COALESCENCE IN MARKOV CHAINS

GEOFFREY R. GRIMMETT AND MARK HOLMES

In celebration of Svante Janson's 70th birthday

ABSTRACT. A Markov chain X^i on a finite state space S has transition matrix P and initial state i. We may run the chains $(X^i : i \in S)$ in parallel, while insisting that any two such chains coalesce whenever they are simultaneously at the same state. There are $|S|$ trajectories which evolve separately, but not necessarily independently, prior to coalescence. What can be said about the number $k(\mu)$ of coalescence classes of the process, and what is the set $K(P)$ of such numbers $k(\mu)$, as the coupling μ of the chains ranges over couplings that are consistent with P? We continue earlier work of the authors ('Non-coupling from the past', in *In and Out of Equilibrium 3*, Springer, 2021) on these two fundamental questions, which have special importance for the 'coupling from the past' algorithm.

We concentrate partly on a family of couplings termed block measures, which may be viewed as couplings of lumpable chains with coalescing lumps. Constructions of such couplings are presented, and also of non-block measure with similar properties.

1. INTRODUCTION

Finite-state space Markov chains form a key topic in probability theory, and they are taught in many undergraduate courses worldwide. They are considered to be well understood, and their theory fully established. Applications across science are significant and multifarious. In this paper we explore some interesting questions concerning coalescence that remain unresolved, and indeed in part unasked.

Throughout this paper, S is a finite, non-empty set (without loss of generality, we may take $S = \{1, 2, \dots, n\}$), and $P = (p_{i,j} : i, j \in S)$ is an irreducible stochastic matrix. Let $X = (X_t : t = 0, 1, \dots)$ be a

Date: 15 October 2025.

2010 *Mathematics Subject Classification.* 60J10, 60J22.

Key words and phrases. Markov chain, grand coupling, coupling from the past, CFTP, lumpability, block measure, coalescence time, coalescence number, avoidance coupling.

Markov chain on the state space S with transition matrix P. The long-term behaviour of X is well known: if X is (irreducible and) aperiodic, then P has a unique invariant distribution $\boldsymbol{\pi} = (\pi_j : j \in S)$, and $\mathbb{P}(X_t = j) \to \pi_j$ for $j \in S$, as $t \to \infty$. For introductions to the theory of Markov chains, the reader is referred to the books [9, 10, 14].

For $i \in S$, denote by $X^i = (X_t^i : t = 0, 1, \dots)$ the Markov chain $X = (X_t)$ conditioned on starting in state i, that is, conditioned on the event that $X_0 = i$. In the current article we explore the simultaneous evolution of the chains X^i, $i \in S$, in the setting where any two paths stay together after they meet; that is, if $X_t^i = X_t^j$ then $X_{t+s}^i = X_{t+s}^j$ for all $s \in \mathbb{N} := \{1, 2, 3, \dots\}$. For general P, multiple joint distributions of evolutions exist that are 'consistent' with P in the sense that they have the required marginal distributions.

In this setting, we say that X^i and X^j *coalesce* (or just that i and j coalesce) if they meet (that is, if $X_t^i = X_t^j$ for some t). The principal issue investigated here is to determine the degree of coalescence of the family $(X^i : i \in S)$ as $t \to \infty$. This question is made more explicit in Section 3, and is connected to earlier work on so-called avoidance coupling; see Remark 3.4.

The current work is a development of earlier work of the authors, [8], directed mainly at coalescence in so-called 'coupling from the past' (CFTP). CFTP is an important technique for exact simulation from a given distribution on a state space S, and is related to the topic of 'Monte Carlo Markov chains' (MCMC). Whereas MCMC runs forwards in time, CFTP runs backwards. The CFTP algorithm is successful in situations where all the chains (X^i) coalesce. We explain this connection in Remark 3.2. Some background information from [8] is included in this work as an aid to the reader.

Here is a summary of the contents of this article. A 'grand coupling' is a coupling of the chains $(X^i : i \in S)$. Such couplings are defined in Section 2, where the special 'independence coupling' is introduced. In Theorem 2.3 we answer the question of when the independence coupling is the unique grand coupling. Section 3 is devoted to forward coalescence. Lumpable chains and block measures are introduced in Section 4, and the associated main results appear in Sections 5 and 6. A condition for a lumpable chain to give rise to a block measure is presented in Theorem 5.3; conversely, the existence of non-block measures is explored in Theorem 6.1 in the special case of the transition matrix with equally probable transitions. In the final Section 7, we initiate an investigation of the structure of the set of grand couplings corresponding to transition matrices.

We close with a remark. The general questions approached here seem quite fundamental to a complete theory of finite state-space Markov chains, and yet they have received only limited attention so far. This is reflected in the partial state of knowledge revealed in [8] and the current paper. This work may be viewed as a contribution to the theory of iterated random functions.

We write $\mathbb{N}$ for the set $\{1, 2, \dots\}$ of natural numbers, and $\mathbb{P}$ for a generic probability measure.

2. Grand couplings of Markov chains

We shall consider only *irreducible* Markov chains on a finite state space S. Let $\mathcal{F}_S$ be the set of functions from S to S, and let $\mathcal{P}_S$ be the set of irreducible stochastic matrices on S.

Definition 2.1 ([8]). A probability measure μ on $\mathcal{F}_S$ is *consistent* with $P \in \mathcal{P}_S$, in which case we say that the pair (P, μ) is *consistent*, if

$$p_{i,j} = \mu\big(\{f \in \mathcal{F}_S : f(i) = j\}\big), \qquad i, j \in S. \tag{2.1}$$

Let $\mathcal{L}_P$ denote the set of probability measures μ on $\mathcal{F}_S$ that are consistent with $P \in \mathcal{P}_S$. ◀

Let $P \in \mathcal{P}_S$. Any probability measure $\mu \in \mathcal{L}_P$ may be used as the basis of a coupling of the chains $(X^i : i \in S)$, as we now demonstrate. Let $F_1, F_2, \dots$ be random functions that are independent and distributed on $\mathcal{F}_S$ according to μ, and set

$$X_t^i = F_t \circ F_{t-1} \circ \cdots \circ F_1(i), \qquad i \in S. \tag{2.2}$$

(This is a convenient abuse of notation; more properly, the right side of (2.2) defines processes having the same law as the sequence (X^i).) For $t \in \{0, 1, \dots\}$, let $Z_t = (X_t^i : i \in S)$. Then $Z = (Z_t : t = 0, 1, \dots)$ is a Markov chain with state space S^S, started in the state $(1, 2, \dots, n)$. The 'multichain' Z is in general reducible, since the number of distinct entries in Z_{t+1} is less than or equal to that in Z_t.

By the consistency property (2.1), the sequence $X^i = (X_t^i : t = 0, 1, \dots)$ has the same law as X conditioned on the event that $X_0 = i$. Moreover, if $X_t^i = X_t^j$ then trivially $X_{t+s}^i = X_{t+s}^j$ for all $s \in \mathbb{N}$ (since each F_t is a function); that is, once the paths from i and j meet, they stay together thenceforth.

A measure $\mu \in \mathcal{L}_P$ is sometimes called a *grand coupling* of P. It is elementary that $\mathcal{L}_P \neq \varnothing$, as indicated in the following example.

Example 2.2 (Independence coupling). Let $P \in \mathcal{P}_S$, and let μ_{ind} be given by the product form

$$\mu_{\text{ind}}(\{f\}) = \prod_{i \in S} p_{i,f(i)}, \qquad f \in \mathcal{F}_S.$$

Then

$$\begin{aligned}\mu_{\text{ind}}(\{f : f(i) = j\}) &= \sum_{f:f(i)=j} \prod_{u \in S} p_{u,f(u)} = p_{i,j} \sum_{f:f(i)=j} \prod_{u \neq i} p_{u,f(u)} \\ &= p_{i,j} \sum_{(j_1,\dots,j_{i-1},j_{i+1},\dots,j_n) \in S^{n-1}} \prod_{u \neq i} p_{u,j_u} \\ &= p_{i,j} \prod_{u \neq i} \Big[\sum_{j_u \in S} p_{u,j_u} \Big] = p_{i,j} \prod_{u \neq i} 1,\end{aligned}$$

so that $\mu_{\text{ind}} \in \mathcal{L}_P$. The measure μ_{ind} is called the *independence coupling* as it gives rise to $n = |S|$ chains X^i with transition matrix P, starting from each $i \in S$ respectively, that evolve independently until they meet. If P is aperiodic then, by Doeblin's theorem (see Theorem 2.5), for every pair $i, j \in S$ the chains X^i, X^j meet in finite time. Such chains stick together thenceforth, and thus all n chains a.s. meet in finite time (in some common random state). ◀

The next theorem contains a criterion for μ_{ind} to be the unique grand coupling.

Theorem 2.3. *For $P \in \mathcal{P}_S$, we have $|\mathcal{L}_P| \geq 2$ if and only if P has at least two rows each of which contains some entry lying in the open interval $(0, 1)$.*

Proof. It can be useful to think in terms of the transition diagram of the chain, i.e., the labelled directed graph G with vertex-set S, and with a directed edge from any i to any j such that $p_{i,j} > 0$; such an edge is labelled with the value $p_{i,j}$. The *in-degree* (respectively, *out-degree*) of a state is the number of edges directed towards it (respectively, away from it).

Let R be the set of rows that contain some entry in the interval $(0, 1)$. Since P is stochastic, such a row must contain at least two such entries.

(a) Let $|R| \leq 1$; we will show that the independence coupling is the unique grand coupling. We explain first the case $|R| = 0$, for which every row contains a single non-zero entry 1, and thus all edge-labels in G are 1. Each state i has out-degree 1, and we write $[i, t_i\rangle$ for the unique directed edge leading from i. There are exactly $n = |S|$ edges in G. Since P is irreducible, G is a directed self-avoiding cycle of length

n. Therefore, $\mathcal{L}_P$ contains only the probability measure $\mu = \delta_f$ that is supported on the single function f defined by $f : i \mapsto t_i$ for $i \in S$. In this trivial situation, μ is the independence coupling μ_{ind}.

Suppose that $|R| = 1$, and assume without loss of generality that the row labelled 1 is the unique row containing an entry in $(0, 1)$. Let $J = \{j : p_{1,j} \in (0, 1)\}$. Edges of G of the form $[1, j\rangle$ for $j \in J$ have labels in $(0, 1)$ and all other edges are labelled 1. For $i \neq 1$, write $[i, t_i\rangle$ for the unique edge directed away from i. We have

$$\mathbb{P}(X_1^1 = j) = p_{i,j} \quad \text{for } j \in J, \tag{2.3}$$

$$\mathbb{P}(X_1^i = t_i) = 1 \qquad \text{for } i \neq 1. \tag{2.4}$$

There is a unique probability measure μ consistent with the above, namely

$$\mu(\{f\}) = p_{1,f(1)} \prod_{i \neq 1} \delta_{f(i),t_i}, \qquad f \in \mathcal{F}_S, \tag{2.5}$$

where δ is the Kronecker delta. This is simply the independence coupling. Therefore, $|\mathcal{L}_P| = 1$.

(b) Conversely, suppose $|R| \geq 2$, and pick two rows lying in R, say the ith and jth with $i \neq j$. Find r, s such that $p_{i,r}p_{j,s} > 0$. There is no loss of generality for the proof that follows if we assume $i = 1$ and $j = 2$, and we assume this henceforth. We may assume further that

$$0 < p_{2,s} \leq p_{1,r} < 1. \tag{2.6}$$

Before defining μ explicitly, we describe the stochastic evolution of the first time-step of the family $(X^i : i \in S)$ of Markov chains.

For $i \in \{2, 3, \dots, n\}$, we let X_1^i be chosen according to P, that is,

$$\mathbb{P}(X_1^i = j) = p_{i,j}, \qquad j \in S. \tag{2.7}$$

Furthermore, $Z := \{X_1^i : i = 2, 3, \dots\}$ is a set of independent random variables.

We turn to X_1^1, which we take to be independent of $Z \setminus \{X_1^2\}$. Furthermore,

$$\mathbb{P}(X_1^1 = r \mid X_1^2 = s) = 1, \tag{2.8}$$

$$\mathbb{P}(X_1^1 = j \mid X_1^2 \neq s) = \begin{cases} \dfrac{p_{1,r} - p_{2,s}}{1 - p_{2,s}} & \text{if } j = r, \\ \dfrac{p_{1,j}}{1 - p_{2,s}} & \text{if } j \neq r. \end{cases} \tag{2.9}$$

It is an exercise in conditional probability that the mass function of X_1^1 is $(p_{1,j} : j \in S)$. Equations (2.8)–(2.9) amount to a coupling of X_1^1 and X_1^2 under which (2.8) holds. Since $p_{1,r} < 1$, (2.8) fails by (2.6)

under the independence coupling μ_{ind}. Therefore, $\mu \neq \mu_{\text{ind}}$ and hence $|\mathcal{L}_P| \geq 2$.

Equations (2.7)–(2.9) may be expressed by defining μ as follows:

$$\mu(\{f\}) = \begin{cases} p_{2,s} \prod_{i \neq 1,2} p_{i,f(i)} & \text{if } f(1) = r,\ f(2) = s, \\ \dfrac{p_{1,r} - p_{2,s}}{1 - p_{2,s}} \prod_{i \neq 1} p_{i,f(i)} & \text{if } f(1) = r,\ f(2) \neq s, \\ \dfrac{1}{1 - p_{2,s}} \prod_{i} p_{i,f(i)} & \text{if } f(1) \neq r,\ f(2) \neq s, \\ 0 & \text{otherwise.} \end{cases} \tag{2.10}$$

One may check directly that $\mu \in \mathcal{L}_P$, but it is quicker to verify the probabilities implied by (2.7)–(2.9). ■

Remark 2.4 (Random transition matrix). By Theorem 2.3, a 'typical' transition matrix P has multiple grand couplings. We make this statement more precise as follows. For given S, we model a 'typical' transition matrix as the $|S| \times |S|$ matrix with elements

$$p_{i,j} = q_{i,j}/Q_i,$$

where the $q_{i,j}$ are independent and uniformly distributed on $(0, 1)$, and $Q_i = \sum_j q_{i,j}$. Let $\mathbb{Q}$ denote the law of such P. Note that $\mathbb{Q}$-a.e. P has all entries in $(0, 1)$ and hence is irreducible and aperiodic. Moreover, by Theorem 2.3, $\mathbb{Q}$-a.e. P admits multiple grand couplings. Further results for a random transition matrix may be found in Remark 3.7, Lemma 5.1, and Section 7. ◀

We will make frequent use of the following well known fact due to Doeblin [5] (see also [9, p. 260]).

Theorem 2.5 ([5]). *Let $(X^i : i \in S)$ be* independent *Markov chains on a finite state space S with common irreducible, aperiodic transition matrix Q, with $X^i_0 = i$ a.s. for each $i \in S$. For $i, j \in S$, there exists a.s. a finite time T such that $X^i_T = X^j_T$.*

The following classical result of Birkhoff and von Neumann will be useful later.

Theorem 2.6 ([3, 13]). *A stochastic matrix P on the finite state space S is doubly stochastic if and only if it lies in the convex hull of the set Π_S of permutation matrices.*

3. Coalescence of trajectories

Consider a Markov chain X on the state space $S = \{1, 2, \dots, n\}$ with irreducible transition matrix P, and let $\mu \in \mathcal{L}_P$ be a grand coupling.

We consider the coalescence of the chains X^i in this section, and begin with some notation from [8].

As in Section 2 (see (2.2)), let $F_1, F_2, \dots$ be independent random functions distributed according to μ, and set $X_t^i = \vec{F}_t(i)$ where $i \in S$ and

$$\vec{F}_t := F_t \circ F_{t-1} \circ \cdots \circ F_1. \tag{3.1}$$

Then $(X^i : i \in S)$ is a family of coupled Markov chains running forwards in time, each having transition matrix P, and such that X^i starts in state i.

For $i, j \in S$, let

$$T_{i,j} = \inf\{t : X_t^i = X_t^j\}. \tag{3.2}$$

We say that i and j *coalesce* (and write $i \sim j$) if $T_{i,j} < \infty$. The *forward coalescence time* is given by

$$T = \inf\big\{t : \vec{F}_t(\cdot) \text{ is a constant function}\big\} = \sup_{i,j \in S} T_{i,j}. \tag{3.3}$$

We say that *coalescence occurs* if $\mathbb{P}(T < \infty) = 1$.

Lemma 3.1. *The relation $\sim$ is a (random) equivalence relation.*

Proof. It suffices to prove the transitivity of $\sim$. Let $i \sim j$ and $j \sim k$. Since S is finite, there exists T such that $X_t^i = X_t^j = X_t^k$ for $t \geq T$. The claim follows ∎

The equivalence classes of $\sim$ are termed the *coalescence classes* of $(X^i : i \in S)$. The number of coalescence classes is a.s. constant (see [8, Lemma 2]), and we denote this number by $k(\mu)$ and call it the *coalescence number* of μ. We define

$$K(P) = \{k(\mu) : \mu \in \mathcal{L}_P\}. \tag{3.4}$$

Remark 3.2 (Coupling from the past (CFTP)). CFTP is a prominent algorithm introduced by Propp and Wilson [15, 16, 18] for perfect simulation from the invariant distribution of an irreducible, aperiodic Markov chain on a finite state space. The CFTP algorithm may be defined by replacing the function $\vec{F}_t$ in (3.1) by the function $\overleftarrow{F}_t := F_1 \circ F_2 \circ \cdots \circ F_t$. The *backward coalescence time* is defined by

$$C = \inf\big\{t : \overleftarrow{F}_t(\cdot) \text{ is a constant function}\big\}, \tag{3.5}$$

and *backward coalescence* is said to occur if $\mathbb{P}(C < \infty)$. On the event $\{C < \infty\}$, $\overleftarrow{F}_C$ may be regarded as a random state, and the main theorem of CFTP asserts that, if backward coalescence occurs, then

$\overleftarrow{F}_C$ is distributed as the invariant distribution of the transition matrix P.

The relationship between forward and backward coalescence was explored in [8]. It turns out that C and T are identically distributed. Furthermore, the CFTP process has the same set $K(P)$ of coalescence numbers as the forward process. In contrast, there is a significant difference between forward and backward coalescence in that, if $\vec{F}_t(i) = \vec{F}_t(j)$, then $\vec{F}_{t+1}(i) = \vec{F}_{t+1}(j)$, whereas the corresponding statement for backward coalescence is false in general. The last occurs whenever the coalescing classes in the forward direction are nondeterministic (see, for example, the forthcoming Examples 3.10 and 4.5(b)). ◀

The following two questions are fundamental to understanding coalescence.

Question 3.3.

1. *Can we determine the set $K(P)$ for given P?*
2. *Which $\mu \in \mathcal{L}_P$ have $k(\mu) = 1$?*

We shall see in Theorem 3.13 that the independence coupling μ_{ind} of Example 2.2 satisfies $k(\mu_{\text{ind}}) \le k(\mu)$ for all $\mu \in \mathcal{L}_P$, and moreover $k(\mu_{\text{ind}}) = 1$ if and only if P is aperiodic.

Henceforth, expressions involving the word 'coalescence' shall refer to forward *coalescence.* Let μ be a probability measure on $\mathcal{F}_S$, and let $\text{supp}(\mu)$ denote the support of μ. Let $F = (F_s : s \in \mathbb{N})$ be a vector of independent and identically distributed random functions, each with law μ. The law of F is the product measure $\boldsymbol{\mu} = \prod_{i \in \mathbb{N}} \mu$.

Remark 3.4 (Avoidance coupling). Let $P \in \mathcal{P}_S$ and let

$$k_{\max} = k_{\max}(P) := \max\{k(\mu) : \mu \in \mathcal{L}_P\}.$$

That is, $k_{\max}$ is the maximum k such that: there exists some grand coupling μ for which there exist k (possibly random) initial states whose trajectories avoid one another for all time. The identification of $k_{\max}$ might be termed the 'avoidance problem with simultaneous updating'. The related problem of avoidance coupling with *sequential* updating was initiated in [1] for random walk on a graph and has been developed further by others (see, for example, [2]). ◀

Recall that, since P is finite and irreducible, it has a *unique* invariant distribution (see, for example, [9, Thm 6.4.3]).

Theorem 3.5. *Let $P \in \mathcal{P}_S$ have (unique) invariant distribution $\boldsymbol{\pi}$.*

(a) *Let $m \in \mathbb{N}$, and suppose there exists $s \in S$ with $\pi_s > 1/m$. Then $k_{\max} < m$.*
(b) *Let $m \in \mathbb{N}$, and suppose there exists $s \in S$ such that $p_{i,s} > 1/m$ for all $i \in S$, then $k_{\max} < m$.*

Proof. (a) Assume $\pi_s > 1/m$, and let $i_1, i_2, \dots, i_m$ be distinct elements of S. For each $k \in \{1, 2, \dots, m\}$, there is asymptotic density π_s of times n at which $X_n^{i_k} = s$. Since $\pi_s > 1/m$, there exist (a.s.) distinct $i, j \in \{i_1, i_2, \dots, i_m\}$ such that there is a strictly positive density of times n at which $X_n^i = X_n^j = s$. Therefore, $k(\mu) < m$ for all $\mu \in \mathcal{L}_P$.

(b) Assume the given condition holds. Then

$$\pi_s = \sum_{i \in S} \pi_i p_{i,s} > \frac{1}{m}. \tag{3.6}$$

The conclusion holds by part (a). ■

Remark 3.6. The condition of part (b) may be changed slightly, as follows. Suppose $p_{i,s} > 1/m$ for all $i \in S$ with $i \neq s$. Then (3.6) becomes

$$\pi_s \geq \sum_{i \neq s} \pi_i p_{i,s} > \frac{1}{m}(1 - \pi_s).$$

Therefore, $\pi_s > 1/(m+1)$, whence $k_{\max} < m + 1$ by part (a). ◀

Remark 3.7 (Random transition matrix). By Theorem 3.5, if there exists $s \in S$ with $\pi_s > \frac{1}{2}$, then $k(\mu) = 1$ for all $\mu \in \mathcal{L}_P$. In particular, there is strictly positive $\mathbb{Q}$-probability (see Remark 2.4) that a random transition matrix P satisfies $K(P) = \{1\}$. ◀

Question 3.8. *Is it true that $K(P) = \{1\}$ for $\mathbb{Q}$-a.e. P?*

Whereas $k(\mu)$ is a.s. constant, the coalescence classes of $\sim$ need not themselves be a.s. constant. Here is an example of this, preceded by some notation.

Definition 3.9. Let $f \in \mathcal{F}_S$ where $S = \{1, 2, \dots, n\}$. We write $f = (j_1 j_2 \dots j_n)$ if $f(r) = j_r$ for $r = 1, 2, \dots, n$. ◀

Example 3.10. Take $S = \{1, 2, 3, 4\}$ and any consistent pair (P, μ) with

$$\operatorname{supp}(\mu) = \{f_{1,1}, f_{1,2}, f_{2,1}, f_{2,2}\}$$

where

$$f_{1,1} = (1212), \quad f_{1,2} = (1221), \quad f_{2,1} = (3434), \quad f_{2,2} = (3443).$$

Then $k(\mu) = 2$ but the coalescence classes of $\vec{F}$ may be either $\{1, 3\}$, $\{2, 4\}$ or $\{1, 4\}$, $\{2, 3\}$, each having a strictly positive probability. The

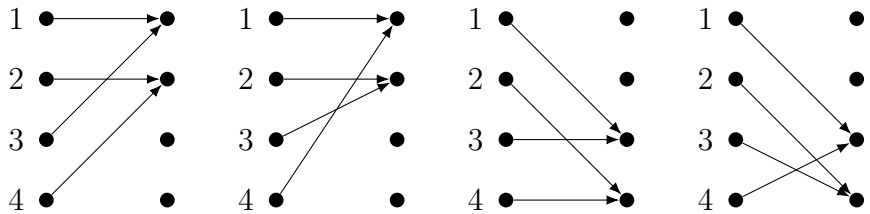

FIGURE 3.1. Diagrammatic representations of the four functions $f_{i,j}$ of Example 3.10.

value of F_1 determines which of these two possibilities occurs.[1] The four functions $f_{i,j}$ are illustrated in Figure 3.1. ◀

Question 3.11. *For given $P \in \mathcal{P}_S$, is the set of cardinalities (possibly with repetition) of coalescing classes a deterministic set?*

A probability measure μ on $\mathcal{F}_S$ may be written in the form

$$\mu = \sum_{f \in \mathcal{F}_S} \alpha_f \delta_f, \tag{3.7}$$

where $\boldsymbol{\alpha}$ is a probability mass function on $\mathcal{F}_S$ with support $\mathrm{supp}(\mu)$, and δ_f is the Dirac delta-mass on the point $f \in \mathcal{F}_S$. Thus, $\alpha_f > 0$ if and only if $f \in \mathrm{supp}(\mu)$. We denote by $M(\mathcal{F}_S)$ the set of all probability measures μ of the form (3.7).

In advance of stating an extension of [8, Lemma 3], we remind the reader of the definition of a cyclic class.

Definition 3.12. Consider an irreducible Markov chain on the state space S with transition matrix P and period d. There exists a unique partition of S into d classes $S_0, S_1, \ldots, S_{d-1}$ such that

$$\sum_{j \in S_{r+1}} p_{i,j} = 1, \qquad i \in S_r,\ r = 0, 1, \ldots, d-1,$$

where by convention $S_0 = S_d$. The sets S_r are called the *cyclic classes* of the chain. ◀

See [4, Sect. 2.3.2] for further details of cyclic classes.

Theorem 3.13.

(a) *We have that $1 \in K(P)$ if and only if $P \in \mathcal{P}_S$ is aperiodic. In this case, $k(\mu_{\mathrm{ind}}) = 1$.*

[1] In the case of backward coalescence, for $r \in \{1,2\}^2$, the 'first' function to be applied is f_r a.s. infinitely often, whence $\overleftarrow{F}_t(1) = \overleftarrow{F}_t(3)$ infinitely often and $\overleftarrow{F}_t(1) \neq \overleftarrow{F}_t(3)$ infinitely often.

(b) *If $P \in \mathcal{P}_S$ has period d, then the coalescence classes of μ_{ind} are a.s. the cyclic classes of P and therefore $k(\mu_{\text{ind}}) = d$. Hence, $d \in K(P)$, and moreover $d \le k$ for all $k \in K(P)$.*

Consider the particular case when all entries of P are strictly positive. Since P is aperiodic, we have by Theorem 3.13(a) that $1 \in K(P)$. However, $K(P)$ may be larger than a singleton; see the forthcoming Theorem 6.1.

Proof of Theorem 3.13. (a) This is found at the proof of [8, Lemma 3], but is included here for clarity. The independence coupling of Example 2.2 gives rise to $n = |S|$ chains with transition matrix P, starting from $1, 2, \dots, n$, respectively, that evolve independently until they meet. If P is aperiodic (and, by assumption, irreducible) then all n chains meet a.s. in finite time. The last holds by Theorem 2.5, since any paths that meet will remain together thenceforth under this coupling.

Conversely, if P is periodic and $p_{i,j} > 0$ then $i \neq j$, and i and j can never coalesce, implying $1 \notin K(P)$.

(b) Suppose P has period $d \ge 2$, and let $S_0, S_1, \dots, S_{d-1}$ be the cyclic classes of P (see Definition 3.12). For $r \neq s$, states $i \in S_r$ and $j \in S_s$ do not coalesce for any $\mu \in \mathcal{L}_P$. Now, P^d has a block diagonal form with (transition) matrices $E_0, E_1, \dots, E_{d-1}$ along its diagonal. Each E_r is irreducible and aperiodic. By the argument above, subject to the independence coupling μ_{ind}, any two states in any given S_r coalesce a.s.

For clarity, we give some more details of the last step. Fix r and let $i, j \in S_r$. Let $Z^k = (Z^k_t : t = 0, 1, \dots)$ be a copy of X^k, and suppose the chains $(Z^k : k \in S_r)$ evolve independently (unlike the sequence (X^k) whose members coalesce when they meet). Now, $Z^{k,d} := (Z^k_{md} : m = 0, 1, \dots)$ has irreducible, aperiodic transition matrix E_r. By Theorem 2.5 there exists a.s. some time $T < \infty$ at which $Z^i_{Td} = Z^j_{Td}$. Therefore, there is some earliest time $U \le Td$ at which $Z^i_U = Z^j_U$. From the pair Z^i, Z^j we construct X^i, X^j by

$$(X^i_t, X^j_t) = \begin{cases} (Z^i_t, Z^j_t) & \text{for } t \le U, \\ (Z^i_t, Z^i_t) & \text{for } t > U. \end{cases}$$

Since the pair (X^i, X^j) has the same law as under $\boldsymbol{\mu}$, this proves that i and j coalesce a.s. under $\boldsymbol{\mu}$. This holds for all pairs of distinct states in S_r, and two states stick together after they coaleasce. The second claim of the theorem is proved.

The minimality of d holds since no two states in distinct cyclic classes may coalesce. ∎

4. Lumpability and block measures

The notion of lumpability was introduced in 1963 by Kemeny and Snell, [11]. A Markov chain X with finite state space S is called lumpable if there exists a partition $\mathcal{S}$ of S such that the projection of X onto $\mathcal{S}$ is itself a Markov chain.

Here is a more precise definition. A partition $\mathcal{S} = \{S_1, S_2, \dots, S_\ell\}$ of S is called *trivial* if $|\mathcal{S}| \in \{1, |S|\}$, (i.e., if either $\mathcal{S} = \{S\}$ or $\mathcal{S} = \{\{s\} : s \in S\}$). Elements of the partition $\mathcal{S}$ are called *blocks*. The chain X is called $\mathcal{S}$-*lumpable* if the sequence $(W_t : t = 0, 1, \dots)$, defined by

$$W_t = j \quad \text{if} \quad X_t \in S_j, \tag{4.1}$$

forms a Markov chain. The process X is called *lumpable* if it is $\mathcal{S}$-lumpable for some non-trivial partition $\mathcal{S}$.

There is a limited literature on lumpable chains, for which the reader may consult, for example, [6, 7, 12] and [9, Exer. 6.1.13]. Kemeny and Snell [11] proved a necessary and sufficient condition for X to be $\mathcal{S}$-lumpable, namely the following.

Theorem 4.1 ([11]). *Let $P = (p_{i,j} : i, j \in S)$. Let $\mathcal{S}$ be a partition of S, and let*

$$\lambda_{r,s}^{(i)} = \sum_{j \in S_s} p_{i,j}, \quad i \in S_r. \tag{4.2}$$

A Markov chain X with transition matrix P is $\mathcal{S}$-lumpable if and only if,

(4.3) *for every r, s, we have that $\lambda_{r,s} := \lambda_{r,s}^{(i)}$ is constant for $i \in S_r$.*

A stochastic matrix $P \in \mathcal{P}_S$ is called $\mathcal{S}$-lumpable if (4.3) holds (where the $\lambda_{r,s}^{(i)}$ are given in (4.2)). For such a pair P and $\mathcal{S}$, let Λ be the $\ell \times \ell$ matrix $(\lambda_{r,s} : 1 \le r, s \le \ell)$. This Λ is the transition matrix of the 'block process' W of (4.1).

The evolution of an $\mathcal{S}$-lumpable chain X may be given in two stages. Suppose $X_t = i \in S_r$; then X_{t+1} is given as follows.

(a) Select a random block B with mass function $\mathbb{P}(B = S_s) = \lambda_{r,s}$.
(b) Conditional on $B = S_s$, choose X_{t+1} with mass function

$$\mathbb{P}(X_{t+1} = j \mid B = S_s) = p_{i,j}/\lambda_{r,s}.$$

We address a certain sub-category of lumpable chains here, namely those for which Λ is doubly stochastic. By Theorem 2.6, such Λ may be expressed as a convex combination of permutation matrices.

Recall the set $M(\mathcal{F}_S)$ of probability measures of the form (3.7).

Definition 4.2 ([8]). Let $\mu \in M(\mathcal{F}_S)$. For a partition $\mathcal{S} = \{S_r : r \in I\}$ of S with index set $I = \{1, 2, \ldots, \ell\}$, we call μ an *$\mathcal{S}$-block measure* if

(a) for $f \in \mathrm{supp}(\mu)$, there exists a unique permutation $\pi = \pi_f$ of I such that, for $r \in I$, $fS_r \subseteq S_{\pi(r)}$, and
(b) $k(\mu) = \ell$.

An $\mathcal{S}$-block measure μ is said to be *trivial* if $\mathcal{S}$ is trivial. A measure μ is called a *block measure* if it is an $\mathcal{S}$-block measure for some partition $\mathcal{S}$. ◀

Since any two states in distinct blocks cannot coalesce under a block measure, the condition $k(\mu) = \ell$ implies that

$$\text{for } r \in I \text{ and } i, j \in S_r, \text{ the pair } i,\ j \text{ coalesce a.s.}, \tag{4.4}$$

so that the coalescence classes of the coalescence relation $\sim$ are a.s. the blocks $S_1, S_2, \ldots, S_\ell$.

By combining the definitions of lumpability and block measures we arrive at a necessary condition for μ to be an $\mathcal{S}$-block measure. The proof may be found in [8].

Theorem 4.3 ([8, Thm 6])**.** *Let S be a non-empty, finite set, let $P \in \mathcal{P}_S$, and let $\mathcal{S} = \{S_r : r \in I\}$ be a partition of S with index set $I = \{1, 2, \ldots, \ell\}$. For $i \in S_r$, let*

$$\lambda_{r,s}^{(i)} = \sum_{j \in S_s} p_{i,j}. \tag{4.5}$$

(a) *If $\mu \in \mathcal{L}_P$ is an $\mathcal{S}$-block measure, then, for $r, s \in I$,*

$$\lambda_{r,s} := \lambda_{r,s}^{(i)} \text{ is constant for } i \in S_r. \tag{4.6}$$

(b) *The 'block-transition matrix' $\Lambda = (\lambda_{r,s} : 1 \le r, s \le \ell)$ is irreducible. Moreover, Λ is doubly stochastic, which may be expressed as*

$$\sum_{i \in S} \sum_{j \in S_s} \frac{1}{|S_{r(i)}|} p_{i,j} = 1, \qquad r, s \in I, \tag{4.7}$$

where $r(i)$ is the index $r \in I$ such that $i \in S_r$.

Whether or not $\mu \in M(\mathcal{F}_S)$ is a block measure turns out to depend on whether or not its coalescing classes are deterministic, that is, a.s. constant. (Recall the definition of coalescing class after Lemma 3.1.) We state this as a theorem.

Theorem 4.4. *A probability measure $\mu \in M(\mathcal{F}_S)$ is a block measure if and only if its coalescing classes $\mathcal{C} = \{C_1, C_2, \ldots, C_\ell\}$ are a.s. constant. If the last holds, μ is a $\mathcal{C}$-block measure.*

Proof. If $\mu \in M(\mathcal{F}_S)$ is an $\mathcal{S}$-block measure then (recall (4.4) and the preceding discussion) the coalescing classes are a.s. the elements of $\mathcal{S}$.

Conversely, let the coalescing classes be $C_1, \dots, C_\ell$, and assume they are a.s. constant. We claim that Definition 4.2 is satisfied with $\mathcal{S} = \mathcal{C}$. Clearly $k(\mu) = \ell$, so (b) of Definition 4.2 holds. Let $f \in \mathrm{supp}(\mu)$, and let $i_1, i_2 \in C_j$. Since i_1 and i_2 a.s. coalesce, $f(i_1)$ and $f(i_2)$ a.s. coalesce, so $f(i_1) \in C_s \iff f(i_2) \in C_s$. Suppose instead that $i_1 \in C_{j_1}$ and $i_2 \in C_{j_2}$ where $j_1 \neq j_2$. This implies that i_1 and i_2 cannot coalesce, and hence $f(i_1) \in C_s$ implies $f(i_2) \notin C_s$.

We have shown that f permutes classes, which verifies Definition 4.2(a) with $\mathcal{S} = \mathcal{C}$, and hence completes the proof. ∎

Example 4.5. Here is an illustration of Theorem 4.4. Consider the transition matrix

$$P = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix},$$

and note that it is irreducible and aperiodic.

(a) Let μ_f put mass $\frac{1}{2}$ on each of the two permutations

$$f_1 = (1234), \qquad f_2 = (3421).$$

(Recall Definition 3.9.) It is easily checked that $\mu_f \in \mathcal{L}_P$ is a (trivial) block measure with blocks $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, so that $k(\mu_f) = 4$.

(b) Let μ_g put mass $\frac{1}{2}$ on each of the two functions

$$g_1 = (3434), \qquad g_2 = (1221).$$

Then $\mu_g \in \mathcal{L}_P$. Under μ_g either: 1 and 3 coalesce, *and* 2 and 4 coalesce *and there is no other coalescence* (this happens if g_1 is applied first) or: 2 and 3 coalesce, *and* 1 and 4 coalesce *and there is no other coalescence* (this happens if g_2 is applied first). Therefore, $k(\mu_g) = 2$, but the coalescence sets of μ_g are not a.s. constant. By Theorem 4.4, μ_g is not a block measure. This case contains the essence of Example 3.10. ◀

5. Existence of block measures

This section is concerned with the existence of block measures. The main question of interest is, for what state space S, partition $\mathcal{S}$, and matrix $P \in \mathcal{P}_S$ does there exist an $\mathcal{S}$-block measure $\mu \in \mathcal{L}_P$? Let $\mathcal{C} = \mathcal{C}(\mu)$ denote the set of coalescing classes of μ. In general $\mathcal{C}(\mu)$ is random, and the support of $\mathcal{C}(\mu)$ depends only on $\mathrm{supp}(\mu)$. By

Theorem 4.4, finding an $\mathcal{S}$-block measure $\mu \in \mathcal{L}_P$ is equivalent to finding $\mu \in \mathcal{L}_P$ for which $\mathcal{C} = \mathcal{S}$ a.s.

By Theorem 3.13, for every $P \in \mathcal{P}_S$ there exists a partition $\mathcal{S}$ of S (comprising the cyclic classes of P) such that the independence coupling μ_{ind} is an $\mathcal{S}$-block measure. Recall the probability measure $\mathbb{Q}$ in Remark 2.4.

Lemma 5.1. *For $\mathbb{Q}$-a.e. $P \in \mathcal{P}_S$, every block measure is trivial.*

Proof. We need only consider cases with $|S| \geq 3$. If there exists a non-trivial block measure, then there exists $T \subseteq S$ such that $2 \leq |T| \leq |S| - 1$ and, by (4.5)–(4.6),

$$\sum_{j \in T} p_{i,j} \text{ is constant for } i \in T.$$

Let $Z_i := \sum_{j \in T} p_{i,j}$. Then $Z_1, Z_2, \ldots$ are independent with a common absolutely continuous distribution. Therefore, $\mathbb{Q}(Z_i = Z_k) = 0$ for all distinct pairs i, k, whence the set of such P has $\mathbb{Q}$-measure 0. ■

By Lemma 5.1, for every S, for $\mathbb{Q}$-a.s. $P \in \mathcal{P}_S$, and for every non-trivial $\mathcal{S}$, there exists no $\mathcal{S}$-block measure. In contrast, the following holds.

Theorem 5.2. *For $S = \{1, 2, \ldots, n\}$ and any partition $\mathcal{S}$ of S there exists $\mu \in M(\mathcal{F}_S)$ such that μ is an $\mathcal{S}$-block measure.*

Proof. Fix S and $\mathcal{S} = \{S_r : r \in I\}$. Consider the set $\widehat{\mathcal{F}}$ of functions f such that

(i) for all r, $f(i) = f(j)$ for every $i, j \in S_r$, and
(ii) there exists a permutation $\pi = \pi_f$ of I such that, for $r \in I$, $fS_r \subseteq S_{\pi(r)}$.

For example, if $\mathcal{S} = \{\{1, 2\}, \{3, 4, 5\}\}$ then $\widehat{\mathcal{F}}$ consists of the 12 functions denoted

$$(11333), (11444), (11555), (22333), (22444), (22555),$$
$$(33111), (44111), (55111), (33222), (44222), (55222),$$

in the notation of Definition 3.9.

Let μ be a probability measure with $\text{supp}(\mu) = \widehat{\mathcal{F}}$. Then $p_{i,j} := \sum_{f: f(i)=j} \mu(\{f\}) > 0$ for every pair i, j, whence $P = (p_{i,j})$ is irreducible. The claim follows by assumptions (i) and (ii), and Definition 4.2. ■

Henceforth, for certain P, $\mathcal{S}$, we will present a natural measure $\mu \in \mathcal{L}_P$, and then determine conditions under which it is an $\mathcal{S}$-block measure.

Let the pair $P \in \mathcal{P}_S$ and $\mathcal{S} = \{S_r : r \in I\}$ (a partition of the state space S with $I = \{1, 2, \dots, \ell\}$) satisfy (4.6) and be such that

$$\Lambda \text{ is doubly stochastic.} \tag{5.1}$$

By Theorem 4.3 these two conditions are necessary for the existence of an $\mathcal{S}$-block measure $\mu \in \mathcal{L}_P$.

Since P is assumed irreducible, the stochastic matrix Λ is irreducible also. By (5.1) and Theorem 2.6, we may find a measure $\rho \in \mathcal{L}_\Lambda$ supported on a subset of the set of permutations of I, and we let Π be a random permutation with law ρ. (Note that ρ is not generally unique.) Let $i \in S_r$ and $j \in S_s$. In order for our forthcoming μ to be consistent with P (recall Definition 2.1) we need that

$$\sum_{f : i \mapsto j} \mu(\{f\}) = p_{i,j}. \tag{5.2}$$

In order for i to be mapped to j, it is necessary that $\Pi(r) = s$; the last occurs with probability $\lambda_{r,s}$. Conditional on Π, we shall then map the states independently in such a way as to obtain (5.2).

More precisely, conditional on Π, let $Z = (Z_i : i \in S)$ be independent random variables such that

$$\mathbb{P}(Z_i = j \mid \Pi) = \begin{cases} p_{i,j}/\lambda_{r,s} & \text{if } S_r \ni i,\ S_s \ni j,\ \Pi(r) = s, \\ 0 & \text{otherwise,} \end{cases} \tag{5.3}$$

and let μ be the law of Z. Thus, $\mu \in M(\mathcal{F}_S)$ is given by

$$\mu(\{f\}) = \mathbb{E}\left[\prod_{i \in S} \mathbb{P}(Z_i = f(i) \mid \Pi) \right], \tag{5.4}$$

where the expectation is over the random permutation Π. We call μ the $(P, \mathcal{S}, \rho)$-*product measure*.

Next we check that $\mu \in \mathcal{L}_P$. Let $i \in S_r$ and $j \in S_s$. By (5.3)–(5.4),

$$\sum_{f : i \mapsto j} \mu(\{f\}) = \mathbb{P}(Z_i = j) = \lambda_{r,s} \cdot \frac{p_{i,j}}{\lambda_{r,s}} = p_{i,j},$$

as required. By the definition of μ, no two states in different sets of the partition $\mathcal{S}$ may coalesce. Thus, to determine whether the $(P, \mathcal{S}, \rho)$-product measure $\mu \in \mathcal{L}_P$ is an $\mathcal{S}$-block measure it remains to check whether $k(\mu) = \ell = |\mathcal{S}|$, which is to say that, for all r, all states in S_r coalesce (recall (4.4)).

Recall that $\mathcal{C}$ denotes the (possibly random) set of coalescing classes, and that $i \sim j$ if there exists $C \in \mathcal{C}$ such that $i, j \in C$.

Theorem 5.3. *Let $P \in \mathcal{P}_S$, and let $\mathcal{S} = \{S_r : r \in I\}$ be a partition of S with index set $I = \{1, 2, \dots, \ell\}$. Assume (4.6) and (5.1) hold. The $(P, \mathcal{S}, \rho)$-product measure μ satisfies $k(\mu) = \ell$ if and only if*[2]

$$\mathbb{P}(i \sim j) > 0, \qquad i, j \in S_1. \tag{5.5}$$

Proof. Suppose the conditions of the theorem hold, and also (5.5). Let $r \in I$. We will show that

$$\mathbb{P}(i \sim j) = 1, \qquad i, j \in S_r. \tag{5.6}$$

The claim $k(\mu) = \ell$ follows since there are finitely many pairs $i, j \in S_r$ and indices r.

Recall from (3.2) that $T_{i,j} = \inf\{t \geq 0 : X_t^i = X_t^j\}$, so that $i \sim j$ if and only if $T_{i,j} < \infty$.

We first prove (5.6) with $r = 1$. Since $|S_1| < \infty$, by (5.5) there exists $\epsilon > 0$, and $M < \infty$ such that

$$\mathbb{P}(T_{i,j} \leq M) > \epsilon, \qquad i, j \in S_1. \tag{5.7}$$

Let $i, j \in S_1$. Either $T_{i,j} \leq M$ or not. Assume that $T_{i,j} > M$. We continue the two chains from time M with initial states X_M^i and X_M^j until the next epoch ($M + K$, say) at which these two processes lie in S_1; note that $\mathbb{P}(K < \infty) = 1$ since Λ is irreducible. Having arrived back in S_1, we apply the argument above to deduce that coalescence occurs by time $2M + K$ with (conditional) probability at least ϵ.

It follows that

$$\begin{aligned} \mathbb{P}(T_{i,j} > 2M + K) &= \mathbb{P}(T_{i,j} > 2M + K \mid T_{i,j} > M)\mathbb{P}(T_{i,j} > M) \\ &\leq (1 - \epsilon)^2. \end{aligned}$$

By iteration of this argument, $\mathbb{P}(T_{i,j} < \infty) = 1$, and (5.6) follows with $r = 1$.

For $r \neq 1$ and $i, j \in S_r$, since Λ is irreducible, a.s. S_r is eventually mapped to S_1. At this point, i and j have both been mapped to elements of S_1, so (if they have not already done so) they will almost surely coalesce by (5.6) with $r = 1$. This verifies (5.6) for general r as required.

Conversely, if there exist $i, j \in S_1$ such that $\mathbb{P}(i \sim j) = 0$ then it is necessarily the case that $k(\mu) \geq \ell + 1$. ∎

[2] One may work with any given value of r in (5.5), and we have chosen $r = 1$ for concreteness. Condition (5.5) is similar to the sufficient condition of [17, Thm 6] for the quenched ergodicity of a Markov chain with random transition matrices. Theorem 5.3 corrects an error in [8, Thm 6].

Example 5.4. Here is an illustration of Theorem 5.3. As there (with $\ell = 2$ for simplicity) we let $P \in \mathcal{P}_S$, $\mathcal{S} = \{S_1, S_2\}$, and we assume that (4.6) and (5.1) hold.

For $r = 1, 2$ we write $S_r = \{x_{r,j} : 1 \le j \le m_r\}$. Ordering the elements of S as $(x_{1,1}\dots, x_{1,m_1}, x_{2,1}, \dots, x_{2,m_2})$, we may express the $(m_1 + m_2) \times (m_1 + m_2)$ matrix P in the form

$$P = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where A is an $m_1 \times m_1$ matrix and D is an $m_2 \times m_2$ matrix.

The 2×2 matrix Λ is doubly stochastic and irreducible, and may (by Theorem 2.6) be expressed in the form

$$\Lambda = \begin{pmatrix} \lambda_{1,1} & \lambda_{1,2} \\ \lambda_{2,1} & \lambda_{2,2} \end{pmatrix} = \alpha I + (1 - \alpha)\overline{I} \tag{5.8}$$

where $\alpha = \lambda_{1,1} = \lambda_{2,2}$ and

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \overline{I} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

In other words, the measure ρ (which is unique since Λ is a 2×2 matrix) puts mass α on the identity permutation, and mass $1 - \alpha$ on the 'interchange' permutation. We let $(\Pi_i : i \in \mathbb{N})$ be independent permutations with common law ρ, and let μ be the $(P, \mathcal{S}, \rho)$-product measure.

We have shown that

$$\Lambda = \begin{pmatrix} \alpha & 1 - \alpha \\ 1 - \alpha & \alpha \end{pmatrix}.$$

Thus, the row sums of A and D (respectively, B and C) are all α (respectively, $1 - \alpha$) by (4.5) and (4.6)). Therefore, P is a mixture of two stochastic matrices P_1 and P_2,

$$P = \alpha P_1 + (1 - \alpha)P_2,$$

where

$$P_1 = \frac{1}{\alpha}\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}, \qquad P_2 = \frac{1}{1 - \alpha}\begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix}.$$

Suppose $\alpha > 0$ (we have $\alpha < 1$ by irreducibility), and that $A' := A/\alpha$ is irreducible and aperiodic. We claim that $k(\mu) = 2$. This may be shown as follows.

Let $W = (W_n : n \ge 0)$ be a Markov chain on S_1 with transition matrix A', and let $i, j \in S_1$. By Theorem 2.5, if W^i and W^j are two

independent versions of W, starting respectively at i and j, then there exists $M < \infty$ such that

$$\mathbb{P}(W_m^i = W_m^j \text{ for some } m \le M) \ge \tfrac{1}{2}.$$

It follows that

$$\begin{aligned}\mathbb{P}(T_{i,j} < \infty) &\ge \mathbb{P}(T_{i,j} < \infty,\ \Pi_m = I \text{ for } m = 1, 2, \dots, M)\\ &\ge \alpha^M \mathbb{P}(W_m^i = W_m^j \text{ for some } m \le M) \ge \tfrac{1}{2}\alpha^m.\end{aligned}$$

Equation (5.5) follows, and hence $k(\mu) = 2$ by Theorem 5.3. ◀

Finally, in preparation for the next section we state a result from [8] which concerns the transition matrix

$$P_n = \begin{pmatrix} n^{-1} & n^{-1} & \cdots & n^{-1} \\ n^{-1} & n^{-1} & \cdots & n^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ n^{-1} & n^{-1} & \cdots & n^{-1} \end{pmatrix}$$

on the state space $S = \{1, 2, \dots, n\}$ with equal entries.

Theorem 5.5 ([8, Thm 7])**.** *For $n \ge 2$ there exists a block measure $\mu \in \mathcal{L}(P_n)$ with $k(\mu) = \ell$ if and only if $\ell \mid n$. In particular, $K(P_n) \supseteq \{\ell : \ell \mid n\}$. For $n \ge 3$, we have $n - 1 \notin K(P_n)$.*

6. Existence of non-block measures

Many of our measures $\mu \in \mathcal{L}_P$ so far have been block measures, though we encountered a certain μ in Example 3.10 which (by Theorem 4.4) is not a block measure. By Theorem 5.5, $n - 1 \notin K(P_n)$, whence $K(P_n) = \{\ell : \ell \mid n\}$ when $n \le 4$. We do not know whether $K(P_n) = \{\ell : \ell \mid n\}$ when $n > 4$. In this section we shall construct a family of non-block measures for P_n.

Theorem 6.1. *Let $n \ge 4$ and suppose $\ell \ne 1$ and $\ell \mid n$. There exists a non-block measure $\mu \in \mathcal{L}_{P_n}$ with $k(\mu) = \ell$.*

Before embarking on the proof of Theorem 6.1, we present an illustrative example.

Example 6.2. Let $n = 6$, $\ell = 2$, and $b = n/\ell = 3$. Let $S_1 = \{1, 2\}$, $S_2 = \{3, 4\}$, and $S_3 = \{5, 6\}$. Let $\widehat{\mathcal{F}} \subset \mathcal{F}_S$ consist of the six functions

$$\begin{aligned} f_{1,1} &= (121212), & f_{1,2} &= (212121),\\ f_{2,1} &= (343443), & f_{2,2} &= (434334),\\ f_{3,1} &= (566565), & f_{3,2} &= (655656). \end{aligned}$$

We make the following observations.

(i) For each $u, v \in \{1, 2, \dots, 6\}$, $f(u) = v$ for exactly one $f \in \widehat{\mathcal{F}}$. Thus, taking μ to be uniform on $\widehat{\mathcal{F}}$, we obtain that $\mu \in \mathcal{L}_{P_6}$.
(ii) Let $f \in \widehat{\mathcal{F}}$. For each i there exists j such that f maps S_i onto S_j. Thus no pair of states in any S_i will ever coalesce.
(iii) The functions $f_{1,j}$ yield the immediately coalescing classes $\{1, 3, 5\}$, $\{2, 4, 6\}$. The functions $f_{2,j}$ yield the immediately coalescing classes $\{1, 3, 6\}$, $\{2, 4, 5\}$. The functions $f_{3,j}$ yield the immediately coalescing classes $\{1, 4, 6\}$, $\{2, 3, 5\}$. All coalescence takes place on the first step, and depending on which function is chosen first (i.e., which function F_1 is) we obtain different coalescing classes. ◀

Proof of Theorem 6.1. Let n, ℓ be as in the statement of the theorem. We will imitate Example 6.2, and the reader may wish to refer back to that example. Let $b = n/\ell$ and

$$S_r = \big\{(r-1)\ell + 1, (r-1)\ell + 2, \dots, r\ell\big\}, \qquad r = 1, 2, \dots, b,$$

noting that $|S_r| = \ell$ for each r. When considering S_r as a vector rather than as a set, we order its elements in increasing order.

We will construct a set $\widehat{\mathcal{F}}$ comprising n $(= b\ell)$ functions $f_{i,j}$ with $i \in \{1, 2, \dots, b\}$, $j \in \{1, 2, \dots, \ell\}$; μ will be the uniform measure on $\widehat{\mathcal{F}}$. Before giving a formal definition of $\widehat{\mathcal{F}}$, we summarise its main properties as follows:

(i) for every $u, v \in \{1, 2, \dots, n\}$ there is exactly one $f \in \widehat{\mathcal{F}}$ such that $f(u) = v$,
(ii) for each $f \in \widehat{\mathcal{F}}$, there is a block S_f such that every block S_r is mapped by f *onto* S_f (viewed as sets),
(iii) the elements of each S_r (viewed as a vector) are mapped to a certain permutation $\pi^r S_f$ of S_f.

We turn now to a formal definition, beginning with some notation. Following Definition 3.9, each $f_{i,j}$ can be expressed in the form $(x_1 x_2 \cdots x_n)$, which is to say that $f_{i,j}(u) = x_u$. It is convenient to break such a vector into consecutive subsequences of length ℓ, and we do this by adding vertical bars; thus, for example, we write $(x_1 x_2 \cdots x_n)$ as

$$\big(x_1 \cdots x_\ell \big| x_{\ell+1} \cdots x_{2\ell} \big| \cdots \big| x_{(b-1)\ell+1} \cdots x_{b\ell}\big),$$

and we term the b subsequences therein 'image blocks'. For each $f_{i,j}$, the image elements $x_1, \dots, x_\ell$ are distinct; furthermore, when viewed as sets, we have that $\{x_{(i-1)\ell+1}, \dots, x_{i\ell}\} = \{x_1, \dots, x_\ell\}$ for each i, so that each subsequent image block of each $f_{i,j}$ is a permutation of the first image block of $f_{i,j}$.

We explain next the action of the $f_{i,j}$. Let ρ denote the rotation permutation $i_1 i_2 \cdots i_\ell \mapsto i_2 i_3 \cdots i_\ell i_1$. We will express each $f_{i,j}$ in terms of a pair $(h_{i,j}, \boldsymbol{\pi}^i)$ where

(a) $h_{i,1} = S_i$ and $h_{i,j} = \rho^{j-1} S_i$ (with S_i considered as a vector),
(b) for each $i = 1, 2, \ldots, b$, there exists a vector $\boldsymbol{\pi}^i = (\pi_2^i, \pi_3^i, \ldots, \pi_b^i)$ of $b-1$ permutations (not necessarily distinct) of ℓ symbols,
(c) we set

$$f_{i,j} = \big(h_{i,j} \big| \pi_2^i h_{i,j} \big| \cdots \big| \pi_b^i h_{i,j}\big). \tag{6.1}$$

We discuss next how the $\boldsymbol{\pi}^i$ are defined, beginning with the simplest case $i = 1$.

Let each π_j^1 be the identity permutation, so that, by (6.1),

$$f_{1,1} = \big(h_{1,1} \big| \pi_2^1 h_{1,1} \big| \cdots \big| \pi_b^1 h_{1,1}\big) = \big(12\cdots\ell \big| 12\cdots\ell \big| \cdots \big| 12\cdots\ell\big).$$

More generally, we let

$$f_{1,j} = \big(h_{1,j} \big| \pi_2^1 h_{1,j} \big| \cdots \big| \pi_b^1 h_{1,j}\big), \tag{6.2}$$

where

$$h_{1,j} = \pi_m^1 h_{1,j} = \big(j(j+1)\cdots(\ell-1)\ell 12\cdots(j-1)\big), \qquad m \geq 2.$$

We discuss next the case of $f_{i,j}$ with $i \geq 2$. Then $h_{i,1} = S_i$ and $h_{i,j} = \rho^{(j-1)} S_i$, and it remains to choose a suitable vector $\boldsymbol{\pi}^i$ of permutations. For $i = 2, 3, \ldots, b$, let $\boldsymbol{\pi}^i = (\pi_2^i, \ldots, \pi_b^i)$ be an ordered set of permutations satisfying $(\pi_2^i, \ldots, \pi_b^i) \neq (\pi_2^{i'}, \ldots, \pi_b^{i'})$ for each $i' < i$ (in other words, all the ordered sets of permutations will be distinct as ordered sets). We can always find distinct $\boldsymbol{\pi}^1, \boldsymbol{\pi}^2, \ldots, \boldsymbol{\pi}^b$ since there are $(\ell!)^{b-1}$ distinct such vectors and $(\ell!)^{b-1} \geq 2^{b-1} \geq b$ (here we have used the fact that $b \geq 2$). The function $f_{i,j}$ is given by (6.1).

By construction, $\widehat{\mathcal{F}}$ has properties (i)–(iii) above. Let μ be the uniform probability measure on $\widehat{\mathcal{F}}$. By (i),

$$\mu(\{f : f(u) = v\}) = \frac{1}{n}, \qquad u, v \in S,$$

whence $\mu \in \mathcal{L}(P_n)$.

We turn to the issue of coalescence. By (ii)–(iii) above, coalescence occurs at the first stage and not subsequently. Since there are at least two distinct $\boldsymbol{\pi}^i$, the coalescing classes are random (indeed the number of possibilities for the set of coalescing classes is the number of distinct $\boldsymbol{\pi}^i$, $i = 1, 2, \ldots, b$). Since mappings between blocks are surjections, there are exactly ℓ coalescing classes $C_1, C_2, \ldots, C_\ell$, and each such C_r is a transversal of $\mathcal{S}$. Hence $k(\mu) = \ell$. ∎

Remark 6.3. In the above proof, we have required that the $\boldsymbol{\pi}^i$ be distinct. It suffices for the proof that at least two of them are distinct. Suppose, on the contrary, that $\boldsymbol{\pi}^j = \boldsymbol{\pi}^1$ for all j. Then the set of coalescing classes is deterministic, and Theorem 4.4 applies. In particular since $\boldsymbol{\pi}^1$ is the identity permutation we may see that

$$C_j = \{j, \ell + j, \dots, (b-1)\ell + j\}, \quad j = 1, 2, \dots, \ell.$$

We deduce that $k(\mu) = \ell$, and that μ is a block measure with blocks $C_1, C_2, \dots, C_\ell$. ◀

7. Functions that generate transition matrices

In this final section we pose an inverse question. As usual, the state space is $S = \{1, 2, \dots, n\}$. Recall the set $\mathcal{P}_S$ of transition matrices on S and the set $\mathcal{F}_S$ of functions from S to S. Given a subset $\mathcal{G} \subseteq \mathcal{F}_S$, let

$$\mathcal{P}(\mathcal{G}) = \{P \in \mathcal{P}_S : \exists \mu \in \mathcal{L}_P \text{ with } \mathrm{supp}(\mu) \subseteq \mathcal{G}\}.$$

In other words, $\mathcal{P}(\mathcal{G})$ is the set of (irreducible) stochastic matrices that can be obtained by just using functions in $\mathcal{G}$.

Question 7.1. *For given $\mathcal{G} \subseteq \mathcal{F}_S$, what can be said about the set $\mathcal{P}(\mathcal{G})$?*

Example 7.2. For distinct $x, y \in S$ let

$$f_{x,y}(v) = \begin{cases} y & \text{if } v = x, \\ v & \text{otherwise.} \end{cases}$$

This is 'almost' the identity function (except that $x \mapsto y$). Let $\mathcal{G} = \{f_{x,y} : x, y \in S,\ x \neq y\}$, so that $\mathcal{G}$ contains $n(n-1)$ functions. Let $\boldsymbol{\alpha} = (\alpha_{x,y} : x \neq y)$ be a vector of non-negative reals satisfying

$$\sum_{x,y:\, x \neq y} \alpha_{x,y} = 1.$$

Let $\mu_{\boldsymbol{\alpha}}$ be the probability measure on $\mathcal{G}$ that selects $f_{x,y}$ with probability $\alpha_{x,y}$. The corresponding $P = P_{\boldsymbol{\alpha}} = (p_{i,j})$ satisfies

$$p_{i,j} = \mu\big(\{f : f(i) = j\}\big) = \mu(\{f_{i,j}\}) = \alpha_{i,j}, \qquad i \neq j,$$

while

$$p_{i,i} = \mu\big(\{f : f(i) = i\}\big) = 1 - \sum_{j:\, j \neq i} \alpha_{i,j}.$$

Therefore, $\mathcal{P}(\mathcal{G})$ is the set of stochastic matrices whose off-diagonal elements sum to 1. ◀

Some very special cases admit precise answers to Question 7.1. Let $\mathcal{D}_S \subset \mathcal{P}_S$ denote the set of doubly stochastic matrices, and let $\mathcal{F}_S^{\mathrm{perm}} \subset \mathcal{F}_S$ denote the set of permutations of S.

Proposition 7.3.

(a) $\mathcal{P}(\mathcal{G}) = \mathcal{D}_S$ *if and only if* $\mathcal{G} = \mathcal{F}_S^{\mathrm{perm}}$.
(b) $\mathcal{P}(\mathcal{G}) = \mathcal{P}_S$ *if and only if* $\mathcal{G} = \mathcal{F}_S$.

Proof. (a) This is a consequence of Theorem 2.6.

(b) Suppose first that $\mathcal{G} = \mathcal{F}_S$. For $P \in \mathcal{P}_S$ one can construct its independence coupling (recall Example 2.2). Therefore $\mathcal{P}(\mathcal{G}) = \mathcal{P}_S$.

Suppose conversely that $\mathcal{G}$ is such that $\mathcal{P}(\mathcal{G}) = \mathcal{P}_S$. We claim that for every $f \in \mathcal{F}_S$ there exists $P \in \mathcal{P}_S$ (hence $P \in \mathcal{P}(\mathcal{G})$) such that $f \in \mathrm{supp}(\mu)$ for every $\mu \in \mathcal{L}_P$. This implies that $f \in \mathcal{G}$ and therefore $\mathcal{G} = \mathcal{F}_S$ as required.

It remains to verify the above claim. Let $f = (j_1 j_2 \dots j_n) \in \mathcal{F}_S$ (recall Definition 3.9). There exists $P_f = (p_{i,j}) \in \mathcal{P}_S$ such that $p_{i,j_i} > 1 - 1/n$ for every $i \in S$; the remaining terms $p_{i,j}$, $j \neq j_i$, are assumed strictly positive, thus implying that P_f is irreducible. For $\mu \in \mathcal{L}_{P_f}$, we have

$$\mu\big(\{g \in \mathcal{F}_S : g(i) = j_i\}\big) > 1 - \frac{1}{n}, \qquad i = 1, 2, \dots, n,$$

whence

$$\mu(f) = \mu\left(\bigcap_{i \in S} \{g : g(i) = j_i\}\right) > 0.$$

Therefore, $f \in \mathrm{supp}(\mu)$. ∎

We recall the random transition matrix of Remark 2.4, with law $\mathbb{Q}$.

Question 7.4. *For given $\mathcal{G} \subset \mathcal{F}_S$, what can be said about $\mathbb{Q}(\mathcal{P}(\mathcal{G}))$?*

This section closes with some partial answers to this question.

Firstly, in the proof of Proposition 7.3(b) it is shown that, if $f \notin \mathcal{G}$, then any matrix with $p_{i,f(i)} > 1 - n^{-1}$ for all $i \in S$ does not belong to $\mathcal{P}(\mathcal{G})$. It follows that $\mathbb{Q}(\mathcal{P}(\mathcal{G})) < 1$ whenever $\mathcal{G} \neq \mathcal{F}_S$.

Secondly, in Example 7.2, we have $\mathbb{Q}(\mathcal{P}(\mathcal{G})) = 0$ since $\mathcal{G}_S$ is the set of stochastic matrices with off-diagonal elements summing to 1.

Thirdly, in the case $|S| = 2$ there are just 4 functions, namely

$$(11), (12), (21), (22).$$

It is elementary that if $\mathcal{G}$ is missing two of them then $\mathbb{Q}(\mathcal{G}) = 0$. The following lemma shows that (when $n = 2$) removing a single function one retains positive measure.

Proposition 7.5. *If $|S| = 2$ and $f \in \mathcal{F}_S$, then $\mathbb{Q}(\mathcal{P}(\mathcal{F}_S \setminus \{f\})) = \frac{1}{2}$.*

Proof. Given a probability mass function $\boldsymbol{\alpha}$ on $\mathcal{F}_S$, the resulting matrix P is given by

$$P = \begin{pmatrix} \alpha_{(11)} + \alpha_{(12)} & \alpha_{(21)} + \alpha_{(22)} \\ \alpha_{(11)} + \alpha_{(21)} & \alpha_{(12)} + \alpha_{(22)} \end{pmatrix}. \tag{7.1}$$

Henceforth, fix $f = (uv)$ and write $u' \neq u$ and $v' \neq v$. Let $\mathcal{P}_{uv}$ denote the set of $P \in \mathcal{P}_S$ for which $p_{2,v} < p_{1,u'}$. Since $p_{2,v}$ and $p_{1,u'}$ are independent with the same probability density function (under $\mathbb{Q}$), we have that $\mathbb{Q}(\mathcal{P}_{uv}) = \frac{1}{2}$. We claim that

$$\mathcal{P}(\mathcal{F}_S \setminus \{f\}) = \mathcal{P}_{uv}, \tag{7.2}$$

and the proposition will follow immediately. It remains to prove (7.2).

Let $P \in \mathcal{P}(\mathcal{F}_S \setminus \{f\})$. There exists a mass function $\boldsymbol{\alpha}$ on $\mathcal{F}_S$ satisfying $\alpha_{(uv)} = 0$ such that (7.1) holds; note that $\alpha_{(ab)} > 0$ for $(ab) \neq (uv)$, $\mathbb{Q}$-a.s., and we will assume this. By (7.1),

$$p_{2,v} = \alpha_{(u'v)} < \alpha_{(u'v)} + \alpha_{(u'v')} = p_{1,u'}.$$

Since P is stochastic, we have also that $p_{2,v'} > p_{1,u}$, and hence $P \in \mathcal{P}_{uv}$ as required.

Conversely, let $P \in \mathcal{P}_{uv}$, so that $p_{2,v} < p_{1,u'}$. Set

$$\alpha_{(uv)} = 0, \quad \alpha_{(uv')} = p_{1,u}, \quad \alpha_{(u'v)} = p_{2,v}, \quad \alpha_{(u'v')} = 1 - (p_{1,u} + p_{2,v}),$$

and note that

$$1 - \alpha_{(u'v')} = p_{1,u} + p_{2,v} < p_{1,u} + p_{1,u'} = 1.$$

The stochastic matrix $Q = (q_{i,j})$ corresponding to $\boldsymbol{\alpha}$ (according to (7.1)) satisfies

$$\begin{aligned} q_{1,u} &= \alpha_{(uv)} + \alpha_{(uv')} = \alpha_{(uv')} = p_{1,u}, \\ q_{2,v} &= \alpha_{(uv)} + \alpha_{(u'v)} = \alpha_{(u'v)} = p_{2,v}, \end{aligned}$$

whence $Q = P$. Therefore, $\mathcal{P}_{uv} \subseteq \mathcal{P}(\mathcal{F}_S \setminus \{f\})$ and the proof of (7.2) is complete. ∎

Finally, for any $\epsilon > 0$, when $|S|$ is large there exist sets $\mathcal{G} \subseteq \mathcal{F}_S$ with size satisfying $|\mathcal{G}| > (1 - \epsilon)|\mathcal{F}_S|$ for which $\mathcal{P}(\mathcal{G})$ has $\mathbb{Q}$-measure zero. Here are two examples.

Proposition 7.6. *We have $\mathbb{Q}(\mathcal{P}(\mathcal{G})) = 0$ in the following two cases.*

(a) *Let $i, j \in S$ and $\mathcal{G} = \mathcal{G}_{i,j} = \{f \in \mathcal{F}_S : f(i) \neq j\}$.*

(b) *Let $i_1, i_2, j_1, j_2 \in S$ satisfy $i_1 \neq i_2$, and let $\mathcal{G}$ be the set of all $f \in \mathcal{F}_S$ such that*

$$f(i_1) = j_1 \quad \text{if and only if} \quad f(i_2) = j_2.$$

Let $n = |S|$. The cardinality of $\mathcal{G}$ in part (a) is $|\mathcal{F}_S|(1 - n^{-1})$, and in part (b) is $|\mathcal{F}_S|(1 - 2n^{-1} + \mathrm{o}(n^{-1}))$.

Proof. (a) For any μ supported on $\mathcal{G}_{i,j}$ we have $\mu(\{f : f(i) = j\}) = 0$, whence every $P = (p_{i,j}) \in \mathcal{P}(\mathcal{G})$ has $p_{i,j} = 0$. The probability of such P satisfies $\mathbb{Q}(\mathcal{P}(\mathcal{G})) = 0$.

(b) For any μ supported on such $\mathcal{G}$ we have

$$\mu(\{f : f(i_1) = j_1\}) = \mu(\{f : f(i_2) = j_2\}),$$

so every $P = (p_{i,j}) \in \mathcal{P}(\mathcal{G})$ has $p_{i_1,j_1} = p_{i_2,j_2}$. The claim follows. ■

Example 7.7. Let $n = 3$, so that $|\mathcal{F}_S| = 27$. Let $\mathcal{G}$ be the following set of 15 functions:

$$\begin{aligned}\mathcal{G} = \{&(111), (311), (121), (321), \mathbf{(231)},\\ &(112), (312), (122), (322), \mathbf{(232)},\\ &(113), (313), (123), (323), \mathbf{(233)}\}.\end{aligned}$$

Note that $1 \mapsto 2$ if and only if $2 \mapsto 3$ (the relevant functions are in bold above) so, by Proposition 7.6(b), $\mathcal{P}(\mathcal{G})$ has zero $\mathbb{Q}$-measure. ◀

Acknowledgements

The authors are grateful to Tabet Aoul Alaa Eddine Islem for indicating an error in an earlier version of Theorem 4.3 as published in [8].

References

1. O. Angel, A. E. Holroyd, J. Martin, D. B. Wilson, and P. Winkler, *Avoidance coupling*, Electron. Commun. Probab. **18** (2013), no. 58, 13.
2. E. Bates and M. Podder, *Avoidance couplings on non-complete graphs*, Random Struct. Alg. **59** (2021), 25–52.
3. G. Birkhoff, *Three observations on linear algebra*, Univ. Nac. Tucumán. Revista A. **5** (1946), 147–151.
4. P. Brémaud, *Markov Chains*, 2nd ed., Springer Nature, Cham, 2020.
5. W. Doeblin, *Exposé de la théorie des chaînes simples constantes de Markoff à un nombre fini d'états*, Revue Math. de l'Union Interbalkanique **2** (1938), 77–105.
6. B. C. Geiger and C. Temmel, *Lumpings of Markov chains, entropy rate preservation, and higher-order lumpability*, J. Appl. Probab. **51** (2014), 1114–1132.
7. K. Georgiou, G. N. Domazakis, D. Pappas, and A. N. Yannacopoulos, *Markov chain lumpability and applications to credit risk modelling in compliance with the International Financial Reporting Standard 9 framework*, European J. Oper. Res. **292** (2021), 1146–1164.
8. G. R. Grimmett and M. Holmes, *Non-coupling from the past*, In and Out of Equilibrium 3. Celebrating Vladas Sidoravicius, Progr. Probab., vol. 77, Birkhäuser/Springer, Cham, 2021, pp. 471–485.

9. G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, 4th ed., Oxford University Press, Oxford, 2020.
10. G. R. Grimmett and D. Welsh, *Probability—an Introduction*, 2nd ed., Oxford University Press, Oxford, 2014.
11. J. G. Kemeny and J. L. Snell, *Finite Markov Chains*, Van Nostrand, 1963.
12. A. Marin, C. Piazza, and S. Rossi, *Proportional lumpability and proportional bisimilarity*, Acta Inform. **59** (2022), 211–244.
13. J. von Neumann, *A certain zero-sum two-person game equivalent to the optimal assignment problem*, Contributions to the theory of games, vol. 2, Annals of Mathematics Studies, no. 28, Princeton University Press, Princeton, NJ, 1953, pp. 5–12.
14. J. R. Norris, *Markov Chains*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 2, Cambridge University Press, Cambridge, 1998.
15. J. G. Propp and D. B. Wilson, *Exact sampling with coupled Markov chains and applications to statistical mechanics*, Random Structures Algorithms **9** (1996), 223–252, Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995).
16. ———, *How to get a perfectly random sample from a generic Markov chain and generate a random spanning tree of a directed graph*, J. Algorithms **27** (1998), 170–217, 7th Annual ACM-SIAM Symposium on Discrete Algorithms (Atlanta, GA, 1996).
17. Y. Takahashi, *Markov chains with random transition matrices*, Kōdai Mathematical Seminar Reports **21** (1969), 426–447.
18. D. B. Wilson and J. G. Propp, *How to get an exact sample from a generic Markov chain and sample a random spanning tree from a directed graph, both within the cover time*, Proceedings of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms (Atlanta, GA, 1996), ACM, New York, 1996, pp. 448–457.

(GRG) Statistical Laboratory, Centre for Mathematical Sciences, Cambridge University, Wilberforce Road, Cambridge CB3 0WB, UK
Email address: grg@statslab.cam.ac.uk
URL: http://www.statslab.cam.ac.uk/~grg/

(MH) School of Mathematics & Statistics, The University of Melbourne, Parkville, VIC 3010, Australia
Email address: holmes.m@unimelb.edu.au
URL: https://researchers.ms.unimelb.edu.au/~mholmes1@unimelb/