above can be employed to capture the small-world phenomenon on the ring network, with results well conforming to those in [14] and in the well-known work by Watts and Strogatz [15]. In addition, the Markov chain approach appears to have an advantage in that it offers a more rigorous analysis of the ring network by incorporating tools in the fields of Markov chain theory, matrix theory, and differential equations.

In order to characterize the small-world property of the ring network in the Markov chain setting, the notion of mean first passage times, see [11], is adopted in [5, 9]. Extensive numerical evidence suggests that the maximum and average mean first passage times of the Markov chain model resemble the average, or characteristic, path-length as in [14, 15]. These quantitative Markov chain measures, therefore, provide us with an analytical, instead of experimental, way of investigating the small-world phenomenon arising on the ring network. We mention that the notion of mean

first passage times is also utilized in [10] to deal with the greedy path-length problem on the ring network.

The existing results in [5, 9], however, concern only the situation with symmetric transition probabilities, yet real-world networks, in general, involve asymmetric processes or interaction, an element distinct from simple topological structures. Examples in this regard include metabolic networks [1], social and large infrastructures [3], food webs [12], and communication networks [2]. Especially, taking communication networks as one instance, asymmetry occurs when the bandwidth, medium access control overhead, or loss rate is different in one direction than in the other.

In view of asymmetric interaction, it is natural for us to recast the ring network as a weighted digraph, with each weight representing the strength of the respective edge, namely, in terms of a communication network, the capacity, cost, or reliability of that link. In this work, we examine such a ring network based on an adapted Markov chain model stemming from Higham's. In particular, similar to [5, 9], we are primarily interested in the transient behavior of the Markov chain, as measured by its maximum mean first passage time. It turns out that such behavior is significantly influenced by the intensity of asymmetry, a quite unique feature which may find applications in real-world networks, including applications relative to the small-world phenomenon. The outcomes from this study also serve as a follow-up to [5] by extending the results there.

This paper is organized as follows: First, Section 2 contains the adaptation of the current Markov chain model so as to accommodate asymmetric interaction. Some necessary preliminary conclusions are stated in Section 3. Next, asymptotic results on the maximum mean first passage time are developed in Section 4. Section 5 provides further results in the context of the transient behavior of the Markov chain, together with examples, numerical results, and discussions. Finally, a few concluding remarks are presented in Section 6.

2. Markov chain model with asymmetry. Again, consider a ring network with $N + 1$ vertices, labeled clockwise as $0, \dots, N$. Suppose that at first, each vertex is connected to its neighboring vertices by directed and weighted edges. For any $j = 0, \dots, N$, the weights associated with edges $\{j, j + 1\}$ and $\{j, j - 1\}$ ² are assumed to be some fixed w_1 and w_2 , respectively. Let $0 < w_1 < w_2$. Such a setting leads to a one-dimensional asymmetric periodic random walk on the ring network. Specifically, starting from any vertex, the process moves over one time unit either to the clockwise neighboring vertex with probability p or to the counterclockwise neighboring vertex with probability $q = 1 - p$, where p and q are assumed to be proportional to w_1 and w_2 , respectively. Hence a homogeneous ergodic Markov chain is defined, with state

²The indices for edges are interpreted with mod $N + 1$ equivalence.

space $S = \{0, \dots, N\}$ and transition matrix

$$T_0 = \begin{bmatrix} 0 & p & & & q \\ q & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & p \\ p & & & q & 0 \end{bmatrix} \in \mathbb{R}^{(N+1) \times (N+1)} \quad (2.1)$$

such that $0 < p < q < 1$ and $p + q = 1$. The ratio of asymmetry is defined to be $r = q/p$. Throughout this paper, without loss of generality, we assume that $r > 1$. Note that in terms of this ratio r , $p = \frac{1}{r+1}$ and $q = \frac{r}{r+1}$.

Next, the Markov chain as above is modified by introducing long-range jumps to non-neighboring vertices according to a small probability ϵ . The transition matrix of this modified Markov chain can be expressed as

$$T_\epsilon = \begin{bmatrix} \epsilon & \tilde{p} & \epsilon & \cdots & \tilde{q} \\ \tilde{q} & \ddots & \ddots & \ddots & \\ \epsilon & \ddots & \ddots & \ddots & \\ \vdots & \ddots & \ddots & \ddots & \tilde{p} \\ \tilde{p} & & & \tilde{q} & \epsilon \end{bmatrix} \in \mathbb{R}^{(N+1) \times (N+1)}, \quad (2.2)$$

where $\tilde{p} = p - a$ and $\tilde{q} = q - b$. The stochasticity of T_ϵ implies that $a + b = (N - 1)\epsilon$. If we assume that $a/b = p/q$, namely the changes in p and q are proportional to p and q , respectively, then we obtain that for $0 < \epsilon \leq \frac{1}{N-1}$,

$$\tilde{p} = \frac{\mu}{r+1} \geq 0 \quad \text{and} \quad \tilde{q} = \frac{\mu r}{r+1} \geq 0, \quad (2.3)$$

where $r = q/p$, the ratio of asymmetry, and $\mu = 1 - (N - 1)\epsilon \in [0, 1]$.

In the same spirit as [5, 9], with the adapted Markov chain model, we resort to the mean first passage times from states $1, \dots, N$ to state 0 for quantifying the transient behavior of the ring network. It is convenient to denote these quantities on the original and the modified Markov chains as column vectors $z^{(0)}$ and $z^{(\epsilon)}$, respectively. In this work, we focus on investigating $\max_i z_i^{(0)}$ and $\max_i z_i^{(\epsilon)}$, i.e. the worst case analysis concerning the transient behavior of the Markov chain model. The methodology, however, can be extended to tackle the average case analysis as well.

It should be pointed out that, unlike in [5, 9], the relationship between $\max_i z_i^{(0)}$ or $\max_i z_i^{(\epsilon)}$ and the average path-length is to be interpreted with caution as for a weighted graph, there is no unequivocal notion of average path-length. Such maximum

mean first passage times, nevertheless, may be associated with, in the terminology of communication networks, the average relative throughput, overhead, or packet loss.³

Starting from now, for the sake of brevity, we simply identify the ring network, with or without long-range jumps, with the corresponding Markov chain. Thus we use such descriptions as the mean first passage times on the ring network, the transition matrix and the states of the ring network, and symmetry or asymmetry of the ring network.

3. Preliminaries. Consider the following tridiagonal matrix:

$$A = \begin{bmatrix} c_0 & c_1 & & & \\ c_{-1} & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & c_1 \\ & & & c_{-1} & c_0 \end{bmatrix} \in \mathbb{R}^{N \times N}, \quad (3.1)$$

where c_i 's are such that $c_1 \neq 0$ and

$$\delta = c_0^2 - 4c_1c_{-1} > 0. \quad (3.2)$$

Hence the equation

$$c_1x^2 + c_0x + c_{-1} = 0 \quad (3.3)$$

yields two distinct real roots, which we denote by x_1 and x_2 . We comment that condition (3.2) guarantees that A is non-singular. In fact, it is known, see [4] for example, that the spectrum of A can be written as

$$\sigma(A) = \left\{ c_0 - 2\sqrt{c_1c_{-1}} \cos \frac{i\pi}{N+1} \right\}_{i=1}^N,$$

which is obviously bounded away from 0. Suppose that $A^{-1} = [a_{i,j}^{(-1)}]$. Then, according to [8], we have:

³Generally speaking, the weights w_1 and w_2 may be thought of as costs. For instance, a link with a higher capacity or lower rate of loss may be associated with a lower cost. If so we may assume that the probabilities p and q are inversely proportional to w_1 and w_2 , respectively, i.e. that the random walk tends to follow a link with lower cost.

LEMMA 3.1. ([8, Section 3]) *Let A be defined as in (3.1). Assume that $c_1 \neq 0$ and that condition (3.2) holds. Then $A^{-1} = [a_{i,j}^{(-1)}]$ is given by*

$$a_{i,j}^{(-1)} = \begin{cases} \frac{(x_1^{-j} - x_2^{-j})(x_1^{N+1}x_2^i - x_1^i x_2^{N+1})}{c_1(x_1 - x_2)(x_1^{N+1} - x_2^{N+1})}, & j \leq i, \\ -\frac{(x_1^i - x_2^i)(x_1^{N+1-j} - x_2^{N+1-j})}{c_1(x_1 - x_2)(x_1^{N+1} - x_2^{N+1})}, & j \geq i, \end{cases} \quad (3.4)$$

where x_1 and x_2 are the two distinct real roots of (3.3).

Lemma 3.1 results immediately in the following three conclusions, whose proofs are omitted, on the row sums of A^{-1} . In the rest of this paper, we denote by e a column vector of all ones.

LEMMA 3.2. *Under the assumptions of Lemma 3.1, if $x_1 = 1$, then the i -th row sum of A^{-1} is given by*

$$(A^{-1}e)_i = -\frac{i(x_2^{N+1} - 1) - (N+1)(x_2^i - 1)}{c_1(x_2 - 1)(x_2^{N+1} - 1)}. \quad (3.5)$$

LEMMA 3.3. *Under the assumptions of Lemma 3.1, if $x_1, x_2 \neq 1$, then the i -th row sum of A^{-1} is given by*

$$(A^{-1}e)_i = -\frac{x_2^{N+1} - x_1^{N+1} - x_1^i x_2^{N+1} + x_1^{N+1} x_2^i - x_2^i + x_1^i}{c_1(x_2^{N+1} - x_1^{N+1})(1 - x_1)(x_2 - 1)}. \quad (3.6)$$

LEMMA 3.4. *Under the assumptions of Lemma 3.3, (3.6) reduces to (3.5) as x_1 approaches 1, i.e. given that $x_1, x_2 \neq 1$,*

$$\lim_{x_1 \rightarrow 1} \frac{x_2^{N+1} - x_1^{N+1} - x_1^i x_2^{N+1} + x_1^{N+1} x_2^i - x_2^i + x_1^i}{c_1(x_2^{N+1} - x_1^{N+1})(1 - x_1)(x_2 - 1)} = \frac{i(x_2^{N+1} - 1) - (N+1)(x_2^i - 1)}{c_1(x_2 - 1)(x_2^{N+1} - 1)}.$$

We now proceed to our main results regarding the maximum mean first passage times $\max_i z_i^{(0)}$ and $\max_i z_i^{(\epsilon)}$.

4. Main results.

4.1. Original Markov chain. First, we study the maximum mean first passage time to state 0 on the original ring network, whose transition matrix T_0 is expressed as in (2.1).

Let \widehat{T}_0 be the principal submatrix obtained from T_0 by deleting its first row and column. It is well-known [6, 7] that the mean first passage times from states $1, \dots, N$ to state 0 can be written as

$$z^{(0)} = (I - \widehat{T}_0)^{-1}e.$$

The following theorem formulates an explicit expression for $z^{(0)}$.

THEOREM 4.1. *For the homogeneous ergodic Markov chain with transition matrix (2.1), the mean first passage time from state i to state 0 is given by*

$$z_i^{(0)} = \frac{(r+1)[i(r^{N+1}-1) - (N+1)(r^i-1)]}{(r-1)(r^{N+1}-1)}, \quad (4.1)$$

where r is the ratio of asymmetry.

Proof. With $c_0 = 1$, $c_1 = -p$, and $c_{-1} = -q$ in (3.3), it can be easily verified that $x_1 = 1$ and $x_2 = r$. Hence formula (4.1) follows directly from Lemma 3.2. \square

It turns out that as r approaches 1, i.e. when the ring network becomes increasingly predominantly symmetric, $z_i^{(0)}$ in (4.1) approaches the mean first passage time from state i to state 0 on a symmetric ring network, which is known to be $i(N+1-i)$ [5]. We state this conclusion below without proof as it is a matter of straightforward calculation.

THEOREM 4.2. *The mean first passage time from state i to state 0 on a symmetric ring network coincides with the limit of $z_i^{(0)}$ in (4.1) as r approaches 1, i.e.*

$$\lim_{r \rightarrow 1} z_i^{(0)} = i(N+1-i).$$

It is shown in [5] that for a symmetric ring network, the maximum mean first passage time occurs roughly at $i = \frac{N+1}{2}$, namely “half-way” between states 0 and N . When it comes to an asymmetric ring network with $r > 1$, however, $\max_i z_i^{(0)}$ emerges at some state i_m which is much biased towards state N as can be seen in Figure 4.1. Numerical experiment indicates that for any fixed $r > 1$, i_m moves closer to N , as measured by i_m/N , when N increases. The next two theorems provide asymptotic results concerning i_m and $\max_i z_i^{(0)}$.

THEOREM 4.3. *When N is sufficiently large, $z_i^{(0)}$ as in (4.1) attains its maximum at*

$$i_m = N+1 - \frac{\ln[(N+1)(r-1)]}{\ln r}. \quad (4.2)$$

In addition,

$$\max_i z_i^{(0)} = \frac{(N+1)(r+1)}{r-1} \left[1 + O\left(\frac{\ln(N+1)}{N+1}\right) \right]. \quad (4.3)$$

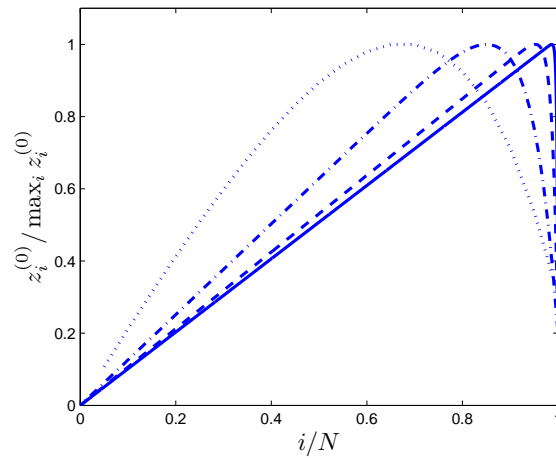


FIG. 4.1. The dotted, dashdot, dashed, and solid curves represent $z_i^{(0)}$ with N -values 20, 100, 500, and 2500, respectively, for the case when $r = 1.2$.

Proof. From (4.1), we see that the difference in $z_i^{(0)}$ can be written as

$$\begin{aligned} z_{i+1}^{(0)} - z_i^{(0)} &= \frac{(r+1) [r^{N+1} - 1 - (N+1)(r-1)r^i]}{(r-1)(r^{N+1} - 1)} \\ &= \frac{(r+1) [1 - r^{-(N+1)} - (N+1)(r-1)r^{-(N+1-i)}]}{(r-1) [1 - r^{-(N+1)}]}. \end{aligned}$$

Notice that the factor $(N+1)(r-1)r^{-(N+1-i)}$ above is increasing in i . Besides, for N large enough, $z_2^{(0)} - z_1^{(0)} > 0$ and $z_N^{(0)} - z_{N-1}^{(0)} < 0$. Consequently, there is a unique $\max_i z_i^{(0)}$, which is attained at some $1 < i < N$.

Continuing, we observe that

$$\begin{aligned} z_{i_m+1}^{(0)} - z_{i_m}^{(0)} &= \frac{(r+1) \left\{ 1 - r^{-(N+1)} - (N+1)(r-1)r^{-\frac{\ln[(N+1)(r-1)]}{\ln r}} \right\}}{(r-1) [1 - r^{-(N+1)}]} \\ &= \frac{-(r+1)r^{-(N+1)}}{(r-1) [1 - r^{-(N+1)}]} < 0 \end{aligned}$$

and, similarly, that

$$z_{i_m}^{(0)} - z_{i_m-1}^{(0)} = \frac{(r+1) [1 - r^{-(N+1)} - r^{-1}]}{(r-1) [1 - r^{-(N+1)}]} > 0,$$

which imply that (4.2) holds.

Finally, substituting (4.2) into (4.1) yields that

$$\begin{aligned} z_{i_m}^{(0)} &= \frac{(N+1)(r+1)}{r-1} - \frac{(r+1) \ln[(N+1)(r-1)]}{(r-1) \ln r} - \frac{(r+1)r^{N+1}}{(r-1)^2(r^{N+1}-1)} \\ &\quad + \frac{(N+1)(r+1)}{(r-1)(r^{N+1}-1)} \\ &= \frac{(N+1)(r+1)}{r-1} + O(\ln(N+1)). \end{aligned}$$

This concludes the proof. \square

It should be mentioned that, in fact, $\max_i z_i^{(0)}$ can be expressed as

$$\max_i z_i^{(0)} = \frac{(r+1)i_m}{r-1} + O(1).$$

The formulation for $\max_i z_i^{(0)}$ in (4.3), however, turns out to be sufficiently accurate for our purpose.

THEOREM 4.4. *Let i_m be given as in (4.2). Then*

$$\lim_{N \rightarrow \infty} \frac{z_{i_m}^{(0)}}{z_N^{(0)}} = \frac{r}{r-1}. \quad (4.4)$$

Proof. It is quite straightforward to verify this limit. \square

We comment that Theorem 4.4 confirms that for N large enough, i_m is indeed strictly smaller than N because it can be seen from (4.4) that $z_{i_m}^{(0)} > z_N^{(0)}$. In addition, this theorem serves as an alternative estimate of $\max_i z_i^{(0)}$ in terms of $z_N^{(0)}$, i.e. $\max_i z_i^{(0)} \approx \frac{rz_N^{(0)}}{r-1}$, provided that N is sufficiently large.

4.2. Modified Markov chain. Next, we investigate the maximum mean first passage time to state 0 on the modified ring network, whose transition matrix T_ϵ is formulated as in (2.2).

Let \widehat{T}_ϵ be the principal submatrix obtained from T_ϵ by deleting its first row and column. Similar to $z^{(0)}$, $z^{(\epsilon)}$ is determined by

$$z^{(\epsilon)} = (I - \widehat{T}_\epsilon)^{-1}e.$$

Note that

$$I - \widehat{T}_\epsilon = A - \epsilon ee^T, \quad (4.5)$$

where the $N \times N$ matrix A is in the form of (3.1) with

$$c_0 = 1, \quad c_1 = \epsilon - \frac{\mu}{r+1}, \quad \text{and} \quad c_{-1} = \epsilon - \frac{\mu r}{r+1}. \quad (4.6)$$

Obviously, $c_{-1} < c_1$.

Following the methodology of [5, 9], we are mainly concerned with the question as to how the small probability ϵ of jumping to non-neighboring states on the ring network affects the maximum mean first passage time to state 0, given that the size of the ring network is large enough. Recall that $0 < \epsilon \leq \frac{1}{N-1}$. Especially, c_1 becomes 0 at $\epsilon = \frac{1}{N+r}$, which lies in the range $(0, \frac{1}{N-1})$, and thus the results in Section 3 no longer apply. To avoid such singularity, we further assume, without much loss of generality, that

$$0 < \epsilon < \epsilon_m = \frac{1}{N+r}. \quad (4.7)$$

This additional assumption implies that $c_{-1} < c_1 < 0$, as is clear from (4.6).

We continue with several technical lemmas.

LEMMA 4.5. *Let c_i 's be given as in (4.6). Suppose that $0 < \epsilon < \epsilon_m$. Then the discriminant as quoted in (3.2) can be written as*

$$\delta = \frac{(r+1)^2 - 4[1 - (N+r)\epsilon][r - (Nr+1)\epsilon]}{(r+1)^2}. \quad (4.8)$$

In addition, we have that $0 < \delta < 1$ and that δ is increasing in ϵ .

Proof. Note that $(N+1)\epsilon < 2$ whenever $\epsilon < \epsilon_m$. It can be readily verified by (4.6) that

$$\delta = (N+1)\epsilon[2 - (N+1)\epsilon] + \frac{\mu^2(r-1)^2}{(r+1)^2} > 0.$$

This formula for δ reduces to (4.8) by combining with $\mu = 1 - (N-1)\epsilon$.

The fact that $\delta < 1$ can be seen from $[1 - (N+r)\epsilon][r - (Nr+1)\epsilon] > 0$. To justify that δ is increasing in ϵ , we observe that from (4.8),

$$\frac{d\delta}{d\epsilon} = \frac{4[r^2 + 2Nr + 1 - 2(N+r)(Nr+1)\epsilon]}{(r+1)^2}.$$

Since

$$\frac{r^2 + 2Nr + 1}{2(N+r)(Nr+1)} - \epsilon_m = \frac{r^2 - 1}{2(N+r)(Nr+1)} > 0,$$

it follows that $\frac{d\delta}{d\epsilon} > 0$ for any $0 < \epsilon < \epsilon_m$. \square

LEMMA 4.6. *For the matrix A as defined in (3.1) with c_i 's as given by (4.6), assuming that $0 < \epsilon < \epsilon_m$, equation (3.3) has two distinct real roots*

$$x_1 = \frac{(r+1)(1-\sqrt{\delta})}{2[1-(N+r)\epsilon]} \quad (4.9)$$

and

$$x_2 = \frac{(r+1)(1+\sqrt{\delta})}{2[1-(N+r)\epsilon]} \quad (4.10)$$

such that $0 < x_1 < 1$, $x_2 > r$, and $x_1 x_2 > 1$. Furthermore, $\lim_{\epsilon \rightarrow 0^+} x_1 = 1$, $\lim_{\epsilon \rightarrow 0^+} x_2 = r$, $\lim_{\epsilon \rightarrow \epsilon_m^-} x_1 = \frac{r-1}{N+r}$, and $\lim_{\epsilon \rightarrow \epsilon_m^-} x_2 = \infty$.

Proof. The derivation of (4.9) and (4.10) is straightforward.

It is obvious that $x_1 > 0$. On the other hand, we have that

$$x_1 = \frac{r+1 - \sqrt{(r+1)^2 - 4[1-(N+r)\epsilon][r-(Nr+1)\epsilon]}}{2[1-(N+r)\epsilon]}.$$

Notice that

$$\begin{aligned} & \left\{ r+1 - 2[1-(N+r)\epsilon] \right\}^2 - \left\{ (r+1)^2 - 4[1-(N+r)\epsilon][r-(Nr+1)\epsilon] \right\} \\ &= -4(N+1)(r+1)[1-(N+r)\epsilon]\epsilon < 0, \end{aligned}$$

which implies that $x_1 < 1$. The claim that $x_2 > r$ can be shown in a similar fashion. In addition, $x_1 x_2 = \frac{c-1}{c_1} > 1$ since $c_{-1} < c_1 < 0$.

The limits of x_1 and x_2 while ϵ is pushed towards 0 or ϵ_m can all be verified directly. \square

Before proceeding, it should be mentioned that there is an interplay between ϵ and N , assuming that r is given. In particular, for any specified ϵ , N can not be arbitrarily large due to the restriction, following (4.7), that

$$N < \frac{1}{\epsilon} - r. \quad (4.11)$$

In view of real-world applications, we shall focus on the scenario where ϵ is sufficiently close to 0 but fixed. Whether N is large enough or not should be construed in this context according to the upper bound in (4.11).

We are now ready to return to (4.5). Due to this rank-one update relationship between $I - \widehat{T}_\epsilon$ and A , it is known [5] that

$$z^{(\epsilon)} = \frac{A^{-1}e}{1 - \epsilon e^T A^{-1}e}. \quad (4.12)$$

By formula (4.12) as well as Lemmas 3.3 and 4.6, it follows:

THEOREM 4.7. *For the homogeneous ergodic Markov chain with transition matrix (2.2), the mean first passage time from state i to state 0 is in the form*

$$z_i^{(\epsilon)} = \frac{(A^{-1}e)_i}{1 - \epsilon e^T A^{-1}e}, \quad (4.13)$$

where

$$(A^{-1}e)_i = \frac{(x_1 + x_2)(x_2^{N+1} - x_1^{N+1} - x_1^i x_2^{N+1} + x_1^{N+1} x_2^i - x_2^i + x_1^i)}{(x_2^{N+1} - x_1^{N+1})(1 - x_1)(x_2 - 1)}, \quad (4.14)$$

with x_1 and x_2 being given by (4.9) and (4.10), respectively.

Proof. The conclusion is immediate. Note that $c_1 = -\frac{1}{x_1 + x_2}$. \square

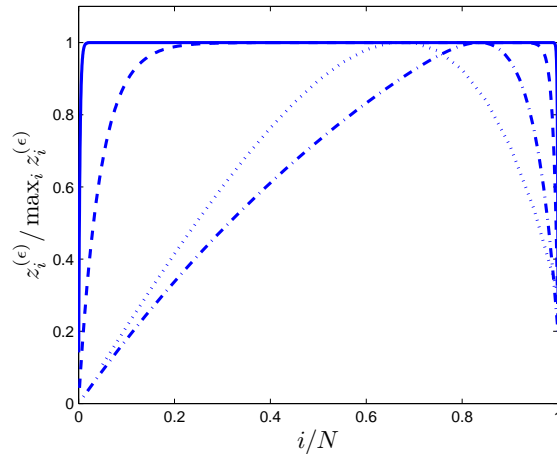


FIG. 4.2. The dotted, dashdot, dashed, and solid curves represent $z_i^{(\epsilon)}$ with N -values 20, 100, 500, and 2500, respectively, for the case when $\epsilon = 1 \times 10^{-5}$ and $r = 1.2$.

Numerical experiments suggest that when ϵ is small but prescribed, and when N is large enough, unlike $\max_i z_i^{(0)}$, $\max_i z_i^{(\epsilon)}$ tends to emerge at some state i_m “half-way” between states 0 and N as illustrated in Figure 4.2. In fact, when N is very large, the graph of $z_i^{(\epsilon)}$ appears to be essentially flat except at i values close to 1 or

N . We recall that according to Lemma 3.4, as ϵ approaches 0, $z_i^{(0)}$ and $z_i^{(\epsilon)}$ become identical. It should be noted, however, that the test result in Figure 4.2 does not contradict that in Figure 4.1 since the value of ϵ here is fixed, rather than being pushed towards 0.

Using Theorem 4.7, we can establish as follows two asymptotic results concerning i_m and $\max_i z_i^{(\epsilon)}$.

THEOREM 4.8. *For any sufficiently small $0 < \epsilon < \epsilon_m$, when N is sufficiently large, $z_i^{(\epsilon)}$ attains its maximum at*

$$i_m = \frac{(N+1) \ln x_2}{\ln \left(\frac{x_2}{x_1} \right)}, \quad (4.15)$$

which satisfies that $i_m > \frac{N+1}{2}$. Furthermore,

$$\max_i z_i^{(\epsilon)} = \frac{N+1}{\sqrt{\delta}} + O \left(\max \left\{ x_1^{\frac{N+1}{2}}, x_2^{-N} \right\} \right), \quad (4.16)$$

where δ is given as in (4.8).

Proof. First of all, it is clear from (4.13) that $z_i^{(\epsilon)}$ is maximized at some i whenever so is $(A^{-1}e)_i$. As a result, it suffices to consider $(A^{-1}e)_i$ in order to estimate i_m . By (4.14), the difference in $(A^{-1}e)_i$ can be expressed as

$$\begin{aligned} & (A^{-1}e)_{i+1} - (A^{-1}e)_i \\ &= \frac{(x_1 + x_2) \left\{ x_1^i (1 - x_1) \left[1 - x_2^{-(N+1)} \right] - x_2^i (x_2 - 1) (1 - x_1^{N+1}) x_2^{-(N+1)} \right\}}{\left[1 - x_1^{N+1} x_2^{-(N+1)} \right] (1 - x_1)(x_2 - 1)}. \end{aligned}$$

According to Lemma 4.6, $0 < x_1 < 1$ and $x_2 > r$, leading to the observation that the term $x_1^i (1 - x_1) \left[1 - x_2^{-(N+1)} \right]$ is decreasing in i , while the term $x_2^i (x_2 - 1) (1 - x_1^{N+1}) x_2^{-(N+1)}$ is increasing in i . Besides, for N large enough, $(A^{-1}e)_2 - (A^{-1}e)_1 > 0$ and $(A^{-1}e)_N - (A^{-1}e)_{N-1} < 0$. Consequently, there is a unique $\max_i z_i^{(\epsilon)}$, which is attained at some $1 < i < N$.

Continuing, with i_m as in (4.15), we estimate $(A^{-1}e)_{i_m+1} - (A^{-1}e)_{i_m}$ for large N :

$$\begin{aligned} (A^{-1}e)_{i_m+1} - (A^{-1}e)_{i_m} &= \frac{(x_1 + x_2)x_1^{i_m}}{\left[1 - x_1^{N+1} x_2^{-(N+1)} \right] (1 - x_1)(x_2 - 1)} \left\{ (1 - x_1) \left[1 - x_2^{-(N+1)} \right] \right. \\ &\quad \left. - \left(\frac{x_2}{x_1} \right)^{\frac{(N+1) \ln x_2}{\ln \left(\frac{x_2}{x_1} \right)}} (x_2 - 1) (1 - x_1^{N+1}) x_2^{-(N+1)} \right\} \\ &\approx \frac{(x_1 + x_2)x_1^{i_m} (2 - x_1 - x_2)}{(1 - x_1)(x_2 - 1)} < 0 \end{aligned}$$

since $x_1 + x_2 = \frac{r+1}{1-(N+r)\epsilon} > 2$. Similarly, we find that

$$(A^{-1}e)_{i_m} - (A^{-1}e)_{i_m-1} \approx \frac{(x_1 + x_2)x_1^{i_m} \left(1 + \frac{x_1}{x_2}\right)}{(1-x_1)(x_2-1)} > 0.$$

Hence (4.15) is valid.

In addition, again by Lemma 4.6, $x_1 x_2 > 1$, which yields that $i_m > \frac{N+1}{2}$. This can be justified by considering $2 \ln x_2 = \ln \left[\left(\frac{x_2}{x_1} \right) x_1 x_2 \right] > \ln \left(\frac{x_2}{x_1} \right)$.

Now, using (4.14), we obtain that

$$\begin{aligned} (A^{-1}e)_{i_m} &= \frac{(x_1 + x_2) \left\{ 1 - x_1^{N+1} x_2^{-(N+1)} - x_1^{i_m} \left[2 - x_1^{N+1} - x_2^{-(N+1)} \right] \right\}}{\left[1 - x_1^{N+1} x_2^{-(N+1)} \right] (1-x_1)(x_2-1)} \\ &= \frac{x_1 + x_2}{(1-x_1)(x_2-1)} + O\left(x_1^{\frac{N+1}{2}}\right) \end{aligned}$$

and that

$$\begin{aligned} e^T A^{-1} e &= \frac{x_1 + x_2}{\left[1 - x_1^{N+1} x_2^{-(N+1)} \right] (1-x_1)(x_2-1)} \left\{ N \left[1 - x_1^{N+1} x_2^{-(N+1)} \right] \right. \\ &\quad \left. - \frac{(x_1 - x_1^{N+1}) \left[1 - x_2^{-(N+1)} \right]}{1-x_1} - \frac{(1-x_1^{N+1})(1-x_2^{-N})}{x_2-1} \right\} \\ &= \frac{x_1 + x_2}{(1-x_1)(x_2-1)} \left[N - \frac{x_1}{1-x_1} - \frac{1}{x_2-1} \right] + O\left(\max\{x_1^{N+1}, x_2^{-N}\}\right). \end{aligned}$$

Thus, we arrive at the following:

$$\begin{aligned} \max_i z_i^{(\epsilon)} &= \frac{(x_1 + x_2)(1-x_1)(x_2-1)}{(1-x_1)^2(x_2-1)^2 - \epsilon(x_1 + x_2) [N(1-x_1)(x_2-1) - x_1(x_2-1) - (1-x_1)]} \\ &\quad + O\left(\max\left\{x_1^{\frac{N+1}{2}}, x_2^{-N}\right\}\right). \end{aligned} \quad (4.17)$$

From (4.9) and (4.10), it is a matter of direct calculation to verify that

$$(1-x_1)(x_2-1) = \frac{(N+1)(r+1)\epsilon}{1-(N+r)\epsilon}.$$

Also note that, as mentioned earlier in the proof, $x_1 + x_2 = \frac{r+1}{1-(N+r)\epsilon}$. These reductions, together with (4.9), (4.10), and (4.17), finally yield (4.16). \square

We mention that from the foregoing proof, it is quite clear that in fact for any i_m in the form of $i_m = (N+1)\gamma$, where $0 < \gamma < 1$ is fixed, $(A^{-1}e)_{i_m+1} - (A^{-1}e)_{i_m} \approx 0$, which explains the plateau-shaped graph of $z_i^{(\epsilon)}$.

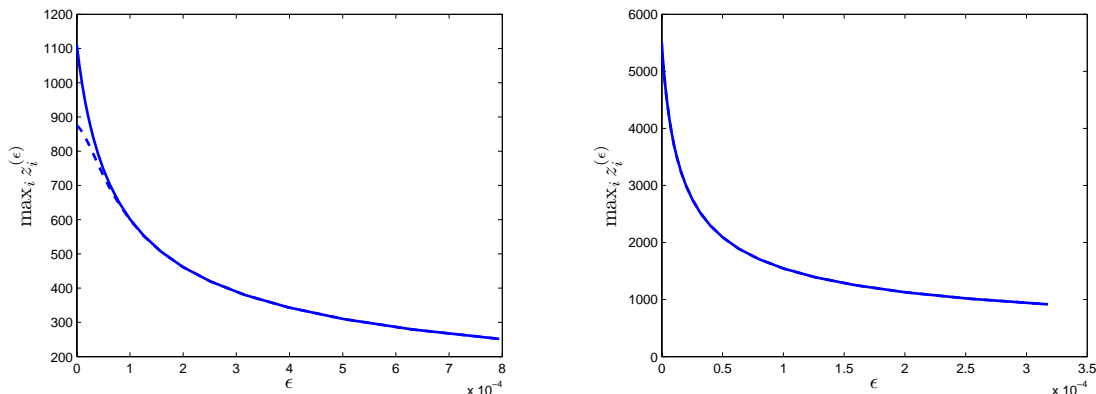


FIG. 4.3. The solid and dashed curves represent the estimate of $\max_i z_i^{(\epsilon)}$ as provided by (4.16) and the exact $\max_i z_i^{(\epsilon)}$, respectively. The graph on the left illustrates the case when $N = 100$ and $r = 1.2$, whereas the graph on the right illustrates the case when $N = 500$ and $r = 1.2$. For each case, the smallest ϵ value is determined by $10^{-8}/N^2$.

The formulation of error terms in the preceding proof is based on the observation that $0 < x_1 < 1$ and $x_2 > r$. Especially, when N is extremely large, i.e. when $\epsilon \approx \frac{1}{N+r}$ by (4.11), we see from (4.9) that $x_1 \approx \frac{r-1}{N+r}$, which is far smaller than 1. This also implies that x_2 is far greater than r since $x_1 x_2 > 1$. Thus, the error terms in the proof, including the one as in (4.16), are negligible when N becomes sufficiently large.

The accuracy of the estimate, without the error term, in formula (4.16) is confirmed by numerical experiment as well. Two such examples are shown in Figure 4.3, where the N -values are indeed quite moderate in light of the ϵ -value. Note that in the graph on the right, the difference between the two curves is barely discernible. Numerical results also evidence that the larger N is, the closer the approximation in (4.16) tends to be towards the actual $\max_i z_i^{(\epsilon)}$.

THEOREM 4.9. Set $\zeta^{(\epsilon)} = \frac{N+1}{\sqrt{\delta}}$, where δ is formulated as in (4.8). Then

$$\lim_{\epsilon \rightarrow 0^+} \zeta^{(\epsilon)} = \frac{(N+1)(r+1)}{r-1},$$

i.e.

$$\lim_{\epsilon \rightarrow 0^+} \zeta^{(\epsilon)} \approx \max_i z_i^{(0)},$$

provided that N is sufficiently large. Moreover, $\zeta^{(\epsilon)}$ is decreasing in ϵ for any $0 < \epsilon < \epsilon_m$.

Proof. The claim regarding $\lim_{\epsilon \rightarrow 0} \zeta^{(\epsilon)}$ follows immediately by referring to (4.3) and (4.16). In addition, the monotonicity of $\zeta^{(\epsilon)}$ is a direct consequence of Lemma 4.5. \square

Before concluding this section, we comment that the estimates as in (4.3) and (4.16) hinge upon the fact that $r > 1$. This means that as r approaches 1, they do not reduce to the corresponding results for the symmetric ring network as can be found in [5]. However, for any fixed $r > 1$, they produce good approximations for the current asymmetric case, provided that N is sufficiently large.

5. Transient behavior of the Markov chain.

5.1. Reduction ratio. Following [5, 9], we define the reduction ratio ρ in the maximum mean first passage time to be

$$\rho = \frac{\max_i z_i^{(\epsilon)}}{\max_i z_i^{(0)}}. \quad (5.1)$$

The question that we are concerned with is how ρ responds to changes in ϵ , the probability of jumping to non-neighboring states, and in r , the ratio of asymmetry, with the assumption that N is sufficiently large. In particular, as in [5, 9], an important issue is whether ρ undergoes a considerable drop as ϵ increases. The development in the preceding section allows us to derive asymptotic results regarding the effects of varying ϵ and r on the behavior of ρ .

According to Theorems 4.3, 4.8, and 4.9, we obtain the following useful result on the reduction ratio ρ :

THEOREM 5.1. *For any sufficiently small $0 < \epsilon < \epsilon_m$, when N is sufficiently large, the reduction ratio in (5.1) can be expressed as*

$$\rho = \tilde{\rho} + O\left(\frac{\ln(N+1)}{N+1}\right), \quad (5.2)$$

where

$$\tilde{\rho} = \frac{r-1}{\sqrt{(r+1)^2 - 4[1 - (N+r)\epsilon][r - (Nr+1)\epsilon]}}. \quad (5.3)$$

In addition, we have that $\lim_{\epsilon \rightarrow 0^+} \tilde{\rho} = 1$ and that $\tilde{\rho}$ is decreasing in ϵ .

The monotonicity of $\tilde{\rho}$ implies the fact that $\tilde{\rho}$ attains its minimum at $\epsilon = \epsilon_m$. This minimum can be interpreted as the best reduction ratio ρ_b . From (5.2), we arrive at the next conclusion.

THEOREM 5.2. *Let $0 < \epsilon < \epsilon_m$ be sufficiently small. If N is sufficiently large, then the best reduction ratio ρ_b can be formulated as*

$$\rho_b = \frac{r-1}{r+1}. \quad (5.4)$$

Proof. The conclusion follows immediately from (5.3). We point out that estimate (5.4) can be derived in an alternative way. Recall that by Theorem 4.6, $\lim_{\epsilon \rightarrow \epsilon_m^-} x_1 = \frac{r-1}{N+r}$ and $\lim_{\epsilon \rightarrow \epsilon_m^-} x_2 = \infty$. On the other hand, we see from (4.17) that, with x_2^2 being a dominant factor as ϵ is near ϵ_m ,

$$\max_i z_i^{(\epsilon)} \approx \frac{1-x_1}{(1-x_1)^2 - N(1-x_1)\epsilon + x_1\epsilon},$$

which further reduces to a value of $N+1$ as ϵ approaches ϵ_m . This limit, together with (4.3), lead to (5.4). \square

Theorem 5.2 reveals the connection between ρ_b and r . Specifically, when r is close to 1, ρ_b can be extremely small; as r increases, however, so does ρ_b , thus less significant the drop in ρ is expected to be. A special case of much interest is stated below.

THEOREM 5.3. *Suppose that $0 < \epsilon < \epsilon_m$ is sufficiently small and that N is sufficiently large. Then the best reduction ratio $\rho_b \geq 1/2$ when $r \geq 3$. Furthermore, $\rho_b \approx 1$ for r large enough.*

It is demonstrated in [5] that for a symmetric ring network, $\rho \leq 1/2$ so long as $\epsilon \gtrsim \frac{32}{(N+1)^3}$, meaning that a substantial drop in ρ can always be achieved as ϵ increases. In contrast to this, we see from Theorem 5.3 the rather striking impact of asymmetry on the transient behavior of the ring network: *No significant reduction in ρ can be attained on a large-scale ring network with ratio of asymmetry $r \geq 3$.*

5.2. Examples. With the estimate in (5.2), we are now in a position to explore numerically the transient behavior of the ring network relative to changes in ϵ and r .

Following [5, 9], we consider that $\epsilon = k/N^\alpha$, where the parameters $k > 0$ and $\alpha > 1$. We choose k -values in the range $N^\alpha \times 10^{-8} < k < N^\alpha \epsilon_m$ based on (4.7). We mention that in all numerical trials, the reduction ratio ρ is computed from formula (5.2) unless stated otherwise.

The role of parameter α is considered in [5, 9]. The results in [9] apply to the case when $\alpha = 3$, while those in [5] apply to the case when $\alpha > 1$. Intuitively, varying α amounts to scaling ϵ , and hence it does not contribute to the intrinsic property of the ring network. This observation carries over to the asymmetric ring network as

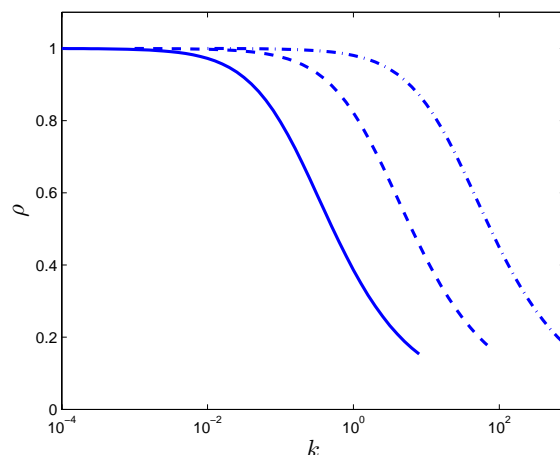


FIG. 5.1. The solid, dashed, and dashdot curves represent the reduction ratio ρ for $\alpha = 1.6$, 2 , and 2.4 , respectively. For each case, N and r are fixed at $N = 500$ and $r = 1.2$.

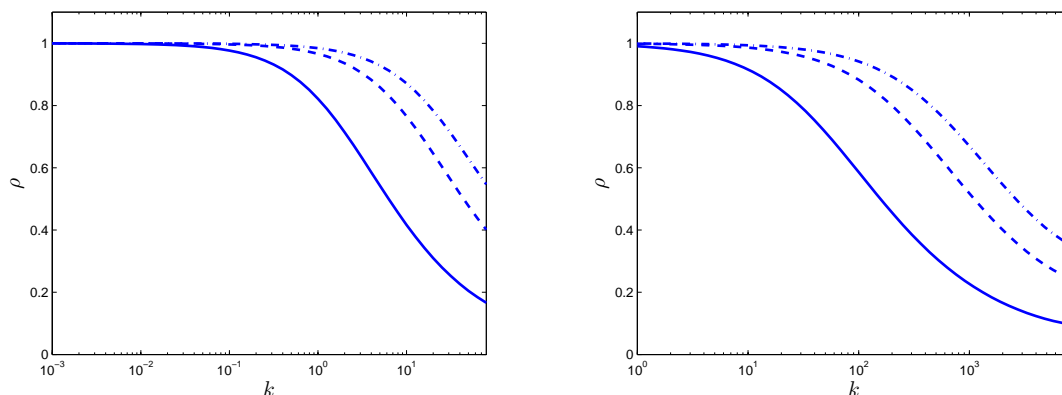


FIG. 5.2. The solid, dashed, and dashdot curves represent the reduction ratio ρ for $r = 1.2$, 1.6 , and 2 , respectively. The graph on the left shows the case when $N = 500$ and $\alpha = 2$, whereas the graph on the right shows the case when $N = 12500$ and $\alpha = 2$.

illustrated in Figure 5.1. Notice that the curves in this figure are in fact on different ϵ -axes.

Figure 5.1 also shows that for $r = 1.2$, there is an abrupt drop in ρ as ϵ moves away from 0 . We comment that it only takes a very small ϵ to induce a 50% reduction in ρ . When $\alpha = 2$, for instance, the example in Figure 5.1 gives $\rho = 0.4990$ at $k = 6.3096$, which translates into $\epsilon = 2.5238 \times 10^{-5}$, far smaller than ϵ_m . It should be noted that

at this ϵ -value, $\tilde{p} = 0.4488$ and $\tilde{q} = 0.5386$ by (2.3), manifesting that the ring network remains to be dominated by asymmetric processes among its neighboring vertices.

Next, we display in Figure 5.2 the effect of varying the ratio of asymmetry r on the reduction ratio ρ . The first observation is that with the growth of r , the graph of ρ shifts towards the right, verifying that the reduction in ρ gradually diminishes. The second observation is that as r goes up, the maximum drop in ρ becomes smaller, confirming the prediction as in Theorem 5.2. When N is large enough, as illustrated in the graph on the right, the best reduction ratios for $r = 1.2, 1.6$, and 2 move closer to their estimates from (5.4), namely $0.0909, 0.2308$, and 0.3333 , respectively.

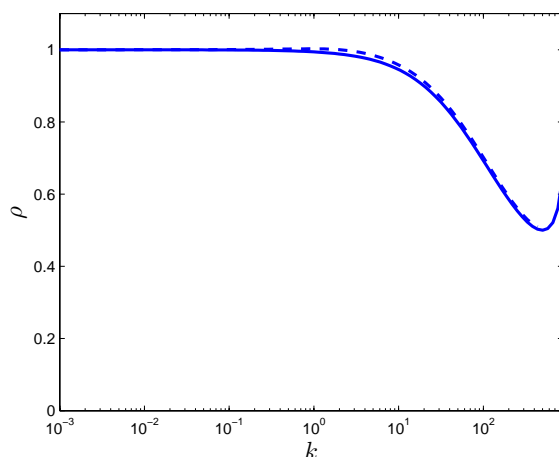


FIG. 5.3. The solid curve represents the reduction ratio ρ as in (5.2) for the case when $N = 500$, $\alpha = 2$, and $r = 3$. Note that in this example, ϵ goes beyond the range $(0, \epsilon_m)$. The dashed curve represents the actual reduction ratio as in (5.1) for the same N , α , and r .

Finally, Figure 5.3 provides an example for the case when $r = 3$. There is approximately a 50% decrease in ρ as ϵ is near ϵ_m , which is again consistent with Theorem 5.2. It is interesting to observe in Figure 5.3 that as ϵ passes beyond ϵ_m , ρ actually starts to bounce back. This indicates in a different perspective, besides consideration of avoiding singularity, that it is indeed reasonable to restrict that $\epsilon < \epsilon_m$. As a comparison, we also plot in the same figure the exact reduction ratio as obtained from (4.1), (4.13), and (5.1). Clearly, the estimated reduction ratio conforms well, even for $\epsilon \geq \epsilon_m$, to the exact reduction ratio.

6. Conclusions. Using the adapted Markov chain model, we study the transient behavior of the ring network which involves asymmetric interaction. We establish useful asymptotic results on the maximum mean first passage time, a key quantity in examining the property of the ring network, for both the original and the modified

Markov chains. In addition, we provide numerical examples to illustrate the conclusions.

It is quite interesting to notice that the ratio of asymmetry r plays an important role in determining the transient behavior of the ring network, as measured by the reduction ratio ρ . Specifically, a significant decrease in ρ can be reached when r is close to 1; as r moves away from 1, however, the decrease in ρ starts to diminish. This discovery may shed light on real-world networks where asymmetric interaction takes place, such as those with unbalanced chemical reaction fluxes, traffic flows, energy transfer rates, and bandwidths. Albeit simplistic and crude, the Markov chain model may well address some fundamental issues in applications.

Finally, it should be pointed out that our results are developed within the framework of the Markov chain model which is associated with the underlying weighted digraph. It is an intriguing, and important, question as to how these results may be interpreted in the context of the small-world phenomenon on the ring network. Seeking for the answer to this question is a part of our ongoing research.

Acknowledgment. The author is grateful to an anonymous reviewer for constructive comments on improving the presentation of the results in this paper.

REFERENCES

- [1] E. Almaas, B. Kovács, T. Vicsek, Z.N. Oltvai, and A.-L. Barabási. Global organization of metabolic fluxes in the bacterium *Escherichia Coli*. *Nature*, 427(6977):839–843, 2004.
- [2] H. Balakrishnan and V.N. Padmanabhan. How network asymmetry affects TCP. *IEEE Comm. Magazine*, 39(4):2–9, 2001.
- [3] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. *Proc. Nat. Acad. Sci.*, 101(11):3747–3752, 2004.
- [4] R. Bellman. *Introduction to Matrix Analysis*. McGraw-Hill, New York, 1970.
- [5] M. Catral, M. Neumann, and J. Xu. Matrix analysis of a Markov chain small-world model. *Linear Algebra Appl.*, 409:126–146, 2005.
- [6] G.E. Cho and C.D. Meyer. Markov chain sensitivity measured by mean first passage times. *Linear Algebra Appl.*, 316:21–28, 2000.
- [7] E. Dietzenbacher. *Perturbations and Eigenvectors: Essays*. Ph.D. Thesis, University of Groningen, The Netherlands, 1990.
- [8] M. Dow. Explicit inverses of Toeplitz and associated matrices. *ANZIAM J.*, 44(E):185–215, 2003.
- [9] D.J. Higham. A matrix perturbation view of the small world phenomenon. *SIAM J. Matrix Anal. Appl.*, 25(4):429–444, 2003.
- [10] D.J. Higham. Greedy pathlengths and small world graphs. *Linear Algebra Appl.*, 416:745–758, 2006.
- [11] J.G. Kemeny and J.L. Snell. *Finite Markov Chains*. Van Nostrand, Princeton, New Jersey, 1960.
- [12] A.E. Krause, K.A. Frank, D.M. Mason, R.E. Ulanowicz, and W.W. Taylor. Compartments revealed in food-web structure. *Nature*, 426(6964):282–285, 2003.

- [13] M.E.J. Newman. The structure and function of complex networks. *SIAM Rev.*, 45(2):167–256, 2003.
- [14] M.E.J. Newman, C. Moore, and D.J. Watts. Mean-field solution of the small-world network model. *Phys. Rev. Lett.*, 84:3201–3204, 2000.
- [15] D.J. Watts and S.H. Strogatz. Collective dynamics of “small-world” networks. *Nature*, 393:440–442, 1998.