Res. Lett. Inf. Math. Sci. (2008) **12**, 1-23
Available online at http://iims.massey.ac.nz/research/letters/

Bounds on Expected Coupling Times in Markov Chains

Jeffrey J. Hunter
I.I.M.S. Massey University, Auckland, New Zealand
j.hunter@massey.ac.nz

Abstract
In the author's paper "Coupling and Mixing Times in Markov Chains" (RLIMS, 11, 1-22, 2007) it was shown that it is very difficult to find explicit expressions for the expected time to coupling in a general Markov chain. In this paper simple upper and lower bounds are given for the expected time to coupling in a discrete time finite Markov chain. Extensions to the bounds under additional restrictive conditions are also given with detailed comparisons provided for two and three state chains.

1. Introduction

In [2] the derivation of the expected time to coupling in a Markov chain and its relation to the expected time to mixing (as introduced in [3], see also [1], [6])) was explored and the two-state cases and three-state cases were examined in detail.

Considerable difficulty was experienced in attempting to obtain closed form expressions for the expected coupling times. The main thrust of this paper is to explore the derivation of easily computable upper and lower bounds on these expectations.

In Section 2 we summarise the main results on coupling. In Section 3 we derive some new bounds and in Section 4 we compare these bounds with special cases considered in [2].

2. Coupling times

Let $P = [p_{ij}]$ be the transition matrix of a finite irreducible, discrete time Markov chain $\{X_n\}$, $(n \geq 0)$, with state space $S = \{1, 2,\ldots, m\}$. Such Markov chains have a unique stationary distribution $\{\pi_j\}$, $(1 \leq j \leq m)$, that, in the case of a regular (finite, irreducible and aperiodic) chain, is also the limiting distribution of the Markov chain ([5, Theorem 7.1.2]). Let $\boldsymbol{\pi}^T = (\pi_1, \pi_2, \ldots, \pi_m)$ be the stationary probability vector of the Markov chain.

Coupling of Markov chains can be described as follows. Start a Markov chain $\{Y_n\}$, with the same transition matrix P and state space S as for $\{X_n\}$, operating under stationary conditions, so that the initial probability distribution for Y_0 is the stationary distribution $\{\pi_j\}$. Start the Markov chain $\{X_n\}$ in an initial state i and allow each

Markov chain to evolve, independently, until time $T = n$ when both chains $\{X_n\}$ and $\{Y_n\}$ reach the same state for the first time at this *n-th* trial. We call this the "coupling time" since after time T each chain is coupled and evolves identically as the $\{Y_n\}$ Markov chain, with each chain having the same distribution at each subsequent trial, the stationary distribution $\{\pi_j\}$.

$\mathbf{Z}_n = (X_n, Y_n)$, $(n \geq 0)$, is a (two-dimensional) Markov chain with state space $S \times S$. The chain is an absorbing chain with absorbing (coupling) states $C = \{(i, i), 1 \leq i \leq m\}$ and transient states $\mathcal{T} = \{(i, j), i \neq j, 1 \leq i \leq m, 1 \leq j \leq m\}$. The transition probabilities, prior to coupling, are given by $P\{\mathbf{Z}_{n+1} = (k, l) \mid \mathbf{Z}_n = (i, j)\} = p_{ik}\, p_{jl}$, (see [2]). Once coupling occurs at time $T = n$, $X_{n+k} = Y_{n+k}$ for all $k \geq 0$.

If $\mathbf{Z}_0 \in C$, coupling of the two Markov chains is instantaneous and the coupling time $T =$ 0. Define $T_{ij,kl}$ to be the first passage time from state (i, j) to state (k, l). The time to coupling in state k, starting in state (i, j), $(i \neq j)$, is the first passage time $T_{ij,kk}$ to the absorbing state (k, k). Let $T_{ij,C,}$ be the first passage time from (i, j), $(i \neq j)$ to the absorbing (coupling) states C. Define $T_{ii,C} = 0$, $(1 \leq i \leq m)$, consistent with the coupling occurring instantaneously if $X_0 = Y_0$ (in state i).

Under the assumption that the embedded Markov chains, X_n and Y_n, are irreducible and aperiodic (i.e regular) the transition matrix for the two dimensional Markov chain can be represented in the canonical form for an absorbing Markov chain, as

$$P = \begin{bmatrix} I & 0 \\ R & Q \end{bmatrix},$$

where I is an $m \times m$ identity matrix, Q is an $m(m-1) \times m(m-1)$ matrix governing the transition probabilities within the transient states $\mathcal{T}$, and R is an $m(m-1) \times m$ matrix governing the transition probabilities from the transient states $\mathcal{T}$ to the absorbing (coupling) states C.

Note that if the Markov chains, X_n and Y_n are periodic (period m) then mixing either occurs initially or never occurs! We restrict attention to embedded regular chains.

In [2] it was shown that, with probability one, starting in state (i, j), coupling will occur in finite time. Let $\kappa_{ij}^{(C)} = E[T_{ij,C}]$ be the expected time to coupling starting in state $X_0 =$ i, $Y_0 = j$, and let $\boldsymbol{\kappa}^{(C)} \equiv (\kappa_{ij}^{(C)})$ be the column vector (of dimension $m(m-1) \times 1$) of the expected times to coupling. Then all the expected values are finite and, [2],

$$\boldsymbol{\kappa}^{(C)} = (I - Q)^{-1}\boldsymbol{e}. \qquad (2.1)$$

Since the states of the Markov chain $\{Y_n\}$ have at each trial the stationary distribution, and since coupling occurs initially if $i = j$ with $T_{ii,C} = 0$, the expected time to coupling starting in state i, $(1 \le i \le m)$ is

$$\tau_{C,i} = \sum_{j=1}^{m} \pi_j E[T_{ij,C}] = \sum_{j \ne i} \pi_j \kappa_{ij}^{(C)}. \tag{2.2}$$

Let $\boldsymbol{\kappa}_1^T = (\kappa_{12}^{(C)},...,\kappa_{1j}^{(C)},...,\kappa_{1m}^{(C)}), \ldots, \boldsymbol{\kappa}_i^T = (\kappa_{i1}^{(C)},...,\kappa_{i,i-1}^{(C)},\kappa_{i,i+1}^{(C)},...\kappa_{im}^{(C)}), \ldots,$ $\boldsymbol{\kappa}_m^T = (\kappa_{m1}^{(C)},...,\kappa_{m,m-1}^{(C)})$, and re-express $\boldsymbol{\kappa}$ as $\boldsymbol{\kappa}^T = (\boldsymbol{\kappa}_1^T,..., \boldsymbol{\kappa}_i^T,..., \boldsymbol{\kappa}_m^T)$.

Define $\boldsymbol{\rho}_i^T = \boldsymbol{\pi}^T[\boldsymbol{e}_1,\boldsymbol{e}_2,...,\boldsymbol{e}_{i-1},\boldsymbol{e}_{i+1},...,\boldsymbol{e}_m] = (\pi_1,...,\pi_{i-1},\pi_{i+1},...\pi_m)$, a modification of $\boldsymbol{\pi}^T$ to yield a vector of dimension $1\times (m\ \text{-}1)$ (with π_i removed at the *i-th* position from $\boldsymbol{\pi}^{\mathrm{T}}$). For $1 \le i \le m$,

$$\tau_{C,i} = \boldsymbol{\rho}_i^T \boldsymbol{\kappa}_i .$$

From (2.1) observe that $\boldsymbol{\kappa}$ can be obtained by solving the set of linear equations

$$(I - Q)\boldsymbol{\kappa}^{(C)} = \boldsymbol{e}. \tag{2.3}$$

The Q-matrix is of dimension $m(m-1) \times m(m-1)$ and governs the transitions within the $m(m-1)$ transient states. This matrix contains some symmetry. The sub-matrix of one-step transition probabilities governing transitions between the states (i, j) and (j, i) $(i \ne j)$ has the structure

$$\begin{array}{c c} & \begin{array}{cc} (i,j) & (j,i) \end{array} \\ \begin{array}{c} (i,j) \\ (j,i) \end{array} & \begin{bmatrix} p_{ii}p_{jj} & p_{ij}p_{ji} \\ p_{ji}p_{ij} & p_{jj}p_{ii} \end{bmatrix} \end{array}.$$

The transition probabilities from (i, j) to the other transient states have some symmetrical reciprocity, i.e. for $i \ne j$ and $r \ne s$,

$P[(X_{n+1},Y_{n+1}) = (r, s) \mid (X_n,Y_n) = (i, j)] = p_{ir}p_{js} = P[(X_{n+1},Y_{n+1}) = (s, r) \mid (X_n,Y_n) = (j, i)]$.

The one step transition to any coupling state (k, k) has the same probability from either (i, j) or (j, i) i.e.

$P[(X_{n+1},Y_{n+1}) = (k, k) \mid (X_n,Y_n) = (i, j)] = p_{ik}p_{jk} = P[(X_{n+1},Y_{n+1}) = (k, k) \mid (X_n,Y_n) = (j, i)]$.

Thus by labelling the states in successive symmetrical pairs, each even numbered row of Q has the same probabilities, but interchanged in pairs, as the previous odd numbered row. Furthermore these pairs of rows have identical probabilities in the same place in the R matrix.

The net effect is that instead of solving the $m(m-1)$ linear equations present in (2.3), we need only solve a reduced number of $m(m-1)/2$ linear equations. This is effected by observing that $\kappa_{ij}^{(C)} = \kappa_{ji}^{(C)}$ so that only these $m(m-1)/2$ quantities (with $i < j$, say) actually need to be solved. We elaborate further on this later.

We introduce some notation.

$$\begin{aligned}\text{Let } \mu_{ij} &= \sum_{r=1}^{m} p_{ir} p_{jr} = \sum_{r=1}^{m} P\{(X_{n+1}, Y_{n+1}) = (r,r) \mid (X_n, Y_n) = (i,j)\} \\ &= P\{(X_{n+1}, Y_{n+1}) \in C \mid (X_n, Y_n) = (i,j)\} \\ &= P\left\{\text{Coupling occurs at the next trial } \middle| \text{The 2-dim MC is in state } (i,j)\right\}.\end{aligned}$$

Observe that $\mu_{ij} = \boldsymbol{p}_i^{(r)T} \boldsymbol{p}_j^{(r)} = \mu_{ji}$ where $\boldsymbol{p}_i^{(r)T} = (p_{i1}, p_{i2}, ..., p_{im})$, the i-th row of the transition matrix P.

3. Bounds

In a general Markov chain setting, elemental expressions of the key equations, Eqn. (2.3), lead, for all $i \neq j$, to

$$\kappa_{ij}^{(C)} - 1 = \sum\sum_{r \neq s} p_{ir} p_{js} \kappa_{rs}^{(C)}. \tag{3.1}$$

We deduce upper and lower bounds for $\kappa_{ij}^{(C)}$ from Eqns. (3.1).

Theorem 1. *If $\mu_{ij} > 0$ for all $i \neq j$, then, for all $i \neq j$,*

$$\kappa_{\min} \leq \kappa_{ij}^{(C)} \leq \kappa_{\max}, \tag{3.2}$$

where $\kappa_{\min} = \dfrac{1}{\max_{i \neq j} \mu_{ij}}$ *and* $\kappa_{\max} = \dfrac{1}{\min_{i \neq j} \mu_{ij}}$.

Proof: Assume that for all $r \neq s$, $\kappa_{rs}^{(C)} \leq \kappa_{\max}$. (3.3)

Observe that

$$1 = (\sum_{r=1}^{m} p_{ir})(\sum_{s=1}^{m} p_{js}) = \sum_{r=s} p_{ir} p_{js} + \sum\sum_{r \neq s} p_{ir} p_{js} = \mu_{ij} + \sum\sum_{r \neq s} p_{ir} p_{js}. \tag{3.4}$$

From Eqn. (3.1) and Eqn. (3.4) it follows that

$$\kappa_{ij}^{(C)} \leq 1 + \left(\sum\sum_{r \neq s} p_{ir} p_{js}\right) \kappa_{\max} = 1 + (1 - \mu_{ij}) \kappa_{\max}. \tag{3.5}$$

Assumption (3.3) implies, using inequality (3.5), that it is sufficient to take $1 + (1 - \mu_{ij}) \kappa_{\max} \leq \kappa_{\max}$ and hence that $\mu_{ij} \kappa_{\max} \geq 1$, i.e. $\kappa_{\max} \geq \dfrac{1}{\mu_{ij}}$ for all $i \neq j$. This is satisfied by taking $\kappa_{\max} = \max\limits_{i \neq j} \dfrac{1}{\mu_{ij}} = \dfrac{1}{\min_{i \neq j} \mu_{ij}} = \dfrac{1}{\mu_{\min}}$.

Similarly let us assume that for all $r \neq s$, $\kappa_{\min} \le \kappa_{rs}^{(C)}$. (3.6)

From Eqn. (3.1) and Eqn.(3.4) we have that

$$\kappa_{ij}^{(C)} \ge 1 + \left(\sum\sum_{r \neq s} p_{ir} p_{js}\right)\kappa_{\min} = 1 + (1 - \mu_{ij})\kappa_{\min}. \tag{3.7}$$

Similar to the argument used above, using assumption (3.6) and inequality (3.7), we require $1 + (1 - \mu_{ij})\kappa_{\min} \ge \kappa_{\min}$ and hence that $\mu_{ij}\kappa_{\min} \le 1$. Thus, $\kappa_{\min} \le \dfrac{1}{\mu_{ij}}$ for all $i \neq j$, which is satisfied by taking $\kappa_{\min} = \min_{i \neq j} \dfrac{1}{\mu_{ij}} = \dfrac{1}{\max_{i \neq j} \mu_{ij}} = \dfrac{1}{\mu_{\max}}$. □

Corollary 1.1: *Provided* $\mu_{ij} > 0$ *for all* $i \neq j$,

$$\frac{(1 - \pi_i)}{\mu_{\max}} = (1 - \pi_i)\kappa_{\min} \le \tau_{C,i} \le (1 - \pi_i)\kappa_{\max} = \frac{(1 - \pi_i)}{\mu_{\min}} . \tag{3.8}$$

Proof: Inequalities (3.8) follow directly from Eqn. (2.2) and Eqn. (3.2). □

If the stationary distribution $\{\pi_i\}$ of the underlying Markov chain is unknown then a simpler, but slightly larger, upper bound for $\tau_{C,i}$ valid for all *i*, follows from (3.8):

$$\tau_{C,i} < \kappa_{\max} = \frac{1}{\min_{i \neq j} \mu_{ij}} = \frac{1}{\mu_{\min}} = \frac{1}{\min_{i \neq j} \sum_{r=1}^{m} p_{ir} p_{jr}}. \tag{3.9}$$

Corollary 1.2: *If the underlying Markov chain consists of independent trials, i.e. the transition probabilities* $p_{ij} = p_j$, *then for all i, j,*

$$\kappa_{ij}^{(C)} = \frac{1}{\sum_{r=1}^{m} p_r^2}. \tag{3.10}$$

Proof: Observe that $\mu_{ij} = \sum_{r=1}^{m} p_r^2 \equiv \mu$. Thus $\min_{i \neq j} \mu_{ij} = \max_{i \neq j} \mu_{ij} = \mu$ and from (3.2) we deduce $\dfrac{1}{\mu} = \kappa_{\min} \le \kappa_{ij}^{(C)} \le \kappa_{\max} = \dfrac{1}{\mu}$ leading to Eqn.(3.10).

Expression (3.10) can also be derived directly in this special case by solving Equations (3.1) (see also Eqn. (5.7) of [2]). □

In [2] it was shown that, under the condition of independent trials,

$$\tau_{C,i} = \frac{\sum_{j \neq i} p_j}{\sum_{k=1}^{m} p_k^2} = \frac{1 - p_i}{\sum_{k=1}^{m} p_k^2}.$$

Since $1 - 2\sum_{r<s} p_r p_s = 1 - [(\sum_{k=1}^{m} p_k)^2 - (\sum_{k=1}^{m} p_k^2)] = \sum_{k=1}^{m} p_k^2$,

$$\tau_{C,i} = \frac{1 - p_i}{\sum_{k=1}^{m} p_k^2} = \frac{1 - \pi_i}{1 - 2\sum_{r<s} p_r p_s}.$$

Thus the bounds given by Corollary 1.1. are tight under independence assumptions. The interval (κ_{min}, κ_{max}), or its width $\kappa_{\max} - \kappa_{\min}$, could be used as a measure of the departure of the underlying MC from independence.

If expression (3.9) is used when the conditions of Corollary 1.1 are violated, the upper bound grossly overestimates the maximum value of $\tau_{C,i}$. In those chains, if at least one $\mu_{ij} = 0$, the upper bound will be ∞. This will occur in those examples where $p_{ij} = 1$ for some pair (i,j), with $i \neq j$, and $p_{rj} = 0$ for some $r \neq i$.

Since there are instances when some of the μ_{ij} could be zero, it is necessary to explore these cases in more detail. We consider the reduced number of linear equations alluded to in Section 2 above.

Define, for all $i \neq j$ and $r \neq s$,

$$\begin{aligned}\alpha_{i,j}^{(r,s)} &= P\left\{(X_{n+1},Y_{n+1}) = (r,s)\middle|(X_n,Y_n) \in \{(i,j),(j,i)\}\right\} \\ &= P\left\{\text{One step transition to state } (r,s) \text{ from either } (i,j) \text{ or } (j,i)\right\} \\ &= p_{ir}p_{js} + p_{jr}p_{is}.\end{aligned}$$

Observe that $\alpha_{i,j}^{(r,s)} = \alpha_{j,i}^{(r,s)} = \alpha_{i,j}^{(s,r)} = \alpha_{j,i}^{(s,r)}$. In each of these situations we shall write the expression in the form $\alpha_{i,j}^{(r,s)}$ with $i < j$ and $r < s$.

Further since for $i \neq j$, $\kappa_{ij}^{(C)} = \kappa_{ji}^{(C)}$ we write the common value as simply κ_{ij} with $i < j$.

Thus from (3.1) above,

$$\begin{aligned}\kappa_{ij}^{(C)} - 1 &= \sum\sum\nolimits_{r\neq s} p_{ir}p_{js}\kappa_{rs}^{(C)} = \sum\sum\nolimits_{r<s} p_{ir}p_{js}\kappa_{rs}^{(C)} + \sum\sum\nolimits_{r>s} p_{ir}p_{js}\kappa_{rs}^{(C)} \\ &= \sum\sum\nolimits_{r<s} p_{ir}p_{js}\kappa_{rs}^{(C)} + \sum\sum\nolimits_{s<r} p_{js}p_{im}\kappa_{sr}^{(C)} = \sum\sum\nolimits_{r<s}(p_{ir}p_{js} + p_{jr}p_{is})\,\kappa_{rs} = \sum\sum\nolimits_{r<s}\alpha_{i,j}^{(r,s)}\kappa_{rs}.\end{aligned}$$

Thus for all $i < j$,

$$\kappa_{ij} - 1 = \sum\sum\nolimits_{r<s}\alpha_{i,j}^{(r,s)}\kappa_{rs}. \tag{3.11}$$

Equation (3.11) is the reduced variant of the linear equations (3.1).

Note that, using Equation (3.4), the parameters $\alpha_{i,j}^{(r,s)}$ have the property that for all $i < j$,

$$\begin{aligned}\sum\nolimits_{r<s}\alpha_{i,j}^{(r,s)} &= \sum\nolimits_{r<s}(p_{ir}p_{js} + p_{jr}p_{is}) = \sum\nolimits_{r<s}p_{ir}p_{js} + \sum\nolimits_{s>r}p_{is}p_{jr} \\ &= \sum\nolimits_{r<s}p_{ir}p_{js} + \sum\nolimits_{r>s}p_{ir}p_{js} = \sum\nolimits_{r\neq s}p_{ir}p_{js} = 1 - \sum\nolimits_{r}p_{ir}p_{jr} \\ &= 1 - \mu_{ij}.\end{aligned} \tag{3.12}$$

Theorem 2. *Without loss of generality, assume $a < b$ and $i < j$. If $\mu_{ab} = 0$ and $\mu_{ij} > 0$ for all $(i, j) \neq (a, b)$, then for all $(i, j) \neq (a, b)$,*

$$\kappa_{\min} \le \kappa_{ij} \le \kappa_{\max}, \tag{3.13}$$

with
$$\frac{1}{1-\alpha_{a,b}^{(a,b)}} + \kappa_{\min} \le \kappa_{ab} \le \frac{1}{1-\alpha_{a,b}^{(a,b)}} + \kappa_{\max}, \tag{3.14}$$

where
$$\kappa_{\min} = \min_{i<j,(i,j)\ne(a,b)}\left[\frac{\lambda_{i,j}^{(a,b)}}{\mu_{ij}}\right] \text{ and } \kappa_{\max} = \max_{i<j,(i,j)\ne(a,b)}\left[\frac{\lambda_{i,j}^{(a,b)}}{\mu_{ij}}\right], \tag{3.15}$$

with
$$\lambda_{i,j}^{(a,b)} = 1 + \frac{\alpha_{i,j}^{(a,b)}}{1-\alpha_{a,b}^{(a,b)}}. \tag{3.16}$$

Proof: From the reduced equations (3.11), with $a < b$ and $i < j$,

$$\kappa_{ab} - 1 = \alpha_{a,b}^{(a,b)}\kappa_{ab} + \sum\sum\nolimits_{r<s,(r,s)\ne(a,b)} \alpha_{a,b}^{(r,s)}\kappa_{rs}$$

implying $\kappa_{ab}(1-\alpha_{a,b}^{(a,b)}) = 1 + \sum\sum\nolimits_{r<s,(r,s)\ne(a,b)} \alpha_{a,b}^{(r,s)}\kappa_{rs}$.
From (3.12) $\sum\nolimits_{r<s}\alpha_{a,b}^{(r,s)} = 1-\mu_{ab} = 1,$ so that $\sum\nolimits_{r<s,(r,s)\ne(a,b)}\alpha_{a,b}^{(r,s)} = 1-\alpha_{a,b}^{(a,b)}$.
Assuming (3.13), i.e. $\kappa_{\min} \le \kappa_{ij} \le \kappa_{\max}$,

$$1+\left\{1-\alpha_{a,b}^{(a,b)}\right\}\kappa_{\min} \le 1+\sum\sum\nolimits_{r<s,(r,s)\ne(a,b)}\alpha_{a,b}^{(r,s)}\kappa_{rs} = \kappa_{ab}(1-\alpha_{a,b}^{(a,b)}) \le 1+\left\{1-\alpha_{a,b}^{(a,b)}\right\}\kappa_{\max}$$

and result (3.14) follows.
The Theorem will follow once we establish the values for the bounds (3.15).

For $i < j$, from equations (3.11),

$$\kappa_{ij} - 1 - \alpha_{i,j}^{(a,b)}\kappa_{ab} = \sum\sum\nolimits_{r<s,(r,s)\ne(a,b)}\alpha_{i,j}^{(r,s)}\kappa_{rs}. \tag{3.17}$$

Now from (3.12), $\sum\sum\nolimits_{r<s,(r,s)\ne(a,b)}\alpha_{i,j}^{(r,s)} = 1-\mu_{ij}-\alpha_{i,j}^{(a,b)}$, so that from Eqn.(3.17),

$$(1-\mu_{ij}-\alpha_{i,j}^{(a,b)})\kappa_{\min} \le \kappa_{ij} - 1 - \alpha_{i,j}^{(a,b)}\kappa_{ab} \le (1-\mu_{ij}-\alpha_{i,j}^{(a,b)})\kappa_{\max},$$

or that $(1-\mu_{ij}-\alpha_{i,j}^{(a,b)})\kappa_{\min} + 1 + \alpha_{i,j}^{(a,b)}\kappa_{ab} \le \kappa_{ij} \le (1-\mu_{ij}-\alpha_{i,j}^{(a,b)})\kappa_{\max} + 1 + \alpha_{i,j}^{(a,b)}\kappa_{ab}$.
Using (3.14), the above expression is bounded above and below as

$$(1-\mu_{ij}-\alpha_{i,j}^{(a,b)})\kappa_{\min} + 1 + \alpha_{i,j}^{(a,b)}\left\{\frac{1}{1-\alpha_{a,b}^{(a,b)}}+\kappa_{\min}\right\} \le \kappa_{ij}$$

$$\le (1-\mu_{ij}-\alpha_{i,j}^{(a,b)})\kappa_{\max} + 1 + \alpha_{i,j}^{(a,b)}\left\{\frac{1}{1-\alpha_{a,b}^{(a,b)}}+\kappa_{\max}\right\}$$

which simplifies, using Eqn. (3.16), to $(1-\mu_{ij})\kappa_{\min} + \lambda_{i,j}^{(a,b)} \le \kappa_{ij} \le (1-\mu_{ij})\kappa_{\max} + \lambda_{i,j}^{(a,b)}$.
Since we require the lower and upper quantities of the above expression to be bounded below by $\kappa_{\min}$ and above by $\kappa_{\max}$, respectively, we further require, for all $i < j$, $\kappa_{\min} \le (1-\mu_{ij})\kappa_{\min} + \lambda_{i,j}^{(a,b)}$ and and $(1-\mu_{ij})\kappa_{\max} + \lambda_{i,j}^{(a,b)} \le \kappa_{\max}$ implying $\mu_{ij}\kappa_{\min} \le \lambda_{ij}^{(a,b)}$ and $\lambda_{i,j}^{(a,b)} \le \mu_{ij}\kappa_{\max}$ leading to expressions (3.15). □

Theorem 2 requires $\mu_{ab} = \sum\nolimits_{r=1}^{m} p_{ar}p_{br} = 0$. This implies that $p_{ar}p_{br} = 0$ for all r.
In particular $p_{aa}p_{ba} = 0$ and $p_{ab}p_{bb} = 0$. Thus there are four possible cases:

$(i)\ p_{aa}=0$ and $p_{bb}=0, (ii)\ p_{aa}=0$ and $p_{ab}=0, (iii)\ p_{ba}=0$ and $p_{bb}=0, (iv)\ p_{ba}=0$ and $p_{ab}=0.$

These conditions will place restrictions, in particular, on $\alpha_{a,b}^{(a,b)} = p_{aa}p_{bb} + p_{ba}p_{ab}$.
For the respective cases: $(i)\ \alpha_{a,b}^{(a,b)} = p_{ba}p_{ab}$, $(ii)\ \alpha_{a,b}^{(a,b)} = 0$, $(iii)\ \alpha_{a,b}^{(a,b)} = 0$, $(iv)\ \alpha_{a,b}^{(a,b)} = p_{aa}p_{bb}$.
A simplification of Eqn. (3.15) and (3.16) can now be carried out for each of these special cases.

Let us extend Theorem 2 to the situation where we have two distinct pairs of states *(a, b)* and *(c, d),* where $\mu_{ab}=0$ and $\mu_{cd}=0$. Without loss of generality, we may assume *a*< *b* and *c* < *d*.

Theorem 3. *Without loss of generality, assume a < b, c < d (with (a, b) ≠ (c, d)) and i < j. If* $\mu_{ab}=0$, $\mu_{cd}=0$ *and* $\mu_{ij}>0$ *for all (i, j) ≠ (a, b) and (c, d), then for all (i, j) ≠ (a, b), (c, d),*

$$\kappa_{\min} \le \kappa_{ij} \le \kappa_{\max}, \tag{3.18}$$

with

$$\frac{1+\alpha_{a,b}^{(c,d)}-\alpha_{c,d}^{(c,d)}}{\tau_2}+\kappa_{\min} \le \kappa_{ab} \le \frac{1+\alpha_{a,b}^{(c,d)}-\alpha_{c,d}^{(c,d)}}{\tau_2}+\kappa_{\max}, \tag{3.19}$$

$$\frac{1+\alpha_{c,d}^{(a,b)}-\alpha_{a,b}^{(a,b)}}{\tau_2}+\kappa_{\min} \le \kappa_{cd} \le \frac{1+\alpha_{c,d}^{(a,b)}-\alpha_{a,b}^{(a,b)}}{\tau_2}+\kappa_{\max}, \tag{3.20}$$

where

$$\kappa_{\min} = \min_{i<j,(i,j)\ne(a,b),(c,d)}\left[\frac{\lambda_{i,j}^{(a,b;c,d)}}{\mu_{ij}}\right], \quad \text{and } \kappa_{\max} = \max_{i<j,(i,j)\ne(a,b),(c,d)}\left[\frac{\lambda_{i,j}^{(a,b;c,d)}}{\mu_{ij}}\right], \tag{3.21}$$

with

$$\lambda_{i,j}^{(a,b;c,d)} = 1+\frac{\alpha_{i,j}^{(a,b)}(1+\alpha_{a,b}^{(c,d)}-\alpha_{c,d}^{(c,d)})+\alpha_{i,j}^{(c,d)}(1+\alpha_{c,d}^{(a,b)}-\alpha_{a,b}^{(a,b)})}{\tau_2}, \tag{3.22}$$

where

$$\tau_2 = (1-\alpha_{a,b}^{(a,b)})(1-\alpha_{c,d}^{(c,d)})-\alpha_{a,b}^{(c,d)}\alpha_{c,d}^{(a,b)}. \tag{3.23}$$

Proof: From the reduced equations (3.11), for distinct pairs (*a, b*), (*c, d*) and (*i, j*) with *a < b, c < d* and *i < j*,

$$\kappa_{ab} = 1+\alpha_{a,b}^{(a,b)}\kappa_{ab}+\alpha_{a,b}^{(c,d)}\kappa_{cd}+\Delta_{ab}, \tag{3.24}$$

$$\kappa_{cd} = 1+\alpha_{c,d}^{(a,b)}\kappa_{ab}+\alpha_{c,d}^{(c,d)}\kappa_{cd}+\Delta_{cd}, \tag{3.25}$$

$$\kappa_{ij} = 1+\alpha_{i,j}^{(a,b)}\kappa_{ab}+\alpha_{i,j}^{(c,d)}\kappa_{cd}+\Delta_{ij}, \tag{3.26}$$

where, for all (*i,j*), $\Delta_{ij} = \sum\sum_{r<s,(r,s)\ne(a,b),(c,d)}\alpha_{i,j}^{(r,s)}\kappa_{rs}$.
From Eqns.(3.24) and (3.25),

$$B\begin{bmatrix}\kappa_{ab}\\ \kappa_{cd}\end{bmatrix} \equiv \begin{bmatrix}1-\alpha_{a,b}^{(a,b)} & -\alpha_{a,b}^{(c,d)}\\ -\alpha_{c,d}^{(a,b)} & 1-\alpha_{c,d}^{(c,d)}\end{bmatrix}\begin{bmatrix}\kappa_{ab}\\ \kappa_{cd}\end{bmatrix} = \begin{bmatrix}1+\Delta_{ab}\\ 1+\Delta_{cd}\end{bmatrix}.$$

.

Since $\det(B) = \tau_2$, as given by (3.23), taking the inverse of B yields

$$\begin{bmatrix}\kappa_{ab}\\ \kappa_{cd}\end{bmatrix} = B^{-1}\begin{bmatrix}1+\Delta_{ab}\\ 1+\Delta_{cd}\end{bmatrix} = \frac{1}{\tau_2}\begin{bmatrix}1-\alpha_{c,d}^{(c,d)} & \alpha_{a,b}^{(c,d)}\\ \alpha_{c,d}^{(a,b)} & 1-\alpha_{a,b}^{(a,b)}\end{bmatrix}\begin{bmatrix}1+\Delta_{ab}\\ 1+\Delta_{cd}\end{bmatrix},$$

so that

$$\begin{bmatrix}\kappa_{ab}\\ \kappa_{cd}\end{bmatrix} = \frac{1}{\tau_2}\begin{bmatrix}1+\alpha_{a,b}^{(c,d)}-\alpha_{c,d}^{(c,d)}+(1-\alpha_{c,d}^{(c,d)})\Delta_{ab}+\alpha_{a,b}^{(c,d)}\Delta_{cd}\\ 1+\alpha_{c,d}^{(a,b)}-\alpha_{a,b}^{(a,b)}+\alpha_{c,d}^{(a,b)}\Delta_{ab}+(1-\alpha_{a,b}^{(a,b)})\Delta_{cd}\end{bmatrix}. \quad (3.27)$$

Since, for all (i,j), $\sum\sum_{r<s,(r,s)\neq(a,b),(c,d)}\alpha_{i,j}^{(r,s)} = 1-\alpha_{i,j}^{(a,b)}-\alpha_{i,j}^{(c,d)}-\mu_{ij}$, (3.28)

$\sum\sum_{r<s,(r,s)\neq(a,b),(c,d)}\alpha_{a,b}^{(r,s)} = 1-\alpha_{a,b}^{(a,b)}-\alpha_{a,b}^{(c,d)}$, and

$\sum\sum_{r<s,(r,s)\neq(a,b),(c,d)}\alpha_{c,d}^{(r,s)} = 1-\alpha_{c,d}^{(a,b)}-\alpha_{c,d}^{(c,d)}$.

Assuming (3.18), i.e. for $(i, j) \neq (a, b), (c, d)$, $\kappa_{\min} \le \kappa_{ij}^{(C)} \le \kappa_{\max}$, from the definition of Δ_{ij},

$$(1-\alpha_{a,b}^{(a,b)}-\alpha_{a,b}^{(c,d)})\kappa_{\min} \le \Delta_{ab} \le (1-\alpha_{a,b}^{(a,b)}-\alpha_{a,b}^{(c,d)})\kappa_{\max}$$

and

$$(1-\alpha_{c,d}^{(a,b)}-\alpha_{c,d}^{(c,d)})\kappa_{\min} \le \Delta_{cd} \le (1-\alpha_{c,d}^{(a,b)}-\alpha_{c,d}^{(c,d)})\kappa_{\max}.$$

From these above two bounds, the bounds given by Eqns (3.19) and (3.20) now follow upon simplification from Eqns.(3.27).

Now, from Eqn.(3.26), using the upper and lower bounds given by Eqns.(3.19) and (3.20) together with (3.28), it is easily shown, using the definition (3.22) that

$$\lambda_{i,j}^{(a,b;c,d)} + (1-\mu_{ij})\kappa_{\min} \le \kappa_{ij} \le \lambda_{i,j}^{(a,b;c,d)} + (1-\mu_{ij})\kappa_{\max}.$$

Now, since the left hand side and the right hand side of the above equation must be bounded below by $\kappa_{\min}$ and bounded above by $\kappa_{\max}$ respectively, the expressions given by (3.21) now follow. □

Theorem 3 can be extended further to incorporate the situation of multiple pairs of states with zero probability of a one step to coupling. Note that there must be at least one pair of states where a single step takes the chain to a coupling state, since coupling occurs with probability one. (Otherwise, the chain is either an absorbing chain or consists of periodic states.)

Theorem 4. *Suppose* $\mu_{a_ib_i} = 0$, $(a_i < b_i)$ *for i = 1, 2, …, n and* $\mu_{ij} > 0$ *otherwise (with* $n < m(m-1)/2$*). Then for* $(i, j) \notin \{(a_1, b_1), \ldots, (a_n, b_n)\}$,

$$\kappa_{\min} \le \kappa_{ij} \le \kappa_{\max}, \quad (3.29)$$

with, for i = 1, 2, …, n, $$\sum_{j=1}^{n} A_{ij} + \kappa_{\min} \le \kappa_{a_ib_i} \le \sum_{j=1}^{n} A_{ij} + \kappa_{\max}, \quad (3.30)$$

where $$\kappa_{\min} = \min_{i<j,(i,j)\neq(a_1,b_1),\ldots,(a_n,b_n)}\left[\frac{\lambda_{ij}}{\mu_{ij}}\right] \text{ and } \kappa_{\max} = \max_{i<j,(i,j)\neq(a_1,b_1),\ldots,(a_n,b_n)}\left[\frac{\lambda_{ij}}{\mu_{ij}}\right], \quad (3.31)$$

with $$\lambda_{ij} = 1 + \sum_{r=1}^{n}\sum_{s=1}^{n} A_{rs}\alpha_{i,j}^{(a_r,b_r)}, \tag{3.32}$$

and $\left[A_{rs}\right] = (I-A)^{-1}$ *and* A *is the* $n \times n$ *matrix* $A = \left[a_{rs}\right] = \left[\alpha_{a_r,b_r}^{(a_s,b_s)}\right]$.

Proof: From the reduced equations (3.11), for distinct pairs (a_i, b_i), (i =1, 2, …, n) with $a_i, < b_i$

$$\kappa_{a_ib_i} = 1 + \sum_{k=1}^{n} \alpha_{a_i,b_i}^{(a_k,b_k)}\kappa_{a_kb_k} + \Delta_{a_ib_i}, \tag{3.33}$$

and for $i < j$, (with $(i, j) \neq (a_i, b_i)$),

$$\kappa_{ij} = 1 + \sum_{k=1}^{n} \alpha_{i,j}^{(a_k,b_k)}\kappa_{a_kb_k} + \Delta_{ij}, \tag{3.34}$$

where, for all (i, j), $\Delta_{ij} = \sum\sum_{r<s,(r,s)\neq(a_1,b_1),\dots,(a_n,b_n)} \alpha_{i,j}^{(r,s)}\kappa_{rs}$.

Let $\boldsymbol{\kappa}^T = (\kappa_{a_1b_1}, \kappa_{a_2b_2}, \dots, \kappa_{a_nb_n})$ and $\boldsymbol{\Delta}^T = (\Delta_{a_1b_1}, \Delta_{a_2b_2}, \dots, \Delta_{a_nb_n})$. From Eqn.(3.33) $\boldsymbol{\kappa} = \boldsymbol{e} + A\boldsymbol{\kappa} + \boldsymbol{\Delta}$, i.e. $(I-A)\boldsymbol{\kappa} = \boldsymbol{e} + \boldsymbol{\Delta}$ implying $\boldsymbol{\kappa} = (I-A)^{-1}(\boldsymbol{e} + \boldsymbol{\Delta})$.

Now for i = 1, 2, …, n, using Eqn.(3.29),

$$\left(\sum\sum_{r<s,(r,s)\neq(a_1,b_1),\dots,(a_n,b_n)} \alpha_{a_i,b_i}^{(r,s)}\right)\kappa_{\min} \le \Delta_{a_ib_i} \le \left(\sum\sum_{r<s,(r,s)\neq(a_1,b_1),\dots,(a_n,b_n)} \alpha_{a_i,b_i}^{(r,s)}\right)\kappa_{\max}.$$

Since, from (3.12), $\sum\sum_{r<s} \alpha_{i,j}^{(r,s)} = 1 - \mu_{ij}$, it follows, under the conditions of the theorem for i = 1, 2, .., n, that

$$\left(1 - \sum_{k=1}^{n} \alpha_{a_i,b_i}^{(a_k,b_k)}\right)\kappa_{\min} \le \Delta_{a_ib_i} \le \left(1 - \sum_{k=1}^{n} \alpha_{a_i,b_i}^{(a_k,b_k)}\right)\kappa_{\max}.$$

i.e. $$\left(1 - \sum_{k=1}^{n} a_{ik}\right)\kappa_{\min} \le \Delta_{a_ib_i} \le \left(1 - \sum_{k=1}^{n} a_{ik}\right)\kappa_{\max}.$$

Expressing these element-wise inequalities in matrix form yields,

$$\kappa_{\min}(\boldsymbol{e} - A\boldsymbol{e}) \le \boldsymbol{\Delta} \le (\boldsymbol{e} - A\boldsymbol{e})\kappa_{\max},$$

or $$\kappa_{\min}(I-A)\,\boldsymbol{e} \le (I-A)\boldsymbol{\kappa} - \boldsymbol{e} \le (I-A)\,\boldsymbol{e}\,\kappa_{\max}.$$

Now if $\boldsymbol{x}$ is a non-negative vector ($\boldsymbol{x} \ge \boldsymbol{0}$) and B is nonnegative matrix then $B\boldsymbol{x} \ge \boldsymbol{0}$. Note that A is a sub-stochastic matrix (since there is at least one pair of states (c, d) $\notin \{(a_1,b_1), \dots, (a_n,b_n)\}$ with $\alpha_{a_i,b_i}^{(c,d)} > 0$ for at least one i, so that there is at least one row of A with a row-sum less than 1). Consequently A has a maximal eigenvalue less than 1. This implies that $\sum_{k=o}^{\infty} A^k = (I-A)^{-1}$ with $(I-A)^{-1}$ non-singular. Consequently $(I-A)^{-1} \ge 0$, (see [4, Theorem 4.6.6]), leading to

$$(I-A)^{-1}\boldsymbol{e} + \kappa_{\min}\boldsymbol{e} \le \boldsymbol{\kappa} \le (I-A)^{-1}\boldsymbol{e} + \kappa_{\max}\boldsymbol{e},$$

which leads, in element form, to Eqn.(3.30).

Now $\sum\sum_{r<s,(r,s)\neq(a_1,b_1),\ldots,(a_n,b_n)} \alpha_{i,j}^{(r,s)} = 1 - \sum_{r=1}^{n} \alpha_{i,j}^{(a_r,b_r)} - \mu_{ij}$ so that for $(i,j) \notin \{(a_1,b_1),\ldots,(a_n,b_n)\}$

$$\left(1 - \sum_{r=1}^{n} \alpha_{i,j}^{(a_r,b_r)} - \mu_{ij}\right)\kappa_{\min} \le \Delta_{ij} \le \left(1 - \sum_{r=1}^{n} \alpha_{i,j}^{(a_r,b_r)} - \mu_{ij}\right)\kappa_{\max}.$$

From Eqn.(3.34), for $(i,j) \notin \{(a_1,b_1), \ldots, (a_n,b_n)\}$,

$$1 + \sum_{r=1}^{n} \alpha_{i,j}^{(a_r,b_r)} \kappa_{a_r b_r} + \left(1 - \sum_{r=1}^{n} \alpha_{i,j}^{(a_r,b_r)} - \mu_{ij}\right)\kappa_{\min} \le \kappa_{ij} \le 1 + \sum_{r=1}^{n} \alpha_{i,j}^{(a_r,b_r)} \kappa_{a_r b_r} + \left(1 - \sum_{r=1}^{n} \alpha_{i,j}^{(a_r,b_r)} - \mu_{ij}\right)\kappa_{\max}.$$

Now, from Eqn.(3.30),

$$1 + \sum_{r=1}^{n} \alpha_{i,j}^{(a_r,b_r)} \left(\sum_{s=1}^{n} A_{rs} + \kappa_{\min}\right) + \left(1 - \sum_{r=1}^{n} \alpha_{i,j}^{(a_r,b_r)} - \mu_{ij}\right)\kappa_{\min} \le \kappa_{ij},$$

and

$$\kappa_{ij} \le 1 + \sum_{r=1}^{n} \alpha_{i,j}^{(a_r,b_r)} \left(\sum_{s=1}^{n} A_{rs} + \kappa_{\max}\right) + \left(1 - \sum_{r=1}^{n} \alpha_{i,j}^{(a_r,b_r)} - \mu_{ij}\right)\kappa_{\max}.$$

From Eqn.(3.29) we require, for the lower bound,

$$\kappa_{\min} \le 1 + \left(\sum_{r=1}^{n} \alpha_{i,j}^{(a_r,b_r)}\right)\left(\sum_{s=1}^{n} A_{rs} + \kappa_{\min}\right) + \left(1 - \sum_{r=1}^{n} \alpha_{i,j}^{(a_r,b_r)} - \mu_{ij}\right)\kappa_{\min},$$

implying, for all $(i, j) \notin \{(a_1, b_1), \ldots, (a_n, b_n)\}$, that $\mu_{ij}\kappa_{\min} \le 1 + \sum_{r=1}^{n}\sum_{s=1}^{n} A_{rs}\alpha_{i,j}^{(a_r,b_r)} \equiv \lambda_{ij}$, leading to the first bound in (3.31) and expression (3.22).
Similarly for the upper bound we require, for all $(i, j) \notin \{(a_1, b_1), \ldots, (a_n, b_n)\}$, $\lambda_{ij} = 1 + \sum_{r=1}^{n}\sum_{s=1}^{n} A_{rs}\alpha_{i,j}^{(a_r,b_r)} \le \mu_{ij}\kappa_{\max}$ leading to the second bound in (3.31). □

Note that Theorem 2 follows from Theorem 4 when $n = 1$ with $(a_1, b_1) = (a, b)$ where $A = \left[a_{11}\right] = \left[\alpha_{a_1,b_1}^{(a_1,b_1)}\right] = \left[\alpha_{a,b}^{(a,b)}\right]$, $\left[A_{11}\right] = (I - A)^{-1} = (1 - \alpha_{a,b}^{(a,b)})^{-1}$.

Similarly, Theorem 3 follows from Theorem 4 when $n = 2$ with $(a_1, b_1) = (a, b)$, $(a_2, b_2) = (c, d)$ where $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \alpha_{a_1,b_1}^{(a_1,b_1)} & \alpha_{a_1,b_1}^{(a_2,b_2)} \\ \alpha_{a_2,b_2}^{(a_1,b_1)} & \alpha_{a_2,b_2}^{(a_2,b_2)} \end{bmatrix} = \begin{bmatrix} \alpha_{a,b}^{(a,b)} & \alpha_{a,b}^{(c,d)} \\ \alpha_{c,d}^{(a,b)} & \alpha_{c,d}^{(c,d)} \end{bmatrix}$

and $\left[A_{rs}\right] = (I - A)^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \dfrac{1}{\tau_2}\begin{bmatrix} 1 - \alpha_{c,d}^{(c,d)} & \alpha_{a,b}^{(c,d)} \\ \alpha_{c,d}^{(a,b)} & 1 - \alpha_{a,b}^{(a,b)} \end{bmatrix}$ with $\tau_2 = \det(I - A) = (1 - \alpha_{a,b}^{(a,b)})(1 - \alpha_{c,d}^{(c,d)}) - \alpha_{a,b}^{(c,d)}\alpha_{c,d}^{(a,b)}$.

4. Special cases

Example 1. *Two-state Markov chains*

Let $P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$, $(0 < a \leq 1, 0 < b \leq 1)$, be the transition matrix of a two-state Markov chain with state space $S = \{1, 2\}$. Let $d = 1 - a - b$.

If $-1 < d < 1$, the Markov chain is regular with a unique stationary distribution given by

$$\pi_1 = \frac{b}{a+b},\ \pi_2 = \frac{a}{a+b}.$$

Note that $\mu_{12} = \mu_{21} = p_{11}p_{21} + p_{12}p_{22} = (1-a)b + a(1-b) = a + b - 2ab \equiv \mu.$

.

Note that $\mu \neq 0$, since if $\mu = 0$ then $a(1-b) + b(1-a) = 0$. i.e. $a(1-b) = 0$ and $(1-a)b = 0$. Thus either (*i*) $a = 0$ and $b = 0$ or (*ii*) $a = 1$ and $b = 1$. Case (*i*) is impossible since this implies both states are absorbing, while case 2 implies the chain is periodic period 2. In both cases coupling never occurs.

In this special case, expressions for the expected number of trials to coupling can be found explicitly since the solution of equations (2.3) for, $(I - Q)\boldsymbol{\kappa}^{(C)} = \boldsymbol{e}$, is easily effected with

$$\kappa_{12}^{(C)} = \kappa_{21}^{(C)} = \frac{1}{(a+b-2ab)} = \frac{1}{\mu} = \kappa_{\min} = \kappa_{\max}.$$

Further it was shown in [2] that

$$\tau_{C,1} = \frac{a}{(a+b)(a+b-2ab)},\ \text{implying}\ \tau_{C,1} = \frac{\pi_2}{\mu} = (1-\pi_1)\kappa_{\min} = (1-\pi_1)\kappa_{\max}$$

and $$\tau_{C,2} = \frac{b}{(a+b)(a+b-2ab)} = \frac{\pi_1}{\mu} = (1-\pi_2)\kappa_{\min} = (1-\pi_2)\kappa_{\max}.$$

Thus the inequalities (3.2) and (3.8) are in fact equalities, with $(1-\pi_i)\kappa_{\min} = \tau_{C,i} = (1-\pi_i)\kappa_{\max}$.

Example 2. *Three-state Markov chains (Explicit solutions of the* κ_{ij} *).*

Let $P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} 1-b-c & b & c \\ d & 1-d-f & f \\ g & h & 1-g-h \end{bmatrix}$ be the transition matrix of a Markov chain with state space $S = \{1, 2, 3\}$.

Note that $0 < b + c \leq 1$, $0 < d + f \leq 1$ and $0 < g + h \leq 1$. Let

$$\Delta_1 = p_{23}p_{31} + p_{21}p_{32} + p_{21}p_{31} = fg + dh + dg,$$
$$\Delta_2 = p_{31}p_{12} + p_{32}p_{13} + p_{32}p_{12} = gb + hc + hb,$$
$$\Delta_3 = p_{12}p_{23} + p_{13}p_{21} + p_{13}p_{23} = bf + cd + cf,$$
$$\Delta = \Delta_1 + \Delta_2 + \Delta_3 = fg + dh + dg + gb + hc + hb + bf + cd + cf.$$

The Markov chain, with the above transition matrix, is irreducible (and hence a stationary distribution exists) if and only if $\Delta_1 > 0$, $\Delta_2 > 0$, $\Delta_3 > 0$, with stationary probability vector

$$(\pi_1, \pi_2, \pi_3) = \frac{1}{\Delta}(\Delta_1, \Delta_2, \Delta_3). \tag{4.1}$$

Observe that

$$\mu_{12} = \mu_{21} = p_{11}p_{21} + p_{12}p_{22} + p_{13}p_{23} = (1-b-c)d + b(1-d-f) + cf = b + d - 2bd - cd - bf + cf,$$
$$\mu_{23} = \mu_{32} = p_{21}p_{31} + p_{22}p_{32} + p_{23}p_{33} = dg + (1-d-f)h + f(1-g-h) = h + f - 2fh - dh - fg + dg$$
$$\mu_{13} = \mu_{31} = p_{31}p_{11} + p_{32}p_{12} + p_{33}p_{13} = g(1-b-c) + hb + (1-g-h)c = c + g - 2cg - bg - ch + bh.$$

Using the reduced equations (3.11) with just three parameters κ_{12}, κ_{13}, and κ_{23} yields

$$\kappa_{12} = 1 + \alpha_{1,2}^{(1,2)}\kappa_{12} + \alpha_{1,2}^{(1,3)}\kappa_{13} + \alpha_{1,2}^{(2,3)}\kappa_{23}$$
$$\kappa_{13} = 1 + \alpha_{1,3}^{(1,2)}\kappa_{12} + \alpha_{1,3}^{(1,3)}\kappa_{13} + \alpha_{1,3}^{(2,3)}\kappa_{23}$$
$$\kappa_{23} = 1 + \alpha_{2,3}^{(1,2)}\kappa_{12} + \alpha_{2,3}^{(1,3)}\kappa_{13} + \alpha_{2,3}^{(2,3)}\kappa_{23}$$

where $\alpha_{i,j}^{(r,s)} = p_{ir}p_{js} + p_{jr}p_{is}$. In matrix form,

$$\begin{bmatrix} 1-\alpha_{1,2}^{(1,2)} & -\alpha_{1,2}^{(1,3)} & -\alpha_{1,2}^{(2,3)} \\ -\alpha_{1,3}^{(1,2)} & 1-\alpha_{1,3}^{(1,3)} & -\alpha_{1,3}^{(2,3)} \\ -\alpha_{2,3}^{(1,2)} & -\alpha_{2,3}^{(1,3)} & 1-\alpha_{2,3}^{(2,3)} \end{bmatrix} \begin{bmatrix} \kappa_{12} \\ \kappa_{13} \\ \kappa_{23} \end{bmatrix} = B\boldsymbol{\kappa} = \boldsymbol{e}. \tag{4.2}$$

In [2] we were unable to find compact expressions for the solutions of (4.2) in all cases and special cases were considered. However, the structure exhibited by Eqn.(4.2) now permits a simple solution:

First note that $\boldsymbol{\kappa} = B^{-1}\boldsymbol{e} = \dfrac{1}{\tau_3}\begin{bmatrix} \tau_{11} & \tau_{12} & \tau_{13} \\ \tau_{21} & \tau_{22} & \tau_{23} \\ \tau_{31} & \tau_{32} & \tau_{33} \end{bmatrix}\boldsymbol{e} = \dfrac{1}{\tau_3}\begin{bmatrix} \tau_{11}+\tau_{12}+\tau_{13} \\ \tau_{21}+\tau_{22}+\tau_{23} \\ \tau_{31}+\tau_{32}+\tau_{33} \end{bmatrix}$ (4.3)

where

$$\tau_{11} = (1-\alpha_{1,3}^{(1,3)})(1-\alpha_{2,3}^{(2,3)}) - \alpha_{1,3}^{(2,3)}\alpha_{2,3}^{(1,3)}, \tau_{12} = \alpha_{1,2}^{(1,3)}(1-\alpha_{2,3}^{(2,3)}) + \alpha_{1,2}^{(2,3)}\alpha_{2,3}^{(1,3)},$$
$$\tau_{13} = \alpha_{1,2}^{(1,3)}\alpha_{1,3}^{(2,3)} + \alpha_{1,2}^{(2,3)}(1-\alpha_{1,3}^{(1,3)}), \tau_{21} = \alpha_{1,3}^{(1,2)}(1-\alpha_{2,3}^{(2,3)}) + \alpha_{1,3}^{(2,3)}\alpha_{2,3}^{(1,2)},$$
$$\tau_{22} = (1-\alpha_{1,2}^{(1,2)})(1-\alpha_{2,3}^{(2,3)}) - \alpha_{2,3}^{(1,2)}\alpha_{1,2}^{(2,3)}, \tau_{23} = (1-\alpha_{1,2}^{(1,2)})\alpha_{1,3}^{(2,3)} + \alpha_{1,2}^{(2,3)}\alpha_{1,3}^{(1,2)}$$
$$\tau_{31} = \alpha_{1,3}^{(1,2)}\alpha_{2,3}^{(1,3)} + (1-\alpha_{1,3}^{(1,3)})\alpha_{2,3}^{(1,2)}, \tau_{32} = (1-\alpha_{1,2}^{(1,2)})\alpha_{2,3}^{(1,3)} + \alpha_{1,2}^{(1,3)}\alpha_{2,3}^{(1,2)},$$
$$\tau_{33} = (1-\alpha_{1,2}^{(1,2)})(1-\alpha_{1,3}^{(1,3)}) - \alpha_{1,2}^{(1,3)}\alpha_{1,3}^{(1,2)},$$

and det(B) = τ_3 with the following equivalent forms:

$$\tau_3 = (1-\alpha_{1,2}^{(1,2)})\tau_{11} - \alpha_{1,2}^{(1,3)}\tau_{21} - \alpha_{1,2}^{(2,3)}\tau_{31} = -\alpha_{1,3}^{(1,2)}\tau_{12} + (1-\alpha_{1,3}^{(1,3)})\tau_{22} - \alpha_{1,3}^{(2,3)}\tau_{32},$$
$$= -\alpha_{2,3}^{(1,2)}\tau_{13} - \alpha_{2,3}^{(1,3)}\tau_{23} + (1-\alpha_{2,3}^{(2,3)})\tau_{33}.$$

Using the observations, from Eqns.(3.12), that

$$\alpha_{1,2}^{(1,2)} + \alpha_{1,2}^{(1,3)} + \alpha_{1,2}^{(2,3)} + \mu_{1,2} = 1, \alpha_{1,3}^{(1,2)} + \alpha_{1,3}^{(1,3)} + \alpha_{1,3}^{(2,3)} + \mu_{13} = 1, \alpha_{2,3}^{(1,2)} + \alpha_{2,3}^{(1,3)} + \alpha_{2,3}^{(2,3)} + \mu_{23} = 1, \quad (4.4)$$

it can be shown that τ_3 can be re-expressed as one of the following equivalent forms

$$\tau_3 = \mu_{12}\tau_{11} + \mu_{13}\tau_{12} + \mu_{23}\tau_{13} = \mu_{12}\tau_{21} + \mu_{13}\tau_{22} + \mu_{23}\tau_{23} = \mu_{12}\tau_{31} + \mu_{13}\tau_{32} + \mu_{23}\tau_{33}.$$

Thus from Eqn.(4.3),

$$\kappa_{12} = \frac{\tau_{11}+\tau_{12}+\tau_{13}}{\tau_3},\ \kappa_{13} = \frac{\tau_{21}+\tau_{22}+\tau_{23}}{\tau_3},\ \kappa_{23} = \frac{\tau_{31}+\tau_{32}+\tau_{33}}{\tau_3}. \quad (4.5)$$

Further $\tau_{C,1} = \pi_2\kappa_{12} + \pi_3\kappa_{13},\ \tau_{C,2} = \pi_1\kappa_{12} + \pi_3\kappa_{23},\ \tau_{C,3} = \pi_1\kappa_{13} + \pi_2\kappa_{23}$ so that

$$\tau_{C,1} = \frac{\Delta_2\kappa_{12} + \Delta_3\kappa_{13}}{\Delta},\ \tau_{C,2} = \frac{\Delta_1\kappa_{12} + \Delta_3\kappa_{23}}{\Delta},\ \tau_{C,3} = \frac{\Delta_1\kappa_{13} + \Delta_2\kappa_{23}}{\Delta} \text{ implying}$$

$$\tau_{C,1} = \frac{\Delta_2(\tau_{11}+\tau_{12}+\tau_{13}) + \Delta_3(\tau_{21}+\tau_{22}+\tau_{23})}{\Delta\tau_3},$$
$$\tau_{C,2} = \frac{\Delta_1(\tau_{11}+\tau_{12}+\tau_{13}) + \Delta_3(\tau_{31}+\tau_{32}+\tau_{33})}{\Delta\tau_3},$$
$$\tau_{C,3} = \frac{\Delta_1(\tau_{21}+\tau_{22}+\tau_{23}) + \Delta_2(\tau_{31}+\tau_{32}+\tau_{33})}{\Delta\tau_3}. \quad \square$$

We now explore the derivation of simple bounds for κ_{ij} utilising Theorems 1, 2 and 3 for the special cases considered in [2] where coupling occurred. We initially restrict attention to the cases where all the μ_{ij} are positive (Example 3). Other cases when $\mu_{12} = 0,\ \mu_{13} > 0,\ \mu_{23} > 0,$ (Example 4) and $\mu_{12} = 0,\ \mu_{13} = 0,\ \mu_{23} > 0,$ (Example 5) follow after Example 3.

Example 3. *Three-state Markov chains (with all μ_{ij} positive.).*
First observe that in Example 2, *Case* 1 (when $p_{12} = p_{23} = p_{31} = 1$) and *Case* 2 (when $p_{12} = p_{32} = 1$, $p_{21} + p_{23} = 1$) each involve a periodic Markov chain (period 3 for Case 1 and period 2 for Case 2). In Case 1 coupling either occurs initially or never occurs. In Case 2 coupling either occurs initially, after one step, or never occurs. For coupling to occur with probability one we need to restrict attention to regular (irreducible, aperiodic, finite) Markov chains. Thus we omit further consideration of these two cases.

Case 3: "Constant movement" with $p_{11} = p_{22} = p_{33} = 0$.

The transition matrix $P = \begin{bmatrix} 0 & b & 1-b \\ 1-f & 0 & f \\ g & 1-g & 0 \end{bmatrix} = \begin{bmatrix} 0 & p_{12} & p_{13} \\ p_{21} & 0 & p_{23} \\ p_{31} & p_{32} & 0 \end{bmatrix},$

with $0 < b < 1$, $0 < f < 1$, $0 < g < 1$. It is easily seen that $\mu_{12} = p_{13}p_{23} = (1-b)f$, $\mu_{23} = p_{21}p_{31} = (1-f)g$, and $\mu_{13} = p_{32}p_{12} = b(1-g)$. Under the stated conditions, all of these parameters are positive so that the conditions of Theorem 1 are satisfied.
With $\mu_{\min} = \min\{(1-b)f,(1-f)g,b(1-g)\}$, and $\mu_{\max} = \max\{(1-b)f,(1-f)g,b(1-g)\}$, Theorem 1 leads to $\kappa_{\min} = \dfrac{1}{\mu_{\max}} \le \kappa_{ij} \le \kappa_{\max} = \dfrac{1}{\mu_{\min}}$.

Since
$$\Delta_1 \equiv p_{23}p_{31} + p_{21}p_{32} + p_{21}p_{31} = fg + 1 - f = 1 - f(1-g),$$
$$\Delta_2 \equiv p_{31}p_{12} + p_{32}p_{13} + p_{32}p_{12} = gb + 1 - g = 1 - g(1-b),$$
$$\Delta_3 \equiv p_{12}p_{23} + p_{13}p_{21} + p_{13}p_{23} = bf + 1 - b = 1 - b(1-f),$$
$$\Delta \equiv \Delta_1 + \Delta_2 + \Delta_3 = 3 - f(1-g) - g(1-b) - b(1-f).$$

Using (4.1), the stationary probabilities can be derived. Bounds on the expected coupling times follow from application of Eqn. (3.8) yielding

$$\frac{2-g(1-b)-b(1-f)}{[3-f(1-g)-g(1-b)-b(1-f)]\mu_{\max}} \le \tau_{C,1} \le \frac{2-g(1-b)-b(1-f)}{[3-f(1-g)-g(1-b)-b(1-f)]\mu_{\min}},$$
$$\frac{2-f(1-g)-b(1-f)}{[3-f(1-g)-g(1-b)-b(1-f)]\mu_{\max}} \le \tau_{C,2} \le \frac{2-f(1-g)-b(1-f)}{[3-f(1-g)-g(1-b)-b(1-f)]\mu_{\min}},$$
$$\frac{2-f(1-g)-g(1-b)}{[3-f(1-g)-g(1-b)-b(1-f)]\mu_{\max}} \le \tau_{C,3} \le \frac{2-f(1-g)-g(1-b)}{[3-f(1-g)-g(1-b)-b(1-f)]\mu_{\min}}.$$

Computation of $\tau_{C,i}$, for all values of the parameters in [2] showed that

$$2.6667 \le \min_{1\le i\le 3} \tau_{C,i} < \infty .$$

For all combinations of $b = f = g$, the ratios $r_{L,i} = \dfrac{\text{lower bound of } \tau_{C,i}}{\tau_{C,i}}$ and

$r_{U,i} = \dfrac{\text{upper bound of } \tau_{C,i}}{\tau_{C,i}}$ are both equal to 1, leading to the result that lower bound of $\tau_{C,i} = \tau_{C,i} =$ upper bound of $\tau_{C,i}$. This is not equivalent to the independence condition implied under Corollary 1.2 but arises due to the symmetry of the transition matrix in each situation, with the stationary probabilities all equal to 1/3.

Taking all combinations of *b, f*, and *g* in steps of 0.1 between 0.1 and 0.9 we achieve considerable variability between the ratios.

In particular, $0.097 \le r_{L,i} \le 1$ with the minimal ratio being achieved at *(b, c, f)* = (0.1, 0.1, 0.9) and (0.9, 0.1, 0.9) for $r_{L,1}$, at (0.9, 0.1, 0.1) and (0.9, 0.9, 0.1) for $r_{L,2}$, and at (0.1, 0.9, 0.1) and (0.1, 0.9, 0.9) for $r_{L,3}$.

Further, $1 \le r_{U,i} \le 14.063$ with the maximal ratio being achieved at *(b, c, f)* = (0.5, 0.9, 0.1), for for $r_{U,1}$, at (0.1, 0.5, 0.9) for $r_{U,2}$, and at (0.9, 0.1, 0.5) for $r_{U,3}$.

Case 4: "*Independent trials*"

For this case $P = \begin{bmatrix} p_1 & p_2 & p_3 \\ p_1 & p_2 & p_3 \\ p_1 & p_2 & p_3 \end{bmatrix}$, so that $p_{ij} = p_j$ for all i, j implying that the Markov chain is equivalent to independent trials on the state space S = {1, 2, 3}.

For all $i \neq j$, $\mu_{ij} = p_1^{\;2} + p_2^{\;2} + p_3^{\;2} = 1 - 2p_1p_2 - 2p_2p_3 - 2p_3p_1$.

Now $\Delta_1 = p_1$, $\Delta_2 = p_2$, $\Delta_3 = p_3$, $\Delta = p_1 + p_2 + p_3 = 1$, implying $\pi_1 = p_1,\ \pi_2 = p_2,\ \pi_3 = p_3$.

For all i, it was shown in [2] that $\tau_{C,i} = \dfrac{1-p_i}{1-2p_1p_2-2p_1p_3-2p_2p_3} = \dfrac{1-\pi_i}{\mu_{\min}} = \dfrac{1-\pi_i}{\mu_{\max}}$.

Thus each inequality in (3.8) is in fact an equality, with the upper and lower bounds coinciding, as observed in Corollary 1.2.

Case 5: "Cyclic drift" $p_{13} = p_{21} = p_{32} = 0$ with

$$P = \begin{bmatrix} p_{11} & p_{12} & 0 \\ 0 & p_{22} & p_{23} \\ p_{31} & 0 & p_{33} \end{bmatrix} = \begin{bmatrix} 1-b & b & 0 \\ 0 & 1-f & f \\ g & 0 & 1-g \end{bmatrix}.$$

For this case $\mu_{12} = p_{12}p_{22} = b(1-f)$, $\mu_{23} = p_{23}p_{33} = f(1-g)$, $\mu_{13} = p_{31}p_{11} = g(1-b)$, with $\mu_{\min} = \min\{b(1-f), f(1-g), g(1-b)\}$ and $\mu_{\max} = \max\{b(1-f), f(1-g), g(1-b)\}$.

Thus for $0 < b < 1$, $0 < f < 1$, $0 < g < 1$, all the μ_{ij} parameters are positive and the results of Theorem 1 can be applied.

Further $\Delta_1 = fg$, $\Delta_2 = gb$, $\Delta_3 = bf$, $\Delta = fg + gb + bf$ so that expressions for the stationary probabilities follow from Eqn.(4.1). Using Eqn.(3.8) this leads to the following bounds on the expected times to coupling:

$$\frac{b(g+f)}{[fg+gb+bf]\mu_{\max}} \leq \tau_{C,1} \leq \frac{b(g+f)}{[fg+gb+bf]\mu_{\min}},$$

$$\frac{f(g+b)}{[fg+gb+bf]\mu_{\max}} \leq \tau_{C,2} \leq \frac{f(g+b)}{[fg+gb+bf]\mu_{\min}},$$

$$\frac{g(f+b)}{[fg+gb+bf]\mu_{\max}} \leq \tau_{C,3} \leq \frac{g(f+b)}{[fg+gb+bf]\mu_{\min}}.$$

As for Case 3, we explore the ratios $r_{L,i} = \dfrac{\text{lower bound of } \tau_{C,i}}{\tau_{C,i}}$ and $r_{U,i}$ $= \dfrac{\text{upper bound of } \tau_{C,i}}{\tau_{C,i}}$.

When $b = f = g$, both ratios are equal to 1, leading to the lower bound of $\tau_{C,i} = \tau_{C,i}$ = upper bound of $\tau_{C,i}$. As for Case 3, this is not equivalent to the independence condition implied under Corollary 1.2 but arises due to the symmetry of the transition matrix in each situation with the stationary probabilities all equal to 1/3.

Taking all combinations of *b, f*, and *g* in steps of 0.1 between 0.1 and 0.9 we achieve less variability between the lower ratios $r_{L,i}$, but much more variability between the upper ratios $r_{U,i}$ than was present in Case 3.

In particular, $0.185 \le r_{L,i} \le 1$ with the minimal ratio being achieved at (*b, c, f*) = (0.1, 0.9, 0.1) for $r_{L,1}$, at (0.1, 0.1, 0.9) for $r_{L,2}$, and (0.9, 0.1, 0.1) for $r_{L,3}$.

Further $1 \le r_{U,i} \le 67.69$ with the maximal ratio being achieved at (*b, c, f*) =

(0.9, 0.1, 0.9) for $r_{U,1}$, at (0.9, 0.9, 0.1) for $r_{U,2}$, and (0.1, 0.9, 0.9) for $r_{U,3}$.
From Eqn.(3.9) simple upper bounds, valid for all *i*, can be given as

$$\tau_{C,i} < \frac{1}{\mu_{\min}} = \frac{1}{\min(p_{12}p_{22}, p_{11}p_{31}, p_{23}p_{33})} = \frac{1}{\min(b(1-f), f(1-g), g(1-b))}.$$

Case 6: "Constant probability state selection"

In this case, with $P = \begin{bmatrix} 1-a & \frac{a}{2} & \frac{a}{2} \\ \frac{b}{2} & 1-b & \frac{b}{2} \\ \frac{c}{2} & \frac{c}{2} & 1-c \end{bmatrix}$, $(0 < a \le 1, 0 < b \le 1, 0 < c \le 1.)$

Observe that

$$\mu_{12} = \frac{2(a+b)-3ab}{4},\ \mu_{13} = \frac{2(a+c)-3ac}{4},\ \mu_{23} = \frac{2(b+c)-3bc}{4}$$

with

$$\mu_{\min} = \min\left(\frac{2(a+b)-3ab, 2(b+c)-3bc, 2(a+c)-3ac}{4}\right),$$

$$\mu_{\max} = \max\left(\frac{2(a+b)-3ab, 2(b+c)-3bc, 2(a+c)-3ac}{4}\right).$$

Further $\Delta_1 = \frac{3bc}{4}$, $\Delta_2 = \frac{3ac}{4}$, $\Delta_3 = \frac{3ab}{4}$ and thus $\Delta = \frac{3(bc+ac+ab)}{4}$. This leads to expressions for the stationary probabilities and hence to the following bounds for the $\tau_{C,i}$:

$$\frac{a(b+c)}{[bc+ac+ab]\mu_{\max}} \le \tau_{C,1} \le \frac{a(b+c)}{[bc+ac+ab]\mu_{\min}},$$
$$\frac{b(a+c)}{[bc+ac+ab]\mu_{\max}} \le \tau_{C,2} \le \frac{b(a+c)}{[bc+ac+ab]\mu_{\min}},$$
$$\frac{c(a+b)}{[bc+ac+ab]\mu_{\max}} \le \tau_{C,3} \le \frac{c(a+b)}{[bc+ac+ab]\mu_{\min}}.$$

Paralleling the procedures of cases 3 and 5 we obtain the following observations for the ratios $r_{L,i}$ and $r_{U,i}$. Firstly both $r_{L,i} = r_{U,i} = 1$, implying equality of the lower and upper bounds of $\tau_{C,i}$, and equal to $\tau_{C,i}$ occur at all cases when $a = b = c$, with the stationary probabilities all the same. In this case there is much less variability between the actual values of the expected times to coupling and the associated lower and upper bounds.

In particular it can be shown that for all values of (*a, b, c*) in the ranges 0.1 (0.1) 1.0, $0.277 \le r_{L,i} \le 1$ and $1 \le r_{U,i} \le 2.17$. The lower ratio $r_{L,i} = 0.277$ occurs at the following sets of values of (*a, b, c*): (0.1, 0.1, 1) and (0.1, 1, 0.1) for $r_{L,1}$, (0.1, 0.1, 1) and (1, 0.1, 0.1) for $r_{L,2}$, and (0.1, 1, 0.1) and (1, 0.1, 0.1) for $r_{L,3}$. The upper ratio $r_{U,i} = 2.17$ occurs at (*a, b, c*) = (1, 0.1, 0.1) for $r_{U,1}$, (0.1, 1, 0.1) for $r_{U,2}$, and (0.1, 0.1, 1) for $r_{U,3}$.

These bounds, especially the upper bounds, are much tighter than those exhibited in Cases 3 and 4, highlighting the efficacy of the procedure of Theorem 1 when the transition matrix is a positive matrix.

Example 4. *Three-state Markov chains ($\mu_{12} = 0$, $\mu_{13} > 0$, $\mu_{23} > 0$).*

Let $P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} 1-b-c & b & c \\ d & 1-d-f & f \\ g & h & 1-g-h \end{bmatrix}$ be the transition matrix of a Markov chain with state space $S = \{1, 2, 3\}$.

Note that $0 < b + c \le 1$, $0 < d + f \le 1$ and $0 < g + h \le 1$.

Observe that $\mu_{12} = p_{11}p_{21} + p_{12}p_{22} + p_{13}p_{23} = 0$ implies $p_{11}p_{21} = 0$, $p_{12}p_{22} = 0$, and $p_{13}p_{23} = 0$.

Thus eight cases need to be considered:

(*i*) $p_{11} = 0$, $p_{12} = 0$, and $p_{13} = 0$, (*ii*) $p_{11} = 0$, $p_{12} = 0$, and $p_{23} = 0$,

(*iii*) $p_{11} = 0$, $p_{22} = 0$, and $p_{13} = 0$, (*iv*) $p_{11} = 0$, $p_{22} = 0$, and $p_{23} = 0$,

(*v*) $p_{21} = 0$, $p_{12} = 0$, and $p_{13} = 0$, (*vi*) $p_{21} = 0$, $p_{12} = 0$, and $p_{23} = 0$,

(*vii*) $p_{21} = 0$, $p_{22} = 0$, and $p_{13} = 0$, (*iix*) $p_{21} = 0$, $p_{22} = 0$, and $p_{23} = 0$.

Of these cases (*i*) and (*iix*) are impossible since $p_{11} + p_{12} + p_{13}$ and $p_{21} + p_{22} + p_{23}$ must be 1. Also cases (*v*) and (*vi*) are impossible (since the above restrictions would imply,

respectively, that $p_{11} = 1$ and $p_{22} = 1$ and hence, respectively, that states 1 and 2, are absorbing.)
This leads to four remaining possibilities (with (*ii*), (*iii*), (*iv*), (*vii*) relabelled as (*a*), (*b*), (*c*) (*d*))
(*a*) $p_{11} = 0,\ p_{12} = 0,$ and $p_{23} = 0,$ with $p_{13} = 1,$ (*b*) $p_{11} = 0,\ p_{22} = 0,$ and $p_{13} = 0,$ with $p_{12} = 1,$
(*c*) $p_{11} = 0,\ p_{22} = 0,$ and $p_{23} = 0,$ with $p_{21} = 1,$ (*d*) $p_{21} = 0,\ p_{22} = 0,$ and $p_{13} = 0,$ with $p_{23} = 1.$

For case (*a*): $P_a = \begin{bmatrix} 0 & 0 & 1 \\ d & 1-d & 0 \\ g & h & 1-g-h \end{bmatrix}$, with $\mu_{13} = 1 - g - h > 0,\ \ \mu_{23} = dg + (1-d)h > 0;$
$\alpha_{1,2}^{(1,2)} = 0,\ \alpha_{1,3}^{(1,2)} = 0,\ \alpha_{2,3}^{(1,2)} = dh + (1-d)g,$ and $0 < d \le 1,\ 0 \le g < 1, 0 < h < 1, 0 < g + h < 1.$
For case (*b*): $P_b = \begin{bmatrix} 0 & 1 & 0 \\ d & 0 & 1-d \\ g & h & 1-g-h \end{bmatrix}$, with $\mu_{13} = h > 0,\ \mu_{23} = dg + (1-d)(1-g-h) > 0;$
$\alpha_{1,2}^{(1,2)} = d,\ \alpha_{1,3}^{(1,2)} = g,\ \alpha_{2,3}^{(1,2)} = dh,$ and $0 \le d < 1,\ 0 \le g < 1\ ,\ 0 < h < 1,\ 0 < g + h \le 1.$
For case (*c*): $P_c = \begin{bmatrix} 0 & b & 1-b \\ 1 & 0 & 0 \\ g & h & 1-g-h \end{bmatrix}$, with $\mu_{13} = bh + (1-b)(1-g-h) > 0,\ \mu_{23} = g > 0;$
$\alpha_{1,2}^{(1,2)} = b,\ \alpha_{1,3}^{(1,2)} = bg,\ \alpha_{2,3}^{(1,2)} = h,$ and $0 \le b < 1,\ 0 < g < 1,\ 0 \le h < 1,\ 0 < g + h \le 1.$
For case (*d*): $P_d = \begin{bmatrix} 1-b & b & 0 \\ 0 & 0 & 1 \\ g & h & 1-g-h \end{bmatrix}$ with $\mu_{13} = (1-b)g + bh > 0,\ \mu_{23} = 1 - g - h > 0;$
$\alpha_{1,2}^{(1,2)} = 0,\ \alpha_{1,3}^{(1,2)} = (1-b)h,\ \alpha_{2,3}^{(1,2)} = 0,$ and $0 < b \le 1,\ 0 < g < 1,\ 0 \le h < 1,\ 0 < g + h < 1.$

Note that there is some symmetry between cases (*a*) and (*d*), and between cases (*b*) and (*c*).
Case (*d*) converts to Case (*a*) by relabelling the states {1, 2, 3} as {2, 1, 3} and changing the parameters (*b, g, h*) to (*d, h, g*). This same procedure will also convert Case (*c*) to Case (*b*).
From Theorem 2,

$$\frac{1}{1-\alpha_{1,2}^{(1,2)}} + \kappa_{\min} \le \kappa_{12} \le \frac{1}{1-\alpha_{1,2}^{(1,2)}} + \kappa_{\max},\ \kappa_{\min} \le \kappa_{13} \le \kappa_{\max},\ \kappa_{\min} \le \kappa_{23} \le \kappa_{\max}.$$

$$\kappa_{\min} = \min\left[\frac{1}{\mu_{13}}\left\{1 + \frac{\alpha_{1,3}^{(1,2)}}{1-\alpha_{1,2,}^{(1,2)}}\right\}, \frac{1}{\mu_{23}}\left\{1 + \frac{\alpha_{2,3}^{(1,2)}}{1-\alpha_{1,2}^{(1,2)}}\right\}\right],\ \text{and}$$

$$\kappa_{\max} = \max\left[\frac{1}{\mu_{13}}\left\{1 + \frac{\alpha_{1,3}^{(1,2)}}{1-\alpha_{1,2,}^{(1,2)}}\right\}, \frac{1}{\mu_{23}}\left\{1 + \frac{\alpha_{2,3}^{(1,2)}}{1-\alpha_{1,2}^{(1,2)}}\right\}\right].$$

These expressions, with substitution as above for the special cases, together with explicit calculations for κ_{ij} provided by equations (4.5) lead to the following observations.

For each of the following parameter selections: case (*a*) with (*d, g, h*) = (1, 0.3, 0.5), case (*b*) with (*d, g, h*) = (0, 0.5, 0.3), (0.6, 0.4, 0.4), case (*c*) with (*b, g, h*) = (0, 0.3, 0.5), (0.6, 0.4, 0.4), and case (*d*) with (*b, g, h*) = (1, 0.5, 0.3) the lower bound for each κ_{ij} = upper bound for κ_{ij} = exact value of κ_{ij}, providing an effective way of evaluating κ_{ij}. Further, at each of the above parameter selections, for *i* = 1, 2, 3, the lower bound for each $\tau_{C,i}$ = upper bound for $\tau_{C,i}$ = exact value of $\tau_{C,i}$.

For each (*i, j*) with *i* < *j*, let $s_{L,ij} = \dfrac{lower\ bound\ of\ \kappa_{ij}}{\kappa_{ij}}$ and $s_{U,ij} = \dfrac{upper\ bound\ of\ \kappa_{ij}}{\kappa_{ij}}$,

and for *i* = 1, 2, 3, let $r_{L,i} = \dfrac{lower\ bound\ of\ \tau_{C,i}}{\tau_{C,i}}$ and $r_{U,i} = \dfrac{upper\ bound\ of\ \tau_{C,i}}{\tau_{C,i}}$.

In every case $s_{L,ij} \le 1$, $r_{L,i} \le 1$, $s_{U,ij} \ge 1$ and $r_{U,i} \le 1$.

Minimal extreme values, with the parameters taking increments of 0.1 in the restricted ranges for each case, occur at the following parameter selections:

Case (*a*): $s_{L,12} = 0.305$, $s_{L,23} = 0.173$, $r_{L,2} = r_{L,3} = 0.186$, at (*d, g, h*) = (0.1, 0, 0.1), $s_{L,13} = 0.223$ and $r_{L,1} = 0.234$ at (*d, g, h*) = (1, 0.8, 0.1).

Case (*b*): $s_{L,12} = s_{L,13} = r_{L,1} = r_{L,2} = 0.100$, $s_{L,23} = r_{L,3} = 0.011$ at (*d, g, h*) = (0.9, 0, 0.9).

Case (*c*): $s_{L,12} = s_{L,23} = r_{L,1} = r_{L,3} = 0.100$, $s_{L,13} = r_{L,2} = 0.011$ at (*b, g, h*) = (0.9, 0.9, 0).

Case (*d*): $s_{L,12} = 0.305$, $s_{L,13} = 0.173$, $\tau_{C,1} = r_{L,3} = 0.186$ at (*b, g, h*) = (0.1, 0.1, 0), $s_{L,23} = 0.223$ and $r_{L,2} = 0.234$ at (*b, g, h*) = (1, 0.1, 0.8).

Maximal extreme values, with the parameters have increments of 0.1 in the restricted ranges for each case, occur at the following parameter selections:

Case (*a*): $s_{U,12} = 46.54$, $s_{U,13} = 87.62$, $s_{U,23} = 44.67$, $r_{U,1} = 80.46$, $r_{U,2} = 44.84$, $r_{U,3} = 58.18$ at (*d, g, h*) = (0.9, 0, 0.1).

Case (*b*): $s_{U,12} = r_{U,1} = 82.90$, $s_{U,13} = r_{U,2} = 81.99$, at (*d, g, h*) = (0.9, 0, 0.9),

$s_{U,23} = r_{U,3} = 29.25$ at (*d, g, h*) = (0.9, 0.9, 0.1).

Case (*c*): $s_{U,12} = r_{U,1} = 82.90$, $s_{U,23} = r_{U,3} = 81.99$, at (*b, g, h*) = (0.9, 0.9, 0),

$s_{U,13} = r_{U,2} = 29.25$ at (*b, g, h*) = (0.9, 0.1, 0.9).

Case (*d*): $s_{U,12} = 46.54$, $s_{U,13} = 44.67$, $s_{U,23} = 87.62, r_{U,1} = 44.84$, $r_{U,2} = 80.46$, $r_{U,3} = 58.18$ at (*b, g, h*) = (0.9, 0.1, 0).

These extremal ratios for the lower bound,(resp. the upper bound) are in many instances smaller (resp. larger) that those experienced when the μ_{ij} are all positive.

Example 5. *Three-state Markov chains (μ_{12} = 0, μ_{13} = 0, μ_{23} > 0,)*

Let $P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} 1-b-c & b & c \\ d & 1-d-f & f \\ g & h & 1-g-h \end{bmatrix}$ be the transition matrix of a Markov chain with state space $S = \{1, 2, 3\}$.

Note that $0 < b + c \le 1$, $0 < d + f \le 1$ and $0 < g + h \le 1$.

Consider the four possibilities from the $\mu_{12} = 0$ cases:

(a) $p_{11} = 0,\ p_{12} = 0,$ and $p_{23} = 0,$ with $p_{13} = 1,$

(b) $p_{11} = 0,\ p_{22} = 0,$ and $p_{13} = 0,$ with $p_{12} = 1,$

(c) $p_{11} = 0,\ p_{22} = 0,$ and $p_{23} = 0,$ with $p_{21} = 1,$

(d) $p_{21} = 0,\ p_{22} = 0,$ and $p_{13} = 0,$ with $p_{23} = 1.$

For case (a): $\mu_{13} = 1 - g - h = 0 \Rightarrow h = 1 - g,\ \mu_{23} = dg + (1-d)(1-g) > 0$,

$P_a = \begin{bmatrix} 0 & 0 & 1 \\ d & 1-d & 0 \\ g & 1-g & 0 \end{bmatrix}$ with $0 < d \le 1$, $0 \le g < 1$, $0 < h = 1 - g \le 1$.

For case (b): with $\mu_{13} = 0 \Rightarrow h = 0,\ \mu_{23} = dg + (1-d)(1-g) > 0$,

$P_b = \begin{bmatrix} 0 & 1 & 0 \\ d & 0 & 1-d \\ g & 0 & 1-g \end{bmatrix}$ with $0 \le d < 1$, $0 < g \le 1$, $h = 0$.

For case (c): $\mu_{13} = bh + (1-b)(1-g-h) = 0,\ \mu_{23} = g > 0$, implies $bh = 0$ and $(1 - b)(1 - g - h) = 0$. This leads to four possibilities: $b = 0$ and $b = 1$ (impossible); $b = 0$ and $g + h = 1$; $h = 0$ and $b = 1$ (which is impossible since state 3 is then transient); $h = 0$ and $g = 1$ (which doesn't lead to coupling since the chain is then periodic with period 2). Thus there is only one possibility:

$P_c = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ g & 1-g & 0 \end{bmatrix}$ with $0 < g < 1$, $0 < h = 1 - g < 1$, (which is a special case of (a) with $d = 1$).

For case (d): $\mu_{13} = (1-b)g + bh = 0,\ \mu_{23} = 1 - g - h > 0$ implies $(1 - b)g$ and $bh = 0$. Thus one of four possibilities $b = 1$ and $g = 0$ (impossible since state 1 is then transient); $b = 1$ and $h = 0$; $g = 0$ and $b = 0$ (impossible since state 1 is then absorbing); $g = 0$ and $h = 1$ (which is impossible since state 1 is then transient). Thus there is only one possibility:

$P_d = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ g & 0 & 1-g \end{bmatrix}$ with $b = 1, 0 < g < 1, h = 0$, (which is a special case of (b) with d = 0).

Thus effectively there are only two non trivial cases to consider – case (a) with $0 \le d \le 1$, $0 < g \le 1$ and case (b), with $0 \le d \le 1, 0 < g \le 1$. (The symmetry, as present in Example 4, effectively reduces this to one case.)

In computing the bounds for the special cases above, for κ_{12}, κ_{13} and κ_{23}, using the procedure of Theorem 3, first observe that $\kappa_{\min} = \dfrac{\lambda_{2,3}^{(1,2;1,3)}}{\mu_{23}} = \kappa_{\max}$, leading to

$$\kappa_{23} = \frac{\lambda_{2,3}^{(1,2;1,3)}}{\mu_{23}},\ \kappa_{12} = \frac{1+\alpha_{1,2}^{(1,3)} - \alpha_{1,3}^{(1,3)}}{\tau_2} + \frac{\lambda_{2,3}^{(1,2;1,3)}}{\mu_{23}},\ \kappa_{13} = \frac{1+\alpha_{1,3}^{(1,2)} - \alpha_{1,2}^{(1,2)}}{\tau_2} + \frac{\lambda_{2,3}^{(1,2;1,3)}}{\mu_{23}},$$

where $$\lambda_{2,3}^{(1,2;1,3)} = 1 + \frac{\alpha_{2,3}^{(1,2)}(1+\alpha_{1,2}^{(1,3)} - \alpha_{1,3}^{(1,3)}) + \alpha_{2,3}^{(1,3)}(1+\alpha_{1,3}^{(1,2)} - \alpha_{1,2}^{(1,2)})}{\tau_2},$$

with $$\tau_2 = (1-\alpha_{1,2}^{(1,2)})(1-\alpha_{1,3}^{(1,3)}) - \alpha_{1,2}^{(1,3)}\alpha_{1,3}^{(1,2)}.$$

Simplification using the observations from Eqn. (4.4), that since $\mu_{12} = 0$ and $\mu_{13} = 0$,

$$\alpha_{1,2}^{(1,2)} + \alpha_{1,2}^{(1,3)} + \alpha_{1,2}^{(2,3)} = 1, \alpha_{1,3}^{(1,2)} + \alpha_{1,3}^{(1,3)} + \alpha_{1,3}^{(2,3)} = 1,\ \alpha_{2,3}^{(1,2)} + \alpha_{2,3}^{(1,3)} + \alpha_{2,3}^{(2,3)} + \mu_{23} = 1.$$

Further, in cases (a) and (c):

$$\alpha_{1,2}^{(1,2)} = 0,\ \alpha_{1,2}^{(1,3)} + \alpha_{1,2}^{(2,3)} = 1,\ \alpha_{1,3}^{(1,2)} = 0,\ \alpha_{1,3}^{(1,3)} + \alpha_{1,3}^{(2,3)} = 1,\ \alpha_{2,3}^{(1,3)} = 0,\ \alpha_{2,3}^{(2,3)} = 0,$$

$$\tau_2 = \alpha_{1,3}^{(2,3)},\ \lambda_{23}^{(1,2;1,3)} = \frac{\alpha_{1,3}^{(2,3)} + \alpha_{2,3}^{(1,2)}(1+\alpha_{1,2}^{(1,3)} - \alpha_{1,3}^{(1,3)})}{\alpha_{1,3}^{(2,3)}},$$

while in cases (b) and (d):

$$\alpha_{12}^{(1,3)} = 0,\ \alpha_{12}^{(1,2)} + \alpha_{12}^{(2,3)} = 1,\ \alpha_{13}^{(1,3)} = 0,\ \alpha_{13}^{(1,2)} + \alpha_{13}^{(2,3)} = 1,\ \alpha_{23}^{(1,2)} = 0, \alpha_{23}^{(2,3)} = 0,$$

$$\tau_2 = \alpha_{1,2}^{(2,3)},\ \lambda_{2,3}^{(1,2;1,3)} = \frac{\alpha_{1,2}^{(2,3)} + \alpha_{2,3}^{(1,2)} + \alpha_{2,3}^{(1,3)}(1+\alpha_{1,3}^{(1,2)} - \alpha_{1,2}^{(1,2)})}{\alpha_{1,2}^{(2,3)}}.$$

Thus in this example, all the bounds are exact, with agreement to the explicit solutions of equations (4.1) being obtained as in Example 3. i.e. $\kappa_{ij}(exact) = \kappa_{ij}(bound)$ leading to the ratios

$$r_{L,i} = \frac{lower\ bound\ of\ \tau_{C,i}}{\tau_{C,i}} = \frac{upper\ bound\ of\ \tau_{C,i}}{\tau_{C,i}} = r_{U,i} = 1, \text{ in each case.}$$

The computation procedure of Theorem 3 is thus an alternative procedure for evaluating the κ_{ij} in the case of a three-state chain when any two of the parameters μ_{ab} and μ_{cd} are both zero.

References

[1] D.J. Aldous, J.A. Fill. *Reversible Markov Chains and Random Walks on Graphs* (Book in preparation)
See http://www.stat.Berkeley.EDU/users/aldous/book.html

[2] J. J. Hunter, Coupling and Mixing times in a Markov chain, Res. Lett. Inf. Math. Sci, 11, (2007) 1-22. (Submitted to Linear Algebra Appl.)

[3] J. J. Hunter, Mixing times with applications to perturbed Markov chains, Linear Algebra Appl. 417 (2006) 108-123.

[4] J. J. Hunter, Mathematical Techniques of Applied Probability, Volume 1, Discrete Time Models: Basic Theory, Academic, New York, 1983.

[5] J. J. Hunter, Mathematical Techniques of Applied Probability, Volume 2, Discrete Time Models: Techniques and Applications, Academic, New York, 1983.

[6] L. Lovasz, P. Winkler, Mixing Times, in: D. Aldous, J. Propp (Eds.), Microsurveys in Discrete Probability, DIMACS Series in Discrete Math and Theor. Comp. Sci, AMS, 1998, 85-133.