

Coupling and Mixing Times in a Markov Chains

Jeffrey J. Hunter

I.I.M.S. Massey University, Auckland, New Zealand

j.hunter@massey.ac.nz

Abstract

The derivation of the expected time to coupling in a Markov chain and its relation to the expected time to mixing (as introduced by the author in “Mixing times with applications to perturbed Markov chains” *Linear Algebra Appl.* (417, 108-123 (2006)) are explored. The two-state cases and three-state cases are examined in detail.

1. Introduction

Let $P = [p_{ij}]$ be the transition matrix of a finite irreducible, discrete time Markov chain $\{X_n\}$, ($n \geq 0$), with state space $S = \{1, 2, \dots, m\}$. It is well known that 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 ([4, Theorem 7.1.2]). Let $\boldsymbol{\pi}^T = (\pi_1, \pi_2, \dots, \pi_m)$ be the stationary probability vector of the Markov chain.

Various measures have been proposed regarding the “time to stationary” for a Markov chain, i.e. the number of trials, or steps, that the Markov chain makes until it can be regarded as operating under “stationary conditions”. (See e.g. [1], [9]).

The author considered ([3]) the concept of “time to mixing” as a possible approach, along the following lines. Let Y be a random variable whose probability distribution is the stationary distribution $\{\pi_j\}$. We shall say that the Markov chain $\{X_n\}$, “achieves “mixing”, at time $T = k$, when $X_k = Y$ for the smallest such $k \geq 1$. Thus, we first sample from the stationary distribution $\{\pi_j\}$ to determine a value of the random variable Y , say $Y = j$. We then observe the Markov chain, starting at a given state i and achieve “mixing” at time $T = n$ when $X_n = j$ for the first such $n \geq 1$. i.e., conditional upon $Y = j$, $T = \min\{n \geq 1 \text{ such that } X_n = j \text{ given that } X_0 = i\} = T_{ij}$, the first passage time from state i to state j , (or return to state j when $i = j$).

In [3] we assumed that mixing cannot occur at the initial trial although we could have assumed a mixing time of $T = 0$ when $X_0 = Y$. We consider such a case later when we summarise the key results regarding the expected “time to mixing” in Section 2.

The main thrust of this paper is to explore the concept of “coupling” of Markov chains as a procedure to provide a measure of the time to stationarity. We start a Markov chain $\{Y_n\}$, with the same transition matrix P as for $\{X_n\}$, operating under stationary conditions, so that the initial probability distribution for Y_0 is the stationary distribution $\{\pi_j\}$. We 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 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, with each chain having the same stationary distribution $\{\pi_j\}$. In Section 3 we explore the derivation of the expected “time to coupling”.

In section 4 we compare the two concepts of “time to stationarity” through a variety of special two-state and three-state cases. We show, in particular, that it is difficult to get general inequalities between the expected times to mixing starting in state i , $\tau_{i,M}$, and the expected time to coupling starting in state i , $\tau_{i,C}$. Further periodic Markov chains that give minimal expected times to mixing in general never achieve coupling. Since the expected times to mixing are in general easier to calculate than the expected times to coupling we recommend that as a first approximation such expected values be used as an indication as to how long a general Markov chain can take to achieve stationarity.

We conclude the paper with some observations pertaining to the expected times to mixing and coupling in general finite Markov chains focussing, in particular, on the case of independent trials.

2. Mixing Times

The irreducibility of the Markov chain ensures that the T_{ij} are all proper random variables ([4, Theorem 5.3.6]), ensuring the finiteness of the “mixing time” (a.s.).

Further, under the finite state space restriction, all the moments of T_{ij} are finite, ([5, Theorem 7.3.1]). Let m_{ij} be the mean first passage time from state i to state j , i.e. $m_{ij} = E[T_{ij} | X_0 = i]$ for all $i, j \in S$. (It is possible, in the presence of null states in the case of an infinite state space for $m_{ii} = +\infty$.)

In [3] we showed that if η_i is the expected time to mixing starting at state i , then

$$\eta_i = \sum_{j=1}^m m_{ij} \pi_j. \quad (2.1)$$

We further showed that $\eta_i = \eta$, a constant independent of i , the starting state and obtained various alternative expressions for η using generalized inverses of $I - P$, (where I is the identity matrix). In particular, if $\boldsymbol{\eta}^T = (\eta_1, \eta_2, \dots, \eta_m)$ and $\mathbf{e}^T = (1, 1, \dots, 1)$, we have the following general results initially derived in [3]. (For the properties of g-inverses of $I - P$ see [4], [5] and [6]).

If $G = [g_{ij}]$ is any g-inverse of $I - P$, and $\Pi = \mathbf{e}\boldsymbol{\pi}^T$,

$$\boldsymbol{\eta} = [1 - \text{tr}(G\Pi) + \text{tr}(G)]\mathbf{e} = \boldsymbol{\eta}\mathbf{e}. \quad (2.2)$$

Further, if $g_j = \sum_{k=1}^m g_{jk}$, then

$$\eta = 1 + \sum_{j=1}^m (g_{jj} - g_j \cdot \pi_j). \quad (2.3)$$

Further, if for some g , $G\mathbf{e} = g\mathbf{e}$, $\eta = 1 - g + \text{tr}(G)$. (2.4)

In particular,
$$\eta = \text{tr}(Z) = 1 + \text{tr}(T). \quad (2.5)$$

where Z is Kemeny and Snell's fundamental matrix $Z = [I - P - \Pi]^{-1}$ given initially in [8, Theorem 4.4.7] and [7]; and T is Meyer's group inverse of $I - P$, $Z - \Pi$ as given in [10, Theorem 3.3].

If we assume that mixing can occur initially when $X_0 = Y$, so that we consider "hitting times" rather than "return times" to any state, then the expected time to mixing, starting in state i , is

$$\tau_{M,i} = \eta_i - 1 = \eta - 1 = \tau_M \quad (= \sum_{j \neq i}^m m_{ij} \pi_j, \text{ since } m_{ii} = 1/\pi_i.)$$

Thus from equations (2.2), (2.3), (2.4) and (2.5), using the same notation,

$$\tau_M = \sum_{j=1}^m (g_{jj} - g_j \cdot \pi_j) = \text{tr}(G) - g = \text{tr}(Z) - 1 = \text{tr}(T). \quad (2.6)$$

3. Coupling Times

Under the assumption that $\{X_n\}$ and $\{Y_n\}$ both have a finite state space $S = \{1, 2, \dots, m\}$, $Z_n = (X_n, Y_n)$, ($n \geq 0$), is a (two-dimensional) Markov chain with state space $S \times S = \{(i, j), 1 \leq i \leq m, 1 \leq j \leq m\}$. 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

$$P\{Z_{n+1} = (k, l) \mid Z_n = (i, j)\} = P\{X_{n+1} = k, Y_{n+1} = l \mid X_n = i, Y_n = j\}$$

$$= P\{X_{n+1} = k \mid X_n = i\} \cdot P\{Y_{n+1} = l \mid Y_n = j\} = p_{ik} p_{jl}.$$

Once coupling occurs at time $T = n$, $X_{n+k} = Y_{n+k}$ for all $k \geq 0$.

If $Z_0 \in C$, coupling of the two Markov chains is instantaneous and the coupling time $T = 0$. We define $T_{ij,kl}$ to be the first passage time from state (i, j) to state (k, l) . Observe that 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 . We define $T_{i,C} = 0$, ($1 \leq i \leq m$), 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

$$P = \begin{bmatrix} I & 0 \\ R & Q \end{bmatrix} \quad (3.1)$$

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! (If $(X_n, Y_n) = (i, j)$ ($i \neq j$) then it is impossible for $(X_n, Y_n) = (k, k)$ for any k due to the identical cyclicity of each embedded chain. If the period of the chains is however less than m coupling may be possible. We restrict attention to embedded regular chains.

Let the n -step first passage time matrix be $F^{(n)} \equiv [f_{ij,kl}^{(n)}]$, where $f_{ij,kl}^{(n)} \equiv P\{T_{ij,kl} = n\}$. It is well known that for absorbing chains ([5], Theorem 6.2.1) that with the equivalent partitioning as given by (3.1), $F^{(n)} = \begin{bmatrix} F_1^{(n)} & F_2^{(n)} \\ F_3^{(n)} & F_4^{(n)} \end{bmatrix}$, then $F_1^{(1)} = I, F_1^{(n)} = 0, (n \geq 2); F_2^{(n)} = 0, (n \geq 1); F_3^{(n)} = Q^{n-1}R, (n \geq 1); F_4^{(n)} = Q[F_4^{(n-1)} - F_{4d}^{(n-1)}]; F_4^{(1)} = Q$, using the notation that if $A = [a_{ij}]$ then $A_d = [\delta_{ij} a_{ij}]$.

In particular, the probability that the first passage (coupling) from the starting state, (i, j) , ($i \neq j$) to the coupling states C occurs on the n -th trial ($n \geq 1$), is given by

$$f_{ij,C}^{(n)} = P\{T_{ij,C} = n\} = \sum_{k=1}^m f_{ij,kk}^{(n)} \text{ so that} \\ \mathbf{f}^{(n)} = (f_{ij,C}^{(n)})_{(i,j) \in \mathcal{T}} = F_3^{(n)} \mathbf{e} = Q^{n-1} R \mathbf{e}, \quad (3.2)$$

where \mathbf{e} is an $m \times 1$ column vector of 1's, and $\mathbf{f}^{(n)}$ is an $m(m-1) \times 1$ column vector.

Let the "reaching probability" matrix be $F \equiv [f_{ij,kl}]$, where $f_{ij,kl} \equiv P\{T_{ij,kl} < \infty\}$. For a general absorbing Markov chain, (see [5], Theorem 6.3.1), with the equivalent partitioning as given by (3.1), $F = \begin{bmatrix} F_1 & F_2 \\ F_3 & F_4 \end{bmatrix} = \begin{bmatrix} I & 0 \\ NR & (N-I)N_d^{-1} \end{bmatrix}$, where $N = (I - Q)^{-1}$.

Thus if $f_{ij,kk}$ is the probability that the coupling state (k, k) is ever reached from the initial state (i, j) , ($i \neq j$), then $f_{ij,kk} \equiv \sum_{n=1}^{\infty} f_{ij,kk}^{(n)} = P\{T_{ij,kk} < \infty\}$ and $F_3 \equiv [f_{ij,kk}] = NR$.

Now $f_{ij,C} \equiv \sum_{k=1}^m \sum_{n=1}^{\infty} f_{ij,kk}^{(n)} = P\{T_{ij,C} < \infty\}$ is the probability of the coupling occurring in finite time, starting in state (i, j) ($i \neq j$), so that

$$\mathbf{f} = (f_{ij,C})_{(i,j) \in \mathcal{T}} = F_3 \mathbf{e} = NR \mathbf{e}.$$

Now, since P is a stochastic matrix, $P \mathbf{e} = \mathbf{e}$, from Eqn. (3.1), (with appropriate dimensions), we deduce that $R \mathbf{e} + Q \mathbf{e} = \mathbf{e}$, implying that $R \mathbf{e} = (I - Q) \mathbf{e}$. Consequently

$$\mathbf{f} = NR \mathbf{e} = (I - Q)^{-1} (I - Q) \mathbf{e} = \mathbf{e}. \quad (3.3)$$

i.e. $f_{ij,C} = 1$, implying that starting from any state (i, j) , coupling will occur with probability one.

Let $\kappa_{ij}^{(C)} = E[T_{ij,C}]$ be the expected time to coupling starting in state $X_0 = i, Y_0 = j$. By definition

$$\kappa_{ij}^{(C)} = \sum_{n=1}^{\infty} nP[T_{ij,C} = n] = \sum_{n=1}^{\infty} nf_{ij,C}^{(n)} = \sum_{k=1}^m \sum_{n=1}^{\infty} nf_{ij,kk}^{(n)}. \quad (3.4)$$

Define $\kappa_{ij,kk} = \sum_{n=1}^{\infty} nf_{ij,kk}^{(n)}$ and $K = (\kappa_{ij,kk}) = (\sum_{n=1}^{\infty} nf_{ij,kk}^{(n)})$ then $K = \sum_{n=1}^{\infty} nF_3^{(n)}$.

Using the expression $F_3^{(n)}$, and noting that $(I - Q)^{-1}$ exists (and hence has all eigenvalues less than 1 in modulus), from Exercise 4.5.1 of [4] we deduce that

$$K = \sum_{n=1}^{\infty} nQ^{n-1}R = (I - Q)^{-2}R = N^2R. \quad (3.5)$$

Let $\boldsymbol{\kappa} \equiv (\kappa_{ij}^{(C)})$, a column vector (of dimension $m(m-1) \times 1$) of the expected times to coupling. Since $\kappa_{ij}^{(C)} = \sum_{k=1}^m \kappa_{ij,kk}$, we deduce from (3.5), and simplifying using (3.3), that

$$\boldsymbol{\kappa} = K\mathbf{e} = NN\mathbf{R}\mathbf{e} = N\mathbf{e}. \quad (3.6)$$

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 , ($1 \leq i \leq m$) is

$$\tau_{C,i} = \sum_{j=1}^m \pi_j E[T_{ij,C}] = \sum_{j \neq i} \pi_j \kappa_{ij}^{(C)}. \quad (3.7)$$

To evaluate (3.7), we restrict attention to those entries in $\boldsymbol{\kappa}$ that start in state i . Let $\boldsymbol{\kappa}_1^T = (\kappa_{12}^{(C)}, \dots, \kappa_{1j}^{(C)}, \dots, \kappa_{1m}^{(C)})$, \dots , $\boldsymbol{\kappa}_i^T = (\kappa_{i1}^{(C)}, \dots, \kappa_{i,i-1}^{(C)}, \kappa_{i,i+1}^{(C)}, \dots, \kappa_{im}^{(C)})$, \dots , $\boldsymbol{\kappa}_m^T = (\kappa_{m1}^{(C)}, \dots, \kappa_{m,m-1}^{(C)})$, and hence re-express $\boldsymbol{\kappa}$ as $\boldsymbol{\kappa}^T = (\boldsymbol{\kappa}_1^T, \dots, \boldsymbol{\kappa}_i^T, \dots, \boldsymbol{\kappa}_m^T)$.

If we define $\boldsymbol{\rho}_i^T = \boldsymbol{\pi}^T [e_1, e_2, \dots, e_{i-1}, e_{i+1}, \dots, e_m] = (\pi_1, \dots, \pi_{i-1}, \pi_{i+1}, \dots, \pi_m)$, i.e. a modification of $\boldsymbol{\pi}^T$ to yield a vector of dimension $1 \times (m - 1)$ (with π_i removed at the i -th position from $\boldsymbol{\pi}^T$) then, for $1 \leq i \leq m$,

$$\tau_{C,i} = \boldsymbol{\rho}_i^T \boldsymbol{\kappa}_i. \quad (3.8)$$

There are some key observations, related to the structure of the Q matrix, that lead to simplification of the calculation of these expected values.

First note that from Eqn. (3.6), $\boldsymbol{\kappa} = N\mathbf{e} = (I - Q)^{-1}\mathbf{e} \Rightarrow (I - Q)\boldsymbol{\kappa} = \mathbf{e}$ so that $\boldsymbol{\kappa}$ can be obtained by solving the set of linear equations

$$(I - Q)\boldsymbol{\kappa} = \mathbf{e}. \quad (3.9)$$

In deriving the properties of the coupling process $\mathbf{Z}_n = (X_n, Y_n)$ the original state space S consisting of m states was expanded to $S \times S$, a state space consisting of m^2 states. The Q -matrix is of dimension $m(m-1) \times m(m-1)$ and governs the transitions within the $m(m-1)$ transient states. However, if we look carefully at this matrix we observe some symmetry. In particular, suppose we look at the sub-matrix of one-step transition probabilities governing transitions between the states (i, j) and (j, i) ($i \neq j$). Observe the symmetry as follows

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

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

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

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

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

What this implies is that by appropriately 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 nett effect of this observation is that instead of inverting the $m(m-1)$ dimensional matrix N , or solving the linear equations (3.9), we need only solve the equivalent of $m(m-1)/2$ equations. We illustrate this further in the $m = 3$ case.

4. Special cases

Example 4.1a. *Two-state Markov chains (Mixing)*

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 $d = 1 - a - b$ so that $1 - d = a + b$.

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

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

If $-1 < d < 1$, the Markov chain is regular and this stationary distribution is in fact the limiting distribution. If $d = -1$ the Markov chain is irreducible periodic, period 2.

The expected time to mixing, $\tau_M = \frac{1}{1-d} = \frac{1}{a+b}$.

From [3], for all two-state irreducible Markov chains, $\tau_M \geq 0.5$ with the minimum value of $\tau_M = 0.5$ occurring when $d = -1$, ($a = 1, b = 1$) in which case the Markov chain is periodic, period 2. Arbitrarily large values of τ_M occur as $d \rightarrow 1$, (when both $a \rightarrow 0$ and

$b \rightarrow 0$), when the chain is approaching the situation when the chain is close to being reducible with both states absorbing.

Example 4.1b. *Two-state Markov chains (Coupling)*

With the same transition matrix as given in Example 4.1a, and the notation of Section 3,

$$\text{observe that } Q = \begin{matrix} & (1,2) & (2,1) \\ \begin{matrix} (1,2) \\ (2,1) \end{matrix} & \begin{bmatrix} p_{11}p_{22} & p_{12}p_{21} \\ p_{21}p_{12} & p_{22}p_{11} \end{bmatrix} \end{matrix} = \begin{bmatrix} (1-a)(1-b) & ab \\ ab & (1-a)(1-b) \end{bmatrix}.$$

Now Eqn. (3.9) $(I - Q)\kappa = e$ can be written as

$$\begin{bmatrix} 1 - (1-a)(1-b) & -ab \\ -ab & 1 - (1-a)(1-b) \end{bmatrix} \begin{bmatrix} \kappa_{12}^{(C)} \\ \kappa_{21}^{(C)} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

This can either be solved by taking the inverse of $(I - Q)$, N , or solving the linear equations to give

$$N = \frac{1}{(a+b)(a+b-2ab)} \begin{bmatrix} 1 - (1-a)(1-b) & ab \\ ab & 1 - (1-a)(1-b) \end{bmatrix} \Rightarrow \begin{bmatrix} \kappa_{12}^{(C)} \\ \kappa_{21}^{(C)} \end{bmatrix} = \frac{1}{(a+b-2ab)} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

$$\text{Now } \tau_{C,1} = \pi_2 \kappa_{21}^{(C)} = \frac{a}{(a+b)(a+b-2ab)}, \text{ and } \tau_{C,2} = \pi_1 \kappa_{12}^{(C)} = \frac{b}{(a+b)(a+b-2ab)}.$$

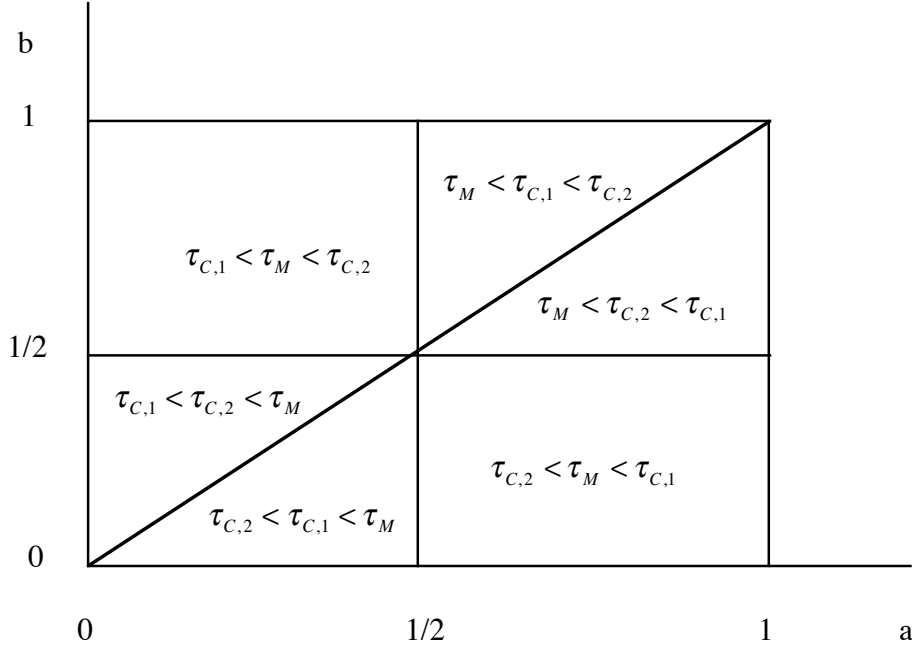
Thus the expected times to coupling are state dependent with

$$\tau_{C,1} < \tau_{C,2} \Leftrightarrow a < b, \quad \tau_{C,1} = \tau_{C,2} \Leftrightarrow a = b, \quad \tau_{C,1} > \tau_{C,2} \Leftrightarrow a > b;$$

$$\tau_{C,1} < \tau_M \Leftrightarrow a < \frac{1}{2}, \quad \tau_{C,1} = \tau_M \Leftrightarrow a = \frac{1}{2}, \quad \tau_{C,1} > \tau_M \Leftrightarrow a > \frac{1}{2};$$

$$\tau_{C,2} < \tau_M \Leftrightarrow b < \frac{1}{2}, \quad \tau_{C,2} = \tau_M \Leftrightarrow b = \frac{1}{2}, \quad \tau_{C,2} > \tau_M \Leftrightarrow b > \frac{1}{2}.$$

Figure 1 gives regions where different inequalities exist between the three different expected values. As can be seen every possible inequality arrangement can be achieved somewhere in the region $(0,1] \times (0,1]$, with selected equalities occurring on the boundaries where $a = 1/2$, $b = 1/2$ and $a = b$. This leads to the observation that we cannot determine any universal inequalities for all a and b .

Figure 1: Two state MC - Inequalities between expected mixing and coupling times**Example 4.2a. Three-state Markov chains (Mixing)**

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

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

It is easily shown that the stationary probability vector is $(\pi_1, \pi_2, \pi_3) = \frac{1}{\Delta}(\Delta_1, \Delta_2, \Delta_3)$,

Define $\tau \equiv p_{12} + p_{13} + p_{21} + p_{23} + p_{31} + p_{32} = b + c + d + f + g + h$.

It can be verified, from [3] using the results for η , that for all three-state MC's,

$$\tau_M = \frac{\tau}{\Delta}. \quad (4.1)$$

The following special cases were explored in detail in [3].

Case 1: “Minimal period 3”

If $p_{12} = p_{23} = p_{31} = 1$ i.e. $b = f = g = 1$, the Markov chain is periodic, period 3, with

transitions cycling from $1 \rightarrow 2 \rightarrow 3 \rightarrow 1 \dots$ and transition matrix $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$.

Note that $\Delta_1 = \Delta_2 = \Delta_3 = 1$, $\Delta = 3$, implying $\boldsymbol{\pi}^T = (1/3, 1/3, 1/3)$. Further $\tau = 3$, implying from (4.1) that $\tau_M = 1$, the smallest possible amongst all 3-state Markov chains.

Case 2: “Period 2”

If $p_{12} = p_{32} = 1$ and $p_{22} = 0$, i.e. $b = 1$, $d + f = 1$, $h = 1$, the Markov chain is periodic, period 2 (with transitions alternating between the states $\{1, 3\}$ and $\{2\}$), and transition matrix

$$P = \begin{bmatrix} 0 & 1 & 0 \\ d & 0 & f \\ 0 & 1 & 0 \end{bmatrix}.$$

Then $\Delta_1 = d$, $\Delta_2 = 1$, $\Delta_3 = f$, $\Delta = 2$, implying $\boldsymbol{\pi}^T = (d/2, 1/2, f/2)$. Further $\tau = 3$, and hence from (4.1), $\tau_M = 1.5$.

Case 3: “Constant movement”

If $p_{11} = p_{22} = p_{33} = 0$, i.e. $b + c = 1$, $d + f = 1$, $g + h = 1$, then at each step the chain does not remain at the state but moves to one of the other states. The Markov chain is irreducible, and regular if $0 < b < 1$, $0 < f < 1$, $0 < g < 1$. The transition matrix is

$$P = \begin{bmatrix} 0 & b & 1-b \\ 1-f & 0 & f \\ g & 1-g & 0 \end{bmatrix}.$$

Now $\Delta_1 = 1 - f(1 - g)$, $\Delta_2 = 1 - g(1 - b)$, $\Delta_3 = 1 - b(1 - f)$, implying that $\Delta = 3 - f(1 - g) - g(1 - b) - b(1 - f)$. Further $\tau = 3$, and hence, from (4.1),

$$\tau_M = \frac{3}{3 - b(1 - f) - f(1 - g) - g(1 - b)} = \frac{3}{3 - p_{12}p_{21} - p_{23}p_{32} - p_{31}p_{13}}.$$

In [3] we showed $1 \leq \tau_M \leq 1.5$. The minimal value of 1 occurs when $b = f = g = 1$, and this case reduces to the “period 3” Case 1 above. Further when $b = f = g = 0$ this case again reduces to a periodic, “period 3” chain but with transitions $1 \rightarrow 3 \rightarrow 2 \rightarrow 1 \dots$.

The maximal value of 1.5 occurs when any pair of (b, f, g) take the values 0 and 1, say $b = 1$, $g = 0$, when this case reduces to the “period 2” Case 2 above. For the regular case $1 < \tau_M < 1.5$.

Case 4: "Independent"

Suppose $p_{ij} = p_j$ for all i, j implying that the Markov chain is equivalent to independent trials on the state space $S = \{1, 2, 3\}$. Now $\Delta_1 = p_1, \Delta_2 = p_2, \Delta_3 = p_3$, so that $\Delta = 1$. Also $\tau = 2$, implying $\tau_M = 2$.

Case 5: "Cyclic drift"

Let $p_{13} = p_{21} = p_{32} = 0$, i.e. $c = d = h = 0$, with $0 < b < 1, 0 < f < 1, 0 < g < 1$, implying

that the Markov chain is regular with transition matrix $P = \begin{bmatrix} 1-b & b & 0 \\ 0 & 1-f & f \\ g & 0 & 1-g \end{bmatrix}$.

At each transition the chain either remains in the same state i or moves to state $i + 1$ (or 1, if $i = 3$).

Now $\Delta_1 = fg, \Delta_2 = gb, \Delta_3 = bf$ so that $\Delta = bf + fg + gb$. Also $\tau = b + f + g$.

Note that $0 < \tau < 3$ and $0 < \Delta < 3$ implying

$$\tau_M = \frac{b + f + g}{bf + fg + gb} = \frac{p_{12} + p_{23} + p_{21}}{p_{12}p_{23} + p_{23}p_{21} + p_{21}p_{12}}.$$

When $b + f + g \rightarrow 3$ then $bf + fg + gb \rightarrow 3$ and $\tau_M \rightarrow 1$ (as in Case 1).

When $b + f + g \rightarrow 0$ then $bf + fg + gb \rightarrow 0$, but the behaviour of τ_M depends upon the rates of convergence and can be large. In this situation the Markov chain resides for a large number of transitions in each state so that there is little movement implying that the mixing time can become excessively large.

Case 6: "Constant probability state selection"

$$\text{Let } P = \begin{bmatrix} 1-a & \frac{a}{2} & \frac{a}{2} \\ \frac{b}{2} & 1-b & \frac{b}{2} \\ \frac{c}{2} & \frac{c}{2} & 1-c \end{bmatrix}, \text{ where } 0 < a < 1, 0 < b < 1, 0 < c < 1.$$

The chain is regular with $\Delta_1 = 3bc/4, \Delta_2 = 3ac/4, \Delta_3 = 3ab/4$ so that

$$\Delta = 3(bc + ca + ab)/4. \text{ Also } \tau = a + b + c, \text{ leading, from (4.1), to } \tau_M = \frac{4(a + b + c)}{3(bc + ca + ab)}.$$

With $a = b = c = e$, $\tau_M = \frac{4}{3e}$ so that $\frac{4}{3} < \tau_M < \infty$. The lower bound is approached as $e \rightarrow 1$ (with the Markov chain approaching the special case of $b = f = g = 1/2$ as in Case 3).

In [3] we showed that for all 3-state irreducible Markov chains $\tau_M \geq 1$, consistent with our observations above.

Example 4.2b. *Three-state Markov chains (Coupling)*

The key to achieving some symmetry in the Q matrix is to take the transient states in a string of ordered pairs. For example if $\mathcal{T} = \{(1, 2), (2, 1), (1, 3), (3, 1), (2, 3), (3, 2)\}$ then Q can be displayed as

$$Q = \begin{bmatrix} p_{11}p_{22} & p_{12}p_{21} & p_{11}p_{23} & p_{13}p_{21} & p_{12}p_{23} & p_{13}p_{22} \\ p_{12}p_{21} & p_{11}p_{22} & p_{13}p_{21} & p_{11}p_{23} & p_{13}p_{22} & p_{12}p_{23} \\ p_{11}p_{32} & p_{12}p_{31} & p_{11}p_{33} & p_{13}p_{31} & p_{12}p_{33} & p_{13}p_{32} \\ p_{12}p_{31} & p_{11}p_{32} & p_{13}p_{31} & p_{11}p_{33} & p_{13}p_{32} & p_{12}p_{33} \\ p_{21}p_{32} & p_{22}p_{31} & p_{21}p_{33} & p_{23}p_{31} & p_{22}p_{33} & p_{23}p_{32} \\ p_{22}p_{31} & p_{21}p_{32} & p_{23}p_{31} & p_{21}p_{33} & p_{23}p_{32} & p_{22}p_{33} \end{bmatrix}.$$

Let $\mathbf{\kappa}^T = (\kappa_{12}^{(C)}, \kappa_{21}^{(C)}, \kappa_{13}^{(C)}, \kappa_{31}^{(C)}, \kappa_{23}^{(C)}, \kappa_{32}^{(C)})$ then the linear equations (3.9) can be expressed as

$$\begin{bmatrix} 1 - p_{11}p_{22} & -p_{12}p_{21} & -p_{11}p_{23} & -p_{13}p_{21} & -p_{12}p_{23} & -p_{13}p_{22} \\ -p_{12}p_{21} & 1 - p_{11}p_{22} & -p_{13}p_{21} & -p_{11}p_{23} & -p_{13}p_{22} & -p_{12}p_{23} \\ -p_{11}p_{32} & -p_{12}p_{31} & 1 - p_{11}p_{33} & -p_{13}p_{31} & -p_{12}p_{33} & -p_{13}p_{32} \\ -p_{12}p_{31} & -p_{11}p_{32} & -p_{13}p_{31} & 1 - p_{11}p_{33} & -p_{13}p_{32} & -p_{12}p_{33} \\ -p_{21}p_{32} & -p_{22}p_{31} & -p_{21}p_{33} & -p_{23}p_{31} & 1 - p_{22}p_{33} & -p_{23}p_{32} \\ -p_{22}p_{31} & -p_{21}p_{32} & -p_{23}p_{31} & -p_{21}p_{33} & -p_{23}p_{32} & 1 - p_{22}p_{33} \end{bmatrix} \begin{bmatrix} \kappa_{12}^{(C)} \\ \kappa_{21}^{(C)} \\ \kappa_{13}^{(C)} \\ \kappa_{31}^{(C)} \\ \kappa_{23}^{(C)} \\ \kappa_{32}^{(C)} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}. \quad (4.1)$$

This matrix equation (4.1) gives rise to 6 linear equations in 6 unknowns. Rather than attempt to invert the 6×6 matrix $(I - Q)$ to obtain N , there is an easier way to solve for the unknowns. The symmetry of $I - Q$ suggests producing another set of three equations by adding each successive pair of equations and a further set of three equations by subtracting the 2nd from the 1st, the 4th from the 3rd and the 6th from the 5th equation.

Let $\kappa_{12}^{(C)} + \kappa_{21}^{(C)} = \alpha_1$, $\kappa_{13}^{(C)} + \kappa_{31}^{(C)} = \alpha_2$, $\kappa_{23}^{(C)} + \kappa_{32}^{(C)} = \alpha_3$ and $\kappa_{12}^{(C)} - \kappa_{21}^{(C)} = \beta_1$, $\kappa_{13}^{(C)} - \kappa_{31}^{(C)} = \beta_2$, $\kappa_{23}^{(C)} - \kappa_{32}^{(C)} = \beta_3$.

With these definitions the six linear equations (4.1) are replaced by two sets of Equations (4.2) and (4.3) below:

$$\begin{bmatrix} 1 - p_{11}p_{22} - p_{12}p_{21} & -p_{11}p_{23} - p_{13}p_{21} & -p_{12}p_{23} - p_{13}p_{22} \\ -p_{11}p_{32} - p_{12}p_{31} & 1 - p_{11}p_{33} - p_{13}p_{31} & -p_{12}p_{33} - p_{13}p_{32} \\ -p_{21}p_{32} - p_{22}p_{31} & -p_{21}p_{33} - p_{23}p_{31} & 1 - p_{22}p_{33} - p_{23}p_{32} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}, \quad (4.2)$$

$$\begin{bmatrix} 1 - p_{11}p_{22} + p_{12}p_{21} & -p_{11}p_{23} + p_{13}p_{21} & -p_{12}p_{23} + p_{13}p_{22} \\ -p_{11}p_{32} + p_{12}p_{31} & 1 - p_{11}p_{33} + p_{13}p_{31} & -p_{12}p_{33} + p_{13}p_{32} \\ -p_{21}p_{32} + p_{22}p_{31} & -p_{21}p_{33} + p_{23}p_{31} & 1 - p_{22}p_{33} + p_{23}p_{32} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \quad (4.3)$$

These equations can be expressed in matrix form as $A\alpha = 2e$, and $B\beta = 0$.

Using Maple (with the p_{ij} elements replaced by the simplified notation of the second expression for P as displayed in Example 4.2a) it can be shown that $\det(B) = \Delta(\tau - \Delta)$. Note that, by the irreducibility of the initial Markov chain, $\Delta > 0$, and since coupling does not occur for a periodic Markov chain, we know that $\tau_M > 1$ (cf. Case1 of Example 4.2a) implying that $\tau > \Delta$ and hence $\det(B) \neq 0$. Thus B is non-singular and B^{-1} exists. From (4.3), we conclude that $\beta = 0$.

This implies that $\kappa_{12}^{(C)} = \kappa_{21}^{(C)}$, $\kappa_{13}^{(C)} = \kappa_{31}^{(C)}$, $\kappa_{23}^{(C)} = \kappa_{32}^{(C)}$, and consequently $2\kappa_{12}^{(C)} = \alpha_1$, $2\kappa_{13}^{(C)} = \alpha_2$, $2\kappa_{23}^{(C)} = \alpha_3$. If we define $\mathbf{\kappa}^{(C)T} = (\kappa_{12}^{(C)}, \kappa_{13}^{(C)}, \kappa_{23}^{(C)})$ then, from (4.2),

$$\begin{bmatrix} 1 - p_{11}p_{22} - p_{12}p_{21} & -p_{11}p_{23} - p_{13}p_{21} & -p_{12}p_{23} - p_{13}p_{22} \\ -p_{11}p_{32} - p_{12}p_{31} & 1 - p_{11}p_{33} - p_{13}p_{31} & -p_{12}p_{33} - p_{13}p_{32} \\ -p_{21}p_{32} - p_{22}p_{31} & -p_{21}p_{33} - p_{23}p_{31} & 1 - p_{22}p_{33} - p_{23}p_{32} \end{bmatrix} \begin{bmatrix} \kappa_{12}^{(C)} \\ \kappa_{13}^{(C)} \\ \kappa_{23}^{(C)} \end{bmatrix} = A\mathbf{\kappa}^{(C)} = e. \quad (4.4)$$

The solution of the set of equations (4.4) can be effected but, even with Maple, we have been unable to find compact expressions for all possible parameter values. We can use MatLab to calculate values for the expected times.

Note however that the elements of $\mathbf{\kappa}^{(C)}$ are subsidiary results in that we are interested in expressions for the elements of $\mathbf{\tau}^{(C)T} = (\tau_{C,1}, \tau_{C,2}, \tau_{C,3})$.

Note from (3.8) that $\tau_{C,i} = \rho_i^T \mathbf{\kappa}_i$ implying

$$\tau_{C,1} = \pi_2 \kappa_{12}^{(C)} + \pi_3 \kappa_{13}^{(C)}, \quad \tau_{C,2} = \pi_1 \kappa_{21}^{(C)} + \pi_3 \kappa_{23}^{(C)}, \quad \tau_{C,3} = \pi_1 \kappa_{31}^{(C)} + \pi_2 \kappa_{32}^{(C)}. \quad (4.5)$$

Thus $\mathbf{\tau}^{(C)} = D\mathbf{\kappa}^{(C)}$ where $D = \begin{bmatrix} \pi_2 & \pi_3 & 0 \\ \pi_1 & 0 & \pi_3 \\ 0 & \pi_1 & \pi_2 \end{bmatrix}$, or that $\mathbf{\kappa}^{(C)} = D^{-1}\mathbf{\tau}^{(C)}$ implying

$AD^{-1}\mathbf{\tau}^{(C)} = e$ where

$$D^{-1} = \frac{1}{2\pi_1\pi_2\pi_3} \begin{bmatrix} \pi_1\pi_3 & \pi_2\pi_3 & -\pi_3^2 \\ \pi_1\pi_2 & -\pi_2^2 & \pi_2\pi_3 \\ -\pi_1^2 & \pi_1\pi_2 & \pi_1\pi_3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1/\pi_2 & 1/\pi_1 & -\pi_3/\pi_1\pi_2 \\ 1/\pi_3 & -\pi_2/\pi_1\pi_3 & 1/\pi_1 \\ -\pi_1/\pi_2\pi_3 & 1/\pi_3 & 1/\pi_2 \end{bmatrix}.$$

In this situation, $\pi_i = \Delta_i/\Delta$, so that substituting in D and D^{-1} , yields

$$\begin{aligned} \boldsymbol{\tau}^{(C)} &= DA^{-1}\mathbf{e} = \frac{1}{\Delta} \begin{bmatrix} \Delta_2 & \Delta_3 & 0 \\ \Delta_1 & 0 & \Delta_3 \\ 0 & \Delta_1 & \Delta_2 \end{bmatrix} A^{-1}\mathbf{e} \\ &= \frac{1}{\Delta} \begin{bmatrix} \Delta_2 & \Delta_3 & 0 \\ \Delta_1 & 0 & \Delta_3 \\ 0 & \Delta_1 & \Delta_2 \end{bmatrix} \begin{bmatrix} 1-p_{11}p_{22}-p_{12}p_{21} & -p_{11}p_{23}-p_{13}p_{21} & -p_{12}p_{23}-p_{13}p_{22} \\ -p_{11}p_{32}-p_{12}p_{31} & 1-p_{11}p_{33}-p_{13}p_{31} & -p_{12}p_{33}-p_{13}p_{32} \\ -p_{21}p_{32}-p_{22}p_{31} & -p_{21}p_{33}-p_{23}p_{31} & 1-p_{22}p_{33}-p_{23}p_{32} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \end{aligned}$$

Alternatively, try to solve for $\boldsymbol{\tau}^{(C)}$ directly from the linear equations $AD^{-1}\boldsymbol{\tau}^{(C)} = \mathbf{e}$:

$$\begin{bmatrix} 1-p_{11}p_{22}-p_{12}p_{21} & -p_{11}p_{23}-p_{13}p_{21} & -p_{12}p_{23}-p_{13}p_{22} \\ -p_{11}p_{32}-p_{12}p_{31} & 1-p_{11}p_{33}-p_{13}p_{31} & -p_{12}p_{33}-p_{13}p_{32} \\ -p_{21}p_{32}-p_{22}p_{31} & -p_{21}p_{33}-p_{23}p_{31} & 1-p_{22}p_{33}-p_{23}p_{32} \end{bmatrix} \begin{bmatrix} \Delta_1\Delta_3 & \Delta_2\Delta_3 & -\Delta_3^2 \\ \Delta_1\Delta_2 & -\Delta_2^2 & \Delta_2\Delta_3 \\ -\Delta_1^2 & \Delta_1\Delta_2 & \Delta_1\Delta_3 \end{bmatrix} \begin{bmatrix} \tau_{C,1} \\ \tau_{C,2} \\ \tau_{C,3} \end{bmatrix} = \frac{2\Delta_1\Delta_2\Delta_3}{\Delta}\mathbf{e}.$$

Case 1: “Minimal period 3” with $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$. Formal substitution in Eqn. (4.4)

leads to $\begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \kappa_{12}^{(C)} \\ \kappa_{13}^{(C)} \\ \kappa_{23}^{(C)} \end{bmatrix} = \mathbf{e}$. This is an inconsistent set of equations with no

solution. This is however consistent with the observation that coupling of periodic chains either occurs initially or never occurs.

Case 2: “Period 2” with $P = \begin{bmatrix} 0 & 1 & 0 \\ d & 0 & f \\ 0 & 1 & 0 \end{bmatrix}$, ($d+f=1$). Eqn. (4.4) yield

$$\begin{bmatrix} 1-d & 0 & -(1-d) \\ 0 & 1 & 0 \\ -d & 0 & d \end{bmatrix} \begin{bmatrix} \kappa_{12}^{(C)} \\ \kappa_{13}^{(C)} \\ \kappa_{23}^{(C)} \end{bmatrix} = \mathbf{e}, \text{ implying } \kappa_{13}^{(C)} = 1 \text{ and } \kappa_{12}^{(C)} = \kappa_{23}^{(C)} (= \infty).$$

There are no finite solutions, consistent with the observation that the periodicity implies that coupling either occurs initially or never occurs.

A key observation is that we will not see any minimal expected time to coupling in the presence of periodicities, as was the case for mixing times.

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

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

$$\text{Eqn. (4.4) yields } A\mathbf{\kappa}^{(C)} = \begin{bmatrix} 1-p_{12}p_{21} & -p_{13}p_{21} & -p_{12}p_{23} \\ -p_{12}p_{31} & 1-p_{13}p_{31} & -p_{13}p_{32} \\ -p_{21}p_{32} & -p_{23}p_{31} & 1-p_{23}p_{32} \end{bmatrix} \begin{bmatrix} \kappa_{12}^{(C)} \\ \kappa_{13}^{(C)} \\ \kappa_{23}^{(C)} \end{bmatrix} = \mathbf{e}.$$

If we define $\eta_2 = p_{12}p_{21} + p_{23}p_{32} + p_{31}p_{13}$, $\eta_3 = p_{12}p_{23}p_{31} - p_{13}p_{21}p_{32}$, then

$$A^{-1} = \frac{1}{1-\eta_2-\eta_3^2} \begin{bmatrix} 1-p_{23}p_{32}-p_{31}p_{13} & p_{13}p_{21}+p_{23}\eta_3 & p_{12}p_{23}-p_{13}\eta_3 \\ p_{12}p_{31}-p_{32}\eta_3 & 1-p_{12}p_{21}-p_{23}p_{32} & p_{13}p_{32}+p_{12}\eta_3 \\ p_{21}p_{32}+p_{31}\eta_3 & p_{23}p_{31}-p_{21}\eta_3 & 1-p_{13}p_{31}-p_{21}p_{12} \end{bmatrix}.$$

This leads to

$$\begin{bmatrix} \kappa_{12}^{(C)} \\ \kappa_{13}^{(C)} \\ \kappa_{23}^{(C)} \end{bmatrix} = \frac{1}{1-\eta_2-\eta_3^2} \begin{bmatrix} 1+p_{13}(p_{21}-p_{31})+p_{23}(p_{12}-p_{32})+(p_{23}-p_{13})\eta_3 \\ 1+p_{12}(p_{31}-p_{21})+p_{32}(p_{13}-p_{23})+(p_{12}-p_{32})\eta_3 \\ 1+p_{21}(p_{32}-p_{12})+p_{31}(p_{23}-p_{13})+(p_{31}-p_{21})\eta_3 \end{bmatrix}.$$

Finally, using Eqn. (4.5) and noting that $\Delta_1 = 1 - p_{23}p_{32}$, $\Delta_2 = 1 - p_{31}p_{13}$, $\Delta_3 = 1 - p_{12}p_{21}$ and $\Delta = 3 - \eta_2$ to obtain expressions for $\tau_{C,i}$.

$$\begin{bmatrix} \tau_{C,1} \\ \tau_{C,2} \\ \tau_{C,3} \end{bmatrix} = \frac{1}{(3-\eta_2)(1-\eta_2-\eta_3^2)} \begin{bmatrix} \Delta_2 + \Delta_3 + (\Delta_2 p_{13} - \Delta_3 p_{12})(p_{21} - p_{31}) + (\Delta_2 p_{23} + \Delta_3 \eta_3)(p_{12} - p_{32}) + (\Delta_3 p_{32} - \Delta_2 \eta_3)(p_{13} - p_{23}) \\ \Delta_1 + \Delta_3 + (\Delta_1 p_{13} - \Delta_3 \eta_3)(p_{21} - p_{31}) + (\Delta_1 p_{23} - \Delta_3 p_{21})(p_{12} - p_{32}) + (\Delta_1 \eta_3 + \Delta_3 p_{31})(p_{23} - p_{13}) \\ \Delta_1 + \Delta_2 + (\Delta_1 p_{12} + \Delta_2 \eta_3)(p_{31} - p_{21}) + (\Delta_1 p_{32} - \Delta_2 p_{31})(p_{13} - p_{23}) + (\Delta_2 p_{21} - \Delta_1 \eta_3)(p_{32} - p_{12}) \end{bmatrix}$$

Computation of $\tau_{C,i}$, for all values of the parameters, shows that

$$1 \leq \tau_M \leq 1.5 < 2.6667 \leq \min_{1 \leq i \leq 3} \tau_{C,i} < \infty.$$

Thus, in all situations for this case, the expected time to mixing is less than the expected time to coupling from any state.

Simple regions, in terms of the parameters (transition probabilities, when inequalities between the $\tau_{C,i}$ and $\tau_{C,j}$ ($i \neq j$) are difficult to determine. We can show, for example, that

$$\tau_{C,1} < \tau_{C,2}$$

$$\Leftrightarrow 0 < (\Delta_1 - \Delta_2) \{ p_{12}(p_{31} + p_{23}) + p_{21}(p_{13} + p_{32}) + \eta_3(p_{23} - p_{13}) \} + \Delta_3 \{ p_{12}p_{32} + p_{21}p_{31} + \eta_3(p_{23} - p_{12}) \}.$$

Note that $p_{12}(p_{31} + p_{23}) + p_{21}(p_{13} + p_{32}) + \eta_3(p_{23} - p_{13}) > 0$ and $\Delta_3 > 0$. Sufficient conditions for $\tau_{C,1} < \tau_{C,2}$ can be derived from the above, in terms of

$\Delta_1 - \Delta_2$, η_3 and $p_{23} - p_{12}$. There is no simple region in terms of the independent parameters p_{12} , p_{23} , p_{31} . Similar conditions can be deduced for other inequalities between the $\tau_{C,i}$.

Case 4: “Independent trials” with $P = \begin{bmatrix} p_1 & p_2 & p_3 \\ p_1 & p_2 & p_3 \\ p_1 & p_2 & p_3 \end{bmatrix}$.

In this situation Eqns. (4.4) become $\begin{bmatrix} 1-2p_1p_2 & -2p_1p_3 & -2p_2p_3 \\ -2p_1p_2 & 1-2p_1p_3 & -2p_2p_3 \\ -2p_1p_2 & -2p_1p_3 & 1-2p_2p_3 \end{bmatrix} \begin{bmatrix} \kappa_{12}^{(C)} \\ \kappa_{13}^{(C)} \\ \kappa_{23}^{(C)} \end{bmatrix} = \mathbf{e}$.

These equations yield $\kappa_{12}^{(C)} = \kappa_{13}^{(C)} = \kappa_{23}^{(C)} = 1/(1-2p_1p_2-2p_1p_3-2p_2p_3)$.

Further $\Delta_1 = p_1, \Delta_2 = p_2, \Delta_3 = p_3$, implying $\Delta = 1$ and $p_1 = p_1, p_2 = p_2, p_3 = p_3$.

Thus, from Eqn. (4.5),

$$\tau_{c,1} = \frac{p_2 + p_3}{1-2p_1p_2-2p_1p_3-2p_2p_3}, \tau_{c,2} = \frac{p_1 + p_3}{1-2p_1p_2-2p_1p_3-2p_2p_3}, \tau_{c,3} = \frac{p_1 + p_2}{1-2p_1p_2-2p_1p_3-2p_2p_3}.$$

Note that $\tau_{c,i} < \tau_{c,j}$ if and only if $p_j < p_i$ for all $i \neq j$.

When $p_1 = p_2 = p_3 = \frac{1}{3}$ it is easily seen that for all $i = 1, 2, 3, \tau_{c,i} = 2 = \tau_M$.

It can be shown that for all possible parameter values $0 < p_1, p_2, p_3, < 1$ with $p_1 + p_2 + p_3 = 1$, that $0 < \tau_{c,i} < 2.22474$ with the maximal value of $\tau_{c,i}$ occurring at $p_1 = 0.18350, p_2 = p_3 = 0.40825$ (and similarly for the other $\tau_{c,i}$.)

Of interest is, when does $\tau_{c,i}$ exceed 2, the value of τ_M ? For each i , the region where $\tau_{c,i} = 2$ is an ellipse. Let $q_1 = p_1 + p_2, q_2 = p_1 - p_2$. Consider the values in the (p_1, p_2) -plane (which lie in the interior of the region bounded by $p_1 = 0, p_2 = 0$ and $p_1 + p_2 = 1$).

For $i = 1$, the ellipse is $\left(q_1 + \frac{1}{4}\right)^2 + 3\left(q_2 - \frac{7}{12}\right)^2 = \left(\frac{1}{2\sqrt{3}}\right)^2$, i.e centred on

$$p_1 = \frac{1}{6} = 0.1667, p_2 = \frac{5}{12} = 0.4167.$$

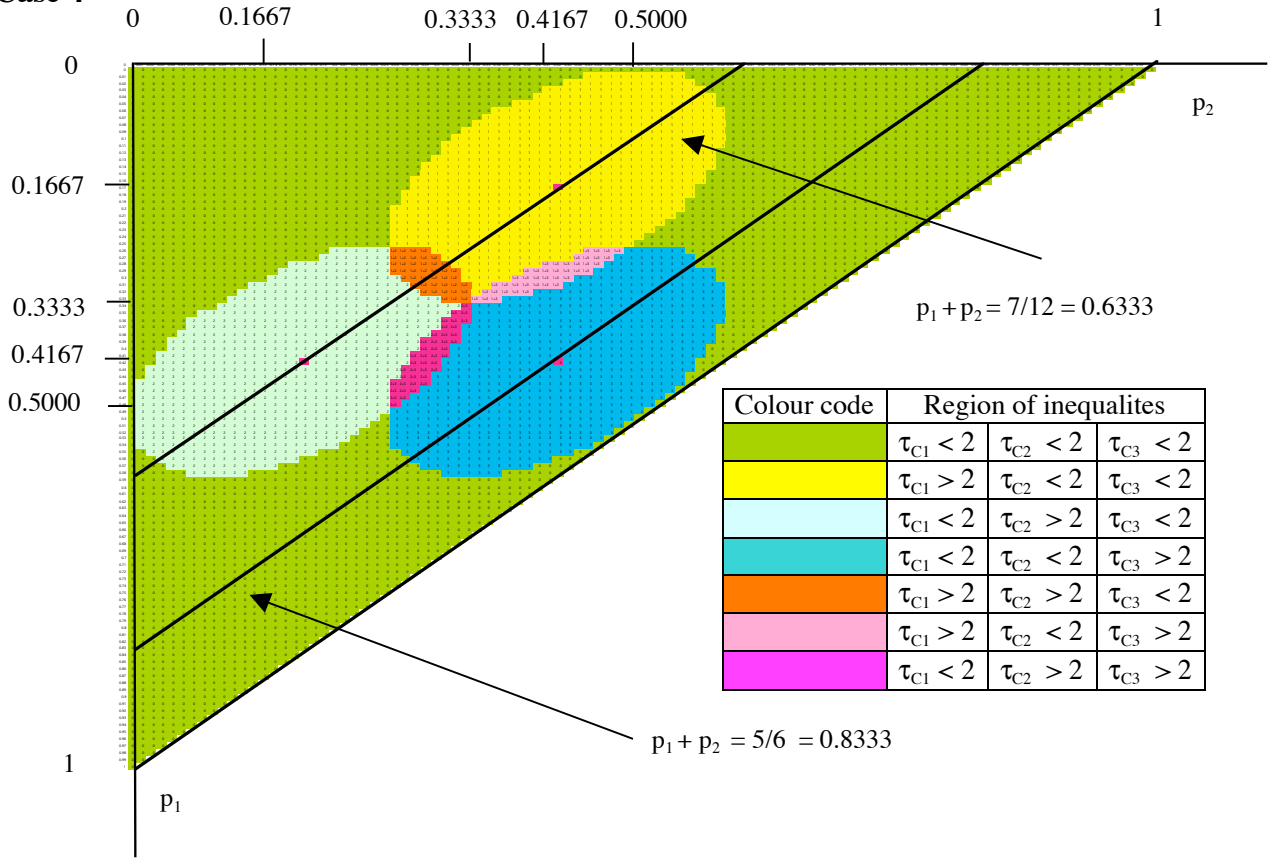
For $i = 2$, the ellipse is $\left(q_1 - \frac{1}{4}\right)^2 + 3\left(q_2 - \frac{7}{12}\right)^2 = \left(\frac{1}{2\sqrt{3}}\right)^2$, i.e. centred on

$$p_1 = \frac{5}{12} = 0.4167, p_2 = \frac{1}{6} = 0.1667.$$

For $i = 3$, the ellipse is $q_1^2 + 3\left(q_2 - \frac{5}{6}\right)^2 = \left(\frac{1}{2\sqrt{3}}\right)^2$, i.e centred on

$$p_1 = \frac{5}{12} = 0.4167, p_2 = \frac{5}{12} = 0.4167.$$

Figure 2: Regions where expected coupling times exceed expected mixing times, Case 4



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

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

Eqn. (4.4) yields $\begin{bmatrix} 1-p_{11}p_{22} & -p_{11}p_{23} & -p_{12}p_{23} \\ -p_{12}p_{31} & 1-p_{11}p_{33} & -p_{12}p_{33} \\ -p_{22}p_{31} & -p_{23}p_{31} & 1-p_{22}p_{33} \end{bmatrix} \begin{bmatrix} \kappa_{12}^{(C)} \\ \kappa_{13}^{(C)} \\ \kappa_{23}^{(C)} \end{bmatrix} = \mathbf{e}.$

Define $\delta_1 = p_{11} + p_{22} + p_{33}$, $\delta_2 = p_{11}p_{22} + p_{22}p_{33} + p_{33}p_{11}$, $\delta_3 = p_{11}p_{22}p_{33} - p_{12}p_{23}p_{31}$, then

$$A^{-1} = \frac{1}{1 - \delta_2 - \delta_3(\delta_1 - \delta_3)} \begin{bmatrix} 1 - p_{33}(\delta_1 - \delta_3 - p_{33}) & p_{23}(p_{11} - \delta_3) & p_{12}p_{23} \\ p_{12}p_{31} & 1 - p_{22}(\delta_1 - \delta_3 - p_{22}) & p_{12}(p_{33} - \delta_3) \\ p_{31}(p_{22} - \delta_3) & p_{23}p_{31} & 1 - p_{11}(\delta_1 - \delta_3 - p_{11}) \end{bmatrix}.$$

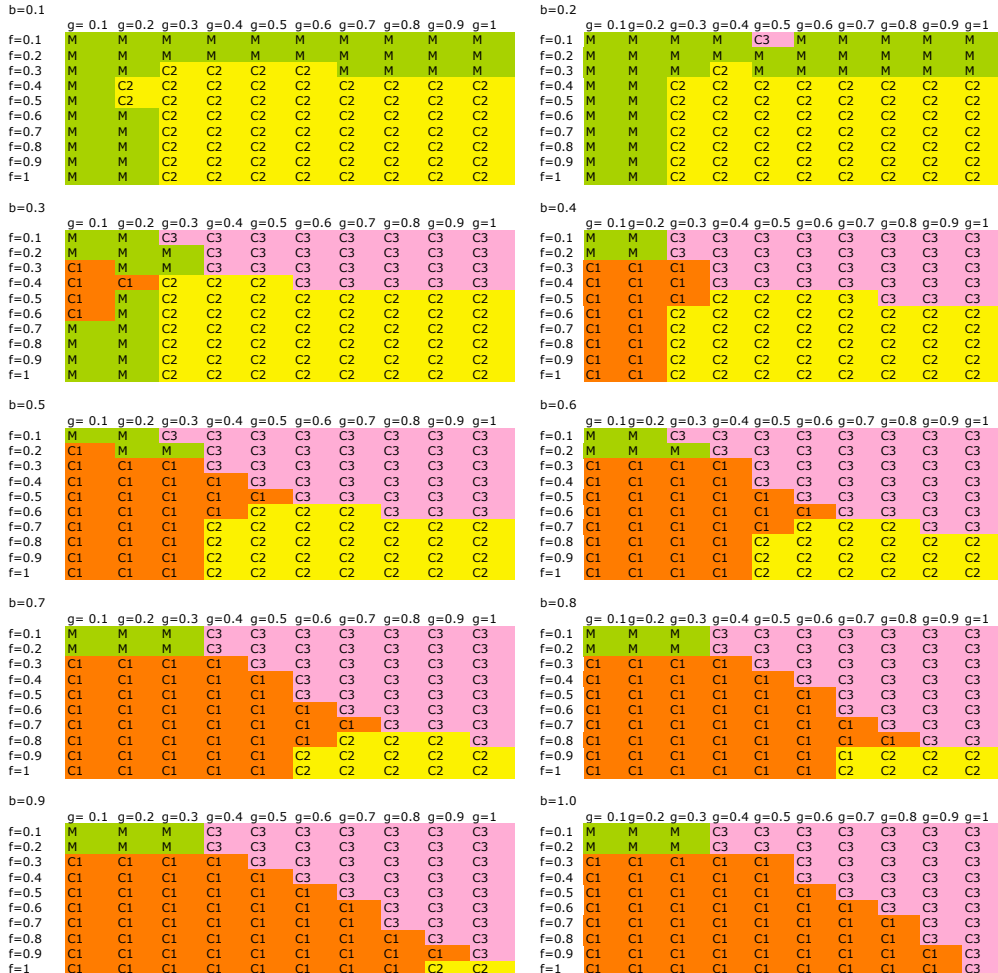
This leads to

$$\begin{bmatrix} \kappa_{12}^{(C)} \\ \kappa_{13}^{(C)} \\ \kappa_{23}^{(C)} \end{bmatrix} = \frac{1}{1 - \delta_2 - \delta_3(\delta_1 - \delta_3)} \begin{bmatrix} 2 - \delta_2 - p_{12}p_{22} + \delta_3(p_{33} - p_{23}) \\ 2 - \delta_2 - p_{31}p_{11} + \delta_3(p_{22} - p_{12}) \\ 2 - \delta_2 - p_{23}p_{33} + \delta_3(p_{11} - p_{31}) \end{bmatrix}.$$

Finally, using Eqn. (4.5) and noting that $\Delta_1 = p_{23}p_{31}$, $\Delta_2 = p_{31}p_{12}$, $\Delta_3 = p_{12}p_{23}$, and $\Delta = p_{23}p_{31} + p_{31}p_{12} + p_{12}p_{23} = 3 - 2\delta_1 + \delta_2$ we can obtain expressions for $\tau_{C,i}$:

$$\begin{bmatrix} \tau_{C,1} \\ \tau_{C,2} \\ \tau_{C,3} \end{bmatrix} = \frac{1}{\Delta(1 - \delta_2 - \delta_3(\delta_1 - \delta_3))} \begin{bmatrix} (p_{31} + p_{23})(2 - \delta_2) - p_{31}(p_{12}p_{22} + p_{23}p_{11}) + \{p_{31}p_{33} + p_{23}(\delta_1 - 2)\}\delta_3 \\ (p_{12} + p_{31})(2 - \delta_2) - p_{12}(p_{31}p_{22} + p_{23}p_{33}) + \{p_{12}p_{11} + p_{31}(\delta_1 - 2)\}\delta_3 \\ (p_{12} + p_{23})(2 - \delta_2) - p_{23}(p_{31}p_{11} + p_{12}p_{33}) + \{p_{23}p_{22} + p_{12}(\delta_1 - 2)\}\delta_3 \end{bmatrix}.$$

Figure 3: Regions where expected coupling times exceed expected mixing times, Case 5



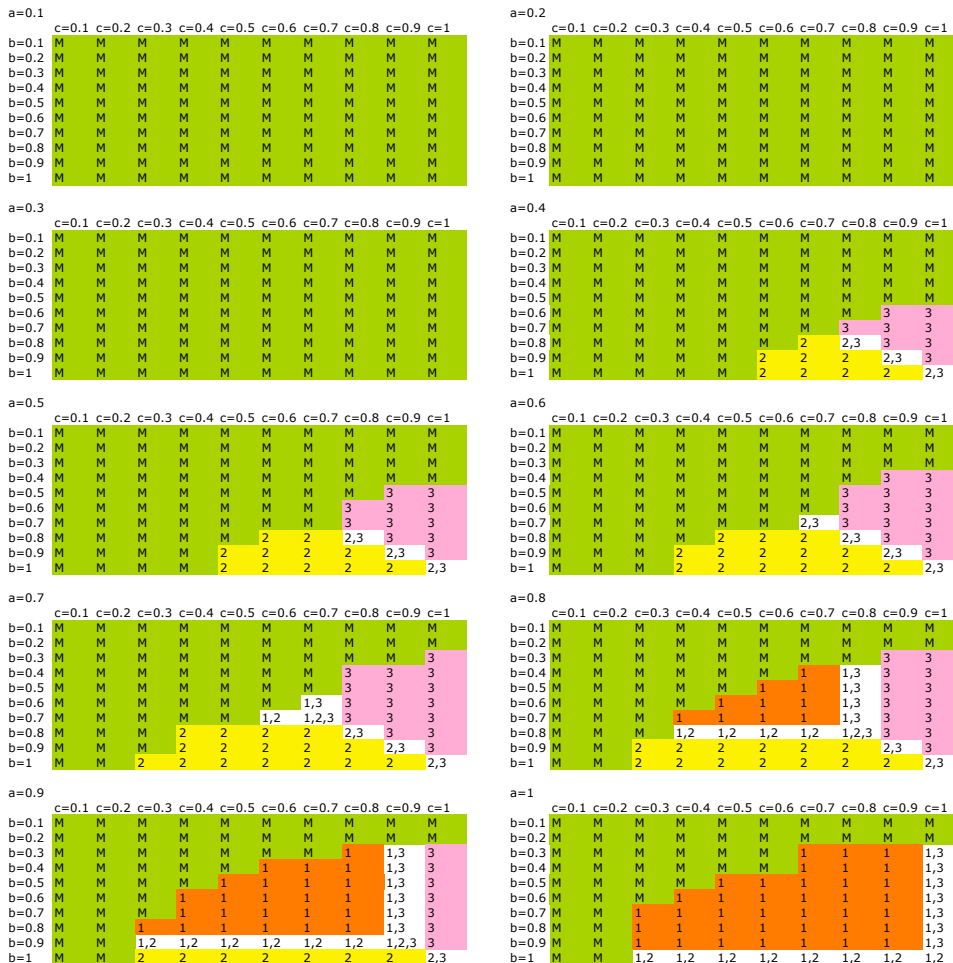
Case 6: “Constant probability state selection”

In this case, with $P = \begin{bmatrix} 1-a & \frac{a}{2} & \frac{a}{2} \\ \frac{b}{2} & 1-b & \frac{b}{2} \\ \frac{c}{2} & \frac{c}{2} & 1-c \end{bmatrix}$, Eqn. (4.4) yields

$$\begin{bmatrix} a+b-\frac{5ab}{4} & -\frac{b}{2}+\frac{ab}{4} & -\frac{a}{2}+\frac{ab}{4} \\ -\frac{c}{2}+\frac{ac}{4} & c+a-\frac{5ac}{4} & -\frac{a}{2}+\frac{ac}{4} \\ -\frac{c}{2}+\frac{bc}{4} & -\frac{b}{2}+\frac{bc}{4} & b+c-\frac{5bc}{4} \end{bmatrix} \begin{bmatrix} \kappa_{12}^{(C)} \\ \kappa_{13}^{(C)} \\ \kappa_{23}^{(C)} \end{bmatrix} = A\kappa^{(C)} = e$$

In this case explicit expressions for the $\tau_{C,i}$ are complicated but, as for Case 5 we can display figures that give regions of the parameter values a, b, c where the expected times, either for mixing or coupling from state i are minimised.

Figure 4: Regions where expected coupling times exceed expected mixing times, Case 6



In attempting to obtain explicit parametric expressions for the $\tau_{C,i}$ let

$\eta_1 = a + b + c, \eta_2 = ab + bc + ca, \eta_3 = abc$. Then $a^2 + b^2 + c^2 = \eta_1^2 - 2\eta_2$, and

$a^2b^2 + b^2c^2 + c^2a^2 = \eta_2^2 - 2\eta_1\eta_3$. With these parameters

$$\det(A) = \frac{3}{4}\eta_1\eta_2 - \frac{1}{2}\eta_3 - \frac{15}{16}\eta_2^2 - \frac{5}{4}\eta_1\eta_3 + \frac{21}{8}\eta_2\eta_3 - \frac{27}{16}\eta_3^2.$$

While it is still difficult to find compact expressions for the $\tau_{C,i}$ we can show that

$$\begin{bmatrix} \kappa_{12}^{(C)} \\ \kappa_{13}^{(C)} \\ \kappa_{23}^{(C)} \end{bmatrix} = \frac{1}{\det(A)} \begin{bmatrix} \left\{ a + c - \frac{5}{4}ac + \frac{b}{2}\left(1 - \frac{a}{2}\right) \right\} \left\{ b + c - \frac{5}{4}bc + \frac{a}{2}\left(1 - \frac{b}{2}\right) \right\} - \frac{ab}{16}(a - c)(b - c) \\ \left\{ c + b - \frac{5}{4}cb + \frac{a}{2}\left(1 - \frac{c}{2}\right) \right\} \left\{ a + b - \frac{5}{4}ab + \frac{c}{2}\left(1 - \frac{a}{2}\right) \right\} - \frac{ca}{16}(c - b)(a - b) \\ \left\{ b + a - \frac{5}{4}ba + \frac{c}{2}\left(1 - \frac{b}{2}\right) \right\} \left\{ c + a - \frac{5}{4}ca + \frac{b}{2}\left(1 - \frac{c}{2}\right) \right\} - \frac{bc}{16}(b - a)(c - a) \end{bmatrix}.$$

It can be shown that the following parameter regions give simple incomplete bounds.

$\eta_3 \leq 0.20 \Rightarrow \tau_{C,i} < \tau_M$ for all i , while $\eta_3 > 0.252 \Rightarrow \tau_M < \tau_{C,i}$ for some i ;

$\eta_2 \leq 1.12 \Rightarrow \tau_{C,i} < \tau_M$ for all i , while $\eta_2 > 1.40 \Rightarrow \tau_M < \tau_{C,i}$ for some i ;

$\eta_1 < 1.9 \Rightarrow \tau_{C,i} < \tau_M$ for all i , while $\eta_1 > 2.2 \Rightarrow \tau_M < \tau_{C,i}$ for some i ;

$\tau_M \geq 2.243$ then $\tau_{C,i} < \tau_M$ for all i , while $\tau_M \leq 2.090$ then $\tau_M < \tau_{C,i}$ for some i .

5. General results

From Theorem 4.2 of [3] it is easy to see that for any irreducible m -state Markov chain,

$$\tau_M \geq \frac{(m-1)}{2}.$$

This is a generalisation of the results for 2-state and 3-state Markov chains as given earlier in this paper. The simplicity of this result, in contrast to the difficulty in obtaining simple expressions for the expected times to coupling, is certainly preferable as a measure of time to stationarity when one can be certain that the expected times to coupling are comparable.

In [3], it was established (Theorem 4.3) that for the special case of independent trials with m possible outcomes, $\tau_M = m - 1$. The three-state, case 4 model, can be generalized in this special situation.

Observe that in this m -state independent trials case with states $1, 2, \dots, m$ and p_i , the probability of outcome i , the elements of the Q matrix, given by (3.1), governing the transitions between state (i, j) to (r, s) are $p_r p_s$. Thus the elements of each row of Q are the same. By ordering the states of the coupled process (X_n, Y_n) in matched pairs as $(1,$

2), (2, 1), (1, 3), (3, 1), (4, 1), (1, 4), ..., (i, j), (j, i), ..., (m-1, m), (m, m-1), it is easy to express the equation $(I - Q)\mathbf{\kappa} = \mathbf{e}$, (Eqn. (3.9)) in element form to show that for the row corresponding to transitions from state (i, j) :

$$(1 - p_i p_j) \kappa_{ij}^{(C)} - p_i p_j \kappa_{ji}^{(C)} - \sum_{r \neq s, (r,s) \neq (i,j), (j,i)} p_r p_s \kappa_{rs}^{(C)} = 1. \quad (5.1)$$

Similarly from state (j, i) :

$$-p_i p_j \kappa_{ij}^{(C)} + (1 - p_i p_j) \kappa_{ji}^{(C)} - \sum_{r \neq s, (r,s) \neq (i,j), (j,i)} p_r p_s \kappa_{rs}^{(C)} = 1. \quad (5.2)$$

Subtract Eqn.(5.2) from Eqn. (5.1), to observe that for all $i \neq j$, $\kappa_{ij}^{(C)} - \kappa_{ji}^{(C)} = 0$, so that

$$\kappa_{ij}^{(C)} = \kappa_{ji}^{(C)}. \quad (5.3)$$

Further now adding Eqns. (5.1) and (5.2):

$$(1 - 2p_i p_j) \kappa_{ij}^{(C)} - 2p_i p_j \kappa_{ji}^{(C)} - 2 \sum_{r \neq s, (r,s) \neq (i,j), (j,i)} p_r p_s \kappa_{rs}^{(C)} = 2.$$

Utilising Eqn. (5.3) and taking, for simplicity, $i < j$,

$$(1 - 2p_i p_j) \kappa_{ij}^{(C)} - 2 \sum_{r < s, (r,s) \neq (i,j)} p_r p_s \kappa_{rs}^{(C)} = 1. \quad (5.4)$$

Thus we have reduced the $m(m-1)$ equations in $m(m-1)$ unknowns to $m(m-1)/2$ equations in $m(m-1)/2$ unknowns. Consider the first two of these reduced equations, (5.4):

$$(1 - 2p_1 p_2) \kappa_{12}^{(C)} - 2p_1 p_3 \kappa_{13}^{(C)} - \sum_{r < s, (r,s) \neq (1,2), (1,3)} p_r p_s \kappa_{rs}^{(C)} = 1, \quad (5.5)$$

$$-2p_1 p_2 \kappa_{12}^{(C)} - (1 - 2p_1 p_3) \kappa_{13}^{(C)} - \sum_{r < s, (r,s) \neq (1,2), (1,3)} p_r p_s \kappa_{rs}^{(C)} = 1. \quad (5.6)$$

Subtracting Eqn. (5.6) from (5.5) yields $\kappa_{12}^{(C)} - \kappa_{13}^{(C)} = 0$. This can be extended in general to show that for all $i < j$ (and hence all $i \neq j$), $\kappa_{ij}^{(C)} = \kappa_{12}^{(C)}$. Thus from Eqn. (5.4) we can deduce

$(1 - 2p_1 p_2) \kappa_{12}^{(C)} - 2 \sum_{r < s, (r,s) \neq (1,2)} p_r p_s \kappa_{rs}^{(C)} = 1$. Hence for all $i \neq j$, since $\sum_{k=1}^m p_k = 1$,

$$\kappa_{ij}^{(C)} = \frac{1}{1 - 2 \sum_{r < s} p_r p_s} = \frac{1}{1 - [(\sum_{k=1}^m p_k)^2 - (\sum_{k=1}^m p_k^2)]} = \frac{1}{\sum_{k=1}^m p_k^2}. \quad (5.7)$$

Now for independent trials the stationary probabilities p_i are simply the p_i . Thus from Eqn. (3.7), $\tau_{C,i} = \sum_{j \neq i} \pi_j \kappa_{ij}^{(C)}$, yielding

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

Observe that $\tau_{C,i} < \tau_{C,i} \Leftrightarrow p_j < p_i$, as noted for the special 3-state case.

When the independent trials model reduces to the simple ‘‘equally likely events’’ model with $p_1 = p_2 = \dots = p_m = 1/m$, then for each i , $\tau_{C,i} = m - 1 = \tau_M$. Thus the expected time

to coupling is the same as the expected time to mixing. Thus in a random “card shuffling” game the expected coupling and mixing times, are both 51 trials.

Note that of course this does not equate to $E[N]$, the expected time for all cards to appear which, from the “coupon collecting” problem, is given as

$$E[N] = n \left(\sum_{i=1}^n \frac{1}{i} \right) = 235.978 \text{ when } n = 52.$$

As in the 3-state case, it would be interesting to know under what conditions the expected time to coupling exceeds the expected time to mixing i.e. conditions on

$$\{p_i\}_{i=1}^m \text{ under which } \frac{1 - p_i}{\sum_{k=1}^m p_k^2} \geq m - 1.$$

Further, in a general Markov chain setting, it would be useful to have some general bounds for $\tau_{C,p}$ the expected time to coupling. As has been seen above these are not easy questions to answer.

References

1. Aldous D.J. and Fill J.A. *Reversible Markov Chains and Random Walks on Graphs* (Book in preparation) See <http://www.stat.Berkeley.EDU/users/aldous/book.html>
2. Hunter J.J. Variances of First Passage Times in a Markov chain with applications to Mixing Times, *Research Letters in Information and Mathematical Sciences, Institute of Information and Mathematical Sciences, Massey University, Auckland, New Zealand*, **10**, 17-48, (2006).(Submitted to *Linear Algebra Appl.*)
3. Hunter J.J. ‘Mixing times with applications to perturbed Markov chains, *Linear Algebra Appl.*, **417**, 108-123, (2006).
4. Hunter, J.J. *Mathematical Techniques of Applied Probability, Volume 1. Discrete Time Models: Basic Theory* Academic, New York, 1983.
5. Hunter, J.J. *Mathematical Techniques of Applied Probability, Volume 2, Discrete Time Models: Techniques and Applications*, Academic, New York, 1983.
6. Hunter J.J. ‘Generalized inverses and their application to applied probability problems’, *Linear Algebra Appl.*, **45**, 157-198, (1982).
7. Hunter J.J. ‘On the moments of Markov renewal processes’, *Adv. Appl. Probab.* **1**, 188-210 (1969)

8. Kemeny J. G. Snell J. L. *Finite Markov Chains*, Van Nostrand, New York, 1960.
9. Lovasz L. and Winkler P. 'Mixing Times', *Microsurveys in Discrete Probability* (Ed Aldous D. and Propp J.) DIMACS Series in Discret Math and Theor. Comp. Sci., 85-133, AMS, 1998.
10. Meyer C.D. 'The role of the group generalized inverse in the theory of finite Markov chains', *SIAM Rev.*, **17**, 443-464, (1975).