The Annals of Probability
2014, Vol. 42, No. 5, 1952–1979
DOI: 10.1214/14-AOP922

THE CUT-AND-PASTE PROCESS[1]

BY HARRY CRANE

Rutgers University

We characterize the class of exchangeable Feller processes evolving on partitions with boundedly many blocks. In continuous-time, the jump measure decomposes into two parts: a σ-finite measure on stochastic matrices and a collection of nonnegative real constants. This decomposition prompts a Lévy–Itô representation. In discrete-time, the evolution is described more simply by a product of independent, identically distributed random matrices.

1. Introduction. For fixed $k = 1, 2, \ldots$, a *k-coloring* of $\mathbb{N} := \{1, 2, \ldots\}$ is an infinite sequence $x = x^1 x^2 \cdots$ taking values in $[k] := \{1, \ldots, k\}$. Two operations bear on our main theorems:

- relabeling: for any permutation $\sigma : \mathbb{N} \to \mathbb{N}$, the *relabeling* of $x = x^1 x^2 \cdots$ by σ is

$$x^\sigma := x^{\sigma(1)} x^{\sigma(2)} \cdots \quad \text{and} \tag{1.1}$$

- restriction: for any finite $n = 1, 2, \ldots$, the *restriction* of x to a k-coloring of $[n]$ is

$$x^{[n]} := x^1 \cdots x^n. \tag{1.2}$$

A Markov process $\mathbf{X} = (X_t, t \geq 0)$ on $[k]^{\mathbb{N}}$, the space of infinite k-colorings, is

(A) *exchangeable* if $\mathbf{X}^\sigma = (X_t^\sigma, t \geq 0)$ is a version of $\mathbf{X}$ for all finite permutations $\sigma : \mathbb{N} \to \mathbb{N}$ and
(B) *consistent (under subsampling)* if $\mathbf{X}^{[n]} = (X_t^{[n]}, t \geq 0)$ is a Markov chain on k-colorings of $[n]$, for all finite $n = 1, 2, \ldots$.

We characterize both $[k]^{\mathbb{N}}$-valued Markov processes satisfying (A) and (B) and a class of partition-valued processes with analogous properties. When $[k]^{\mathbb{N}}$ is endowed with the product-discrete topology, exchangeability and consistency are equivalent to exchangeability and the Feller property; and so our main theorems characterize exchangeable Feller processes on $[k]^{\mathbb{N}}$ and $\mathcal{P}_{\mathbb{N}:k}$, partitions of $\mathbb{N}$ with at most k blocks.

Received July 2012; revised January 2014.
[1]Supported in part by NSF Grant DMS-13-08899 and NSA Grant H98230-13-1-0299.
MSC2010 subject classifications. Primary 60J25; secondary 60G09, 60J35.
Key words and phrases. Exchangeable random partition, de Finetti's theorem, Lévy–Itô decomposition, paintbox process, coalescent process, interacting particle system, Feller process, random matrix product.

1.1. *Discrete-time characterization.* A stochastic matrix $S = (S_{ii'}, 1 \le i, i' \le k)$ has nonnegative entries and all rows summing to one, and it determines the transition probabilities of a time-homogeneous Markov chain $\mathbf{Y} = (Y_m, m \ge 0)$ on $[k]$ by

$$\mathbb{P}_S\{Y_1 = i' \mid Y_0 = i\} = S_{ii'}, \qquad i, i' = 1, \ldots, k. \tag{1.3}$$

From any probability measure Σ on the space of $k \times k$ stochastic matrices, we construct a Markov chain $\mathbf{X}^*_\Sigma := (X^*_m, m \ge 0)$ on $[k]^{\mathbb{N}}$ as follows. First, we let X^*_0 be an exchangeable initial state and $S_1, S_2, \ldots$ be independent, identically distributed (i.i.d.) random matrices from Σ. Then, for $m = 1, 2, \ldots,$ we generate the components of $X^*_m = X^{*1}_m X^{*2}_m \cdots$, given $X^*_{m-1}, \ldots, X^*_0, S_1, S_2, \ldots,$ conditionally independently from transition probability matrix $S_m := (S_m(i, i'), 1 \le i, i' \le k)$,

$$\mathbb{P}\{X^{*j}_m = i' \mid X^*_{m-1}, S_m\} = S_m(i, i') \qquad \text{on the event } X^{*j}_{m-1} = i.$$

Such a construction exists for all exchangeable and consistent Markov chains on $[k]^{\mathbb{N}}$.

THEOREM 1.1. *Let* $\mathbf{X} = (X_m, m \ge 0)$ *be a discrete-time, exchangeable, consistent Markov chain on* $[k]^{\mathbb{N}}$. *Then there exists a unique probability measure* Σ *such that* $\mathbf{X}^*_\Sigma$ *is a version of* $\mathbf{X}$.

To any $x \in [k]^{\mathbb{N}}$, the *asymptotic frequency vector* $|x| := (f_1(x), \ldots, f_k(x))$ is an element of the $(k-1)$-dimensional simplex Δ_k, where

$$f_i(x) := \lim_{n\to\infty} n^{-1} \sum_{j=1}^{n} \mathbf{1}\{x^j = i\}, \qquad i = 1, \ldots, k, \tag{1.4}$$

is the limiting proportion of coordinates labeled i in x, if it exists. With probability one, the asymptotic frequency vector of any exchangeable k-coloring exists and $|\mathbf{X}| := (|X_m|, m \ge 0)$ is a sequence in Δ_k. From the same i.i.d. sequence $S_1, S_2, \ldots$ used to construct $\mathbf{X}^*_\Sigma = (X^*_m, m \ge 0)$ in Theorem 1.1, we can construct $\mathbf{\Phi}_\Sigma := (\Phi_m, m \ge 0)$ in Δ_k by putting $\Phi_0 := |X^*_0|$ and

$$\Phi_m := \Phi_{m-1} S_m = \Phi_0 S_1 \cdots S_m, \qquad m \ge 1, \tag{1.5}$$

where $\Phi_{m-1} S_m$ in (1.5) is the usual right action of a $k \times k$ matrix on a $1 \times k$ row vector.

THEOREM 1.2. *Let* $\mathbf{X} = (X_m, m \ge 0)$ *be a discrete-time, exchangeable, consistent Markov chain on* $[k]^{\mathbb{N}}$. *Then* $\mathbf{\Phi}_\Sigma$ *is a version of* $|\mathbf{X}|$, *where* Σ *is the unique probability measure from Theorem* 1.1.

Together, Theorems 1.1 and 1.2 relate the evolution of discrete-time Markov chains to products of i.i.d. random matrices. Crane and Lalley [7] have combined representation (1.5) with the Furstenberg–Kesten theorem [10] to identify a class of these chains that exhibits the cutoff phenomenon.

1.2. *Continuous-time characterization.* In continuous-time, an exchangeable, consistent Markov process $\mathbf{X} = (X_t, t \geq 0)$ can jump infinitely often, and thus, behaves differently than its discrete-time counterpart; but consistency limits this behavior: since each restriction $\mathbf{X}^{[n]}$ is a finite state space Markov process, it must remain in each visited state for a positive amount of time. The upshot of these observations is a characterization of the transition law of $\mathbf{X}$ by a unique σ-finite measure on $k \times k$ stochastic matrices and a unique collection of nonnegative constants.

Our next theorem yields a Lévy–Itô-type characterization of $\mathbf{X}$ by dividing its discontinuities into two cases. Let $t > 0$ be the time of a discontinuity in $\mathbf{X}$. Then either

(I) a positive proportion of coordinates changes colors at time t, that is,

$$\lim_{n\to\infty} n^{-1} \sum_{j=1}^{n} \mathbf{1}\{X_{t-}^{j} \neq X_t^j\} > 0 \quad \text{or}$$

(II) a zero proportion of coordinates changes colors at time t, that is,

$$\lim_{n\to\infty} n^{-1} \sum_{j=1}^{n} \mathbf{1}\{X_{t-}^{j} \neq X_t^j\} = 0.$$

In discrete-time, Type-(I) jumps are governed by a probability measure Σ and Type-(II) transitions are forbidden. In continuous-time, Type-(I) jumps are governed by a σ-finite measure Σ and Type-(II) transitions include only *single-index flips*, that is, jumps for which exactly one coordinate changes color. Deciding the Type-(II) jump rates is a collection of nonnegative constants $\mathbf{c} = (\mathbf{c}_{ii'}, 1 \leq i \neq i' \leq k)$: independently, each coordinate changes colors from i to i' at rate $\mathbf{c}_{ii'}$. The transition law of $\mathbf{X}$ is characterized by the pair $(\Sigma, \mathbf{c})$.

We do not fully explain $(\Sigma, \mathbf{c})$ and its relation to $\mathbf{X}$ until Section 4. Sparing the details, we write $\mathbf{X}^*_{\Sigma,\mathbf{c}}$ to denote a continuous-time Markov process constructed from a Poisson point process with intensity measure determined by $(\Sigma, \mathbf{c})$. Theorem 1.3 says that any exchangeable, consistent Markov process $\mathbf{X}$ admits a version with this construction.

THEOREM 1.3. *Let $\mathbf{X} = (X_t, t \geq 0)$ be a continuous-time, exchangeable, consistent Markov process on $[k]^{\mathbb{N}}$. Then there exists a unique measure Σ satisfying (1.6) and unique nonnegative constants $\mathbf{c} = (\mathbf{c}_{ii'}, 1 \leq i \neq i' \leq k)$ such that $\mathbf{X}^*_{\Sigma,\mathbf{c}}$ is a version of $\mathbf{X}$.*

In Theorem 1.3, Σ is required to satisfy

$$\Sigma(\{I_k\}) = 0 \quad \text{and} \quad \int_{\mathcal{S}_k} (1 - S_*)\Sigma(dS) < \infty, \tag{1.6}$$

where I_k is the $k \times k$ identity matrix, $S_* := \min(S_{11}, \dots, S_{kk})$ for any $k \times k$ stochastic matrix S, and $\mathcal{S}_k$ is the space of $k \times k$ stochastic matrices. Consistency imposes (1.6): uniqueness requires the first half, finiteness of finite-dimensional jump rates forces the second half.

As in discrete-time, we define the projection of $\mathbf{X} = (X_t, t \ge 0)$ into Δ_k by $|\mathbf{X}| = (|X_t|, t \ge 0)$. Unlike discrete-time, the existence of $|\mathbf{X}|$ does not follow directly from de Finetti's theorem because now $\mathbf{X}$ is an uncountable collection.

THEOREM 1.4. *Let $\mathbf{X} = (X_t, t \ge 0)$ be a continuous-time, exchangeable, consistent Markov process on $[k]^{\mathbb{N}}$. Then $|\mathbf{X}| = (|X_t|, t \ge 0)$ exists almost surely and is a Feller process on Δ_k.*

Theorems 1.3 and 1.4 give the Lévy–Itô representation. The projection $|\mathbf{X}|$ jumps only at the times of Type-(I) discontinuities in $\mathbf{X}$; at other times, it follows a continuous, deterministic trajectory. Thus, Theorem 1.3 warrants the heuristic interpretation that Σ governs the "discrete" component of $\mathbf{X}$ and $\mathbf{c}$ governs the "continuous" component.

1.3. *Partition-valued Markov processes.* Any $x \in [k]^{\mathbb{N}}$ determines a partition $\pi = \mathcal{B}(x)$ of $\mathbb{N}$ through

$$i \text{ and } j \text{ are in the same block of } \pi \iff x^i = x^j. \tag{1.7}$$

If the characteristic pair $(\Sigma, \mathbf{c})$ treats colors symmetrically, that is, Σ is row–column exchangeable and $\mathbf{c}_{ii'} = \mathbf{c}_{jj'} = c$ for all $i \ne i'$ and $j \ne j'$, then the projection $\mathcal{B}(\mathbf{X}^*_{\Sigma,\mathbf{c}}) = (\mathcal{B}(X^*_t), t \ge 0)$ into $\mathcal{P}_{\mathbb{N}:k}$ through (1.7) is an exchangeable, consistent Markov process on $\mathcal{P}_{\mathbb{N}:k}$. Our main theorem for partition-valued processes states that any exchangeable, consistent Markov process on $\mathcal{P}_{\mathbb{N}:k}$ can be generated by projecting an exchangeable, consistent Markov process from $[k]^{\mathbb{N}}$.

THEOREM 1.5. *Let $\mathbf{\Pi}$ be an exchangeable, consistent Markov process on $\mathcal{P}_{\mathbb{N}:k}$.*

- *In discrete-time, there exists a unique, row–column exchangeable probability measure Σ such that $\mathcal{B}(\mathbf{X}^*_\Sigma)$ is a version of $\mathbf{\Pi}$;*
- *in continuous-time, there exists a unique, row–column exchangeable measure satisfying (1.6) and a unique constant $c \ge 0$ such that $\mathcal{B}(\mathbf{X}^*_{\Sigma,\mathbf{c}})$ is a version of $\mathbf{\Pi}$, where $\mathbf{c}_{ii'} = c$ for all $1 \le i \ne i' \le k$.*

Analogously to (1.4), we define the asymptotic frequency of $\pi \in \mathcal{P}_{\mathbb{N}}$ by $|\pi|^{\downarrow}$, the asymptotic block frequencies of π in decreasing order of size. When it exists, $|\pi|^{\downarrow}$ is an element of the *ranked k-simplex* $\Delta^{\downarrow}_k$.

THEOREM 1.6. *Let $\mathbf{\Pi} = (\Pi_t, t \ge 0)$ be a continuous-time, exchangeable, consistent Markov process on $\mathcal{P}_{\mathbb{N}:k}$. Then $|\mathbf{\Pi}|^{\downarrow} := (|\Pi_t|^{\downarrow}, t \ge 0)$ exists almost surely and is a Feller process on $\Delta^{\downarrow}_k$.*

1.4. *The cut-and-paste process.* We call $\mathbf{X}^*_{\Sigma,\mathbf{c}}$ a *cut-and-paste process*: its jumps occur by first *cutting* each color class into subclasses and then *pasting* subclasses together. When $(\Sigma, \mathbf{c})$ treats colors symmetrically, we call $\mathbf{X}^*_{\Sigma,\mathbf{c}}$ and its projection into $\mathcal{P}_{\mathbb{N}:k}$ a *homogeneous cut-and-paste process*.

Cut-and-paste processes should not be conflated with synonymous, but not analogous, *split-and-merge* [15] and *coagulation–fragmentation* processes [8]. The latter processes share aspects, but are not one, with the cut-and-paste process. Each process evolves by operations that divide (cut, split, fragment) and unite (paste, merge, coagulate), but split-and-merge processes evolve on interval partitions, coagulation–fragmentation processes on set partitions, and cut-and-paste processes on k-colorings. At the time of a jump, a cut-and-paste process undergoes two operations simultaneously (cut *and* paste), the others undergo only one operation (split *or* merge, coagulate *or* fragment).

Theorems 1.5 and 1.6 do elicit qualitative connections to exchangeable coalescent and fragmentation processes [2, 14], both of which are characterized by pairs (ν, c), where ν is a unique σ-finite measure on ranked-mass partitions and $c \geq 0$ is a unique constant. For coalescent processes, ν determines the rate of multiple collisions and c the rate of binary coalescence. For fragmentation processes, ν determines the rate of dislocation and c the rate of erosion. In both cases, (ν, c) gives a Lévy–Itô description. But, in a strict sense, processes on $\mathcal{P}_{\mathbb{N}:k}$ behave differently than those on $\mathcal{P}_{\mathbb{N}}$ [5, 6], and Theorem 1.5 neither refines nor is a special case of previous results. In Section 6.1, we further discuss any relationships (and lack thereof) between cut-and-paste, coalescent and fragmentation processes.

1.5. *Applications to DNA sequencing.* Decades ago, population genetics applications motivated the initial study of random partitions and partition-valued processes [9, 11, 13]. Somewhat later, Bertoin [2, 3] and Pitman [14, 16] connected coalescent and fragmentation processes to Brownian motion, Lévy processes and subordinators. In the present, DNA sequencing inspires processes restricted to partitions with a bounded number of blocks.

For let the colors correspond to DNA nucleotides, adenine (A), cytosine (C), guanine (G) and thymine (T). Then, for a sample of n individuals, $X^1 \cdots X^n \in \{A, C, G, T\}^{[n]}$ is a string of DNA nucleotides at a particular chromosomal site, where X^i denotes the nucleotide of individual $i = 1, \ldots, n$. If we observe a DNA sequence $(X^i_m, m \geq 0)$ for each $i = 1, \ldots, n$, then $(X_m, m \geq 0)$ is a sequence in $\{A, C, G, T\}^{[n]}$, with $X_m = X^1_m \cdots X^n_m$. By forgetting colors (in this case nucleotides), we obtain a sequence of set partitions; see Table 1.

In practice, biological phenomena such as recombination induce dependence among nearby chromosomal sites. For modeling this dependence, the Markov property strikes a balance between practical feasibility and mathematical tractability. Exchangeability and consistency incorporate a logical structure that is apt for DNA sequencing. See [4] for a detailed statistical consideration of these applications.

TABLE 1

An array of DNA sequences for 3 *individuals. From this array, we obtain a sequence in* $\{A, C, G, T\}^{[3]}$: (*AAT, ATT, TTT, CCG, CGG, GGC, AAT*, ...). *By ignoring nucleotide labels, we obtain the sequence* (12|3, 1|23, 123, 12|3, 1|23, 12|3, 12|3, ...) *of partitions of the set* $\{1, 2, 3\}$

Individuals/sites	$m = 1$	**2**	**3**	**4**	**5**	**6**	**7**	**...**
X_m^1	A	A	T	C	C	G	A	...
X_m^2	A	T	T	C	G	G	A	...
X_m^3	T	T	T	G	G	C	T	...

1.6. *Discussion of main theorems.* For concreteness, let $\mathbf{X}$ be a discrete-time Markov chain on $\{1, 2\}^{\mathbb{N}}$. According to Theorem 1.1, a transition $X \mapsto X'$ can be generated in two steps:

(i) Draw a random pair (p_1, p_2) of success probabilities from a probability measure Σ on $[0, 1] \times [0, 1]$.
(ii) Given (p_1, p_2), update each coordinate $j = 1, 2, \ldots$ of X independently by the following coin flipping process.

- If $X^j = 1$, flip a p_1-coin ($\mathbb{P}\{\text{heads}\} = p_1$); otherwise, flip a p_2-coin.
- If the outcome is heads, put $X'^j = 1$; otherwise, put $X'^j = 2$.

The pair (p_1, p_2) determines a 2×2 stochastic matrix

$$S = \begin{pmatrix} p_1 & 1 - p_1 \\ p_2 & 1 - p_2 \end{pmatrix},$$

which describes the transition probability matrix for each coordinate, as in (1.3). By the law of large numbers, the proportion of coordinates labeled 1 in X' equals

$$\begin{aligned} f_1(X') &= \mathbb{P}\{X'^1 = 1 \mid X^1 = 1\} f_1(X) + \mathbb{P}\{X'^1 = 1 \mid X^1 = 2\} f_2(X) \\ &= p_1 f_1(X) + p_2 f_2(X). \end{aligned}$$

Overall, the asymptotic frequencies $|X'| = (f_1(X'), f_2(X'))$ of X' are the entries of

$$|X| S = \begin{pmatrix} f_1(X) & f_2(X) \end{pmatrix} \begin{pmatrix} p_1 & 1 - p_1 \\ p_2 & 1 - p_2 \end{pmatrix}.$$

In discrete-time, exchangeability implies that if $X' \neq X$, then the proportion of coordinates changing colors from X to X' is strictly positive. In continuous-time, the transition rate $X \mapsto X'$ need not be bounded, and thus, Σ need not be finite. Furthermore, there is no requirement that a strictly positive proportion of coordinates changes colors at the time of a discontinuity. However, the consistency assumption implies that any finite collection of coordinates jumps at a finite rate, producing condition (1.6). Together, exchangeability and consistency restrict

Type-(II) discontinuities to involve only a single coordinate, called a *single-index flip*. For instance, if "double-index flips" were permitted, that is, a pair of indices changes colors simultaneously while all other coordinates remain unchanged, then the finite restrictions of $\mathbf{X}$ could not be càdlàg. To see this, suppose any pair $(X^n, X^{n'})$, $n < n'$, changes from $(1,1)$ to $(2,2)$ at positive rate $\mathbf{r}$. Then, by exchangeability, any pair $(X^n, X^{n'+j})$, $j \geq 1$, in state $(1,1)$ must also flip at rate $\mathbf{r}$. For any such jump, the restriction of $\mathbf{X}$ to $[n]$ witnesses only a change in coordinate n at rate $\sum_{n'>n} \mathbf{r} = \infty$, which contradicts assumption (B). For similar reasons, condition (1.6) prevents infinitely many Type-(I) discontinuities from bunching up in any finite restriction of $\mathbf{X}$.

Upon observing our main theorems for $[k]^{\mathbb{N}}$-valued processes, the analogous conclusions for $\mathcal{P}_{\mathbb{N}:k}$-valued processes are nearly immediate. The key observation is that the projection of $\mathbf{X}$ into $\mathcal{P}_{\mathbb{N}:k}$ preserves the Markov property only if the transition law of $\mathbf{X}$ treats the labels $[k]$ symmetrically, which requires row–column exchangeability of Σ and $\mathbf{c}_{ii'} = \mathbf{c}_{jj'} = c$ for all $i \neq i'$, $j \neq j'$.

1.7. *Examples*. We illustrate our main theorems with three examples: two exchangeable, consistent Markov processes on $[k]^{\mathbb{N}}$ (one in discrete-time and one in continuous-time) and a family of exchangeable Markov chains that is not consistent (the Ehrenfest walk on the hypercube). Example 1.9 shows why discrete-time chains cannot admit single-index flips.

EXAMPLE 1.7 (A reversible discrete-time chain [6]). For $\alpha > 0$, we define transition probabilities

$$P_n(x, x') := \prod_{i=1}^{k} \frac{\prod_{i'=1}^{k} (\alpha/k)^{\uparrow \mathbf{n}_{ii'}(x,x')}}{\alpha^{\uparrow \mathbf{n}_i(x)}}, \qquad x, x' \in [k]^{[n]}, \tag{1.8}$$

where $\mathbf{n}_{ii'}(x,x') := \#\{j \in [n] : x^j = i \text{ and } x'^j = i'\}$, $\mathbf{n}_i(x) := \#\{j \in [n] : x^j = i\}$, and $\alpha^{\uparrow n} := \alpha(\alpha+1)\cdots(\alpha+n-1)$. This transition probability is reversible with respect to

$$\lambda_{\xi}^{(n)}(x) = \frac{\prod_{i=1}^{k} \alpha^{\uparrow \mathbf{n}_i(x)}}{(k\alpha)^{\uparrow n}}, \qquad x \in [k]^{[n]},$$

and projects to a transition probability on $\mathcal{P}_{[n]:k}$ (partitions of $[n]$ with at most k blocks) with reversible stationary distribution

$$\varrho_{\xi}^{(n)}(\pi) := \frac{k!}{(k-\#\pi)!} \frac{\prod_{b\in\pi} \alpha^{\uparrow \#b}}{(k\alpha)^{\uparrow n}}, \qquad \pi \in \mathcal{P}_{[n]:k},$$

where $\#\pi$ denotes the number of blocks of π and $\#b$ denotes the cardinality of $b \subseteq [n]$.

Namely, in Theorem 1.1, the transition probabilities in (1.8) correspond to the homogeneous cut-and-paste chain with $\Sigma_{\alpha/k} = \xi_{\alpha/k} \otimes \cdots \otimes \xi_{\alpha/k}$, where ξ_α is the symmetric Dirichlet distribution with parameter $(\alpha, \dots, \alpha)$. That is, $S \sim \Sigma_{\alpha/k}$ is a random matrix whose rows are independent and identically distributed from Dirichlet$(\alpha/k, \dots, \alpha/k)$.

EXAMPLE 1.8 (A purely continuous process). For $\mathbf{c}_{12}, \mathbf{c}_{21} > 0$, let each coordinate of $\mathbf{X} = (X_t, t \geq 0)$ evolve independently, each jumping from 1 to 2 at rate $\mathbf{c}_{12}$ and from 2 to 1 at rate $\mathbf{c}_{21}$. The projection of $\mathbf{X}$ into the simplex evolves continuously and deterministically by a constant interchange of mass between the colors 1 and 2. Eventually, the projection settles to the fixed point

$$\left(\frac{c_{21}}{c_{12} + c_{21}}, \frac{c_{12}}{c_{12} + c_{21}}\right).$$

The projection into $\mathcal{P}_{\mathbb{N}:k}$ is Markov only if $c_{12} = c_{21}$. In this case, the projection settles to $(1/2, 1/2)$ and, in equilibrium, there is a constant and equal flow of mass between the two blocks.

EXAMPLE 1.9 (Nonexample: Ehrenfest chain on $\{0, 1\}^{[n]}$). The family of discrete-time Ehrenfest chains on the hypercubes $\{0, 1\}^{[n]}$, $n \in \mathbb{N}$, is not consistent, and thus, not covered by our theory. On $\{0, 1\}^{[n]}$, an Ehrenfest chain $\mathbf{X}^{[n]}$ evolves by choosing a coordinate $1, \dots, n$ uniformly at random and then flipping a fair coin to decide its value at the next time. All other coordinates remain unchanged. In the language of Section 1.6, all transitions of this chain are single-index flips.

The finite-dimensional chains are exchangeable but not consistent. For any $n \in \mathbb{N}$, the probability that $\mathbf{X}^{[n]}$ remains in the same state after a transition is $1/2$, whereas the projection of an Ehrenfest chain $\mathbf{X}^{[n+1]}$ on $\{0, 1\}^{[n+1]}$ into $\{0, 1\}^{[n]}$ remains in the same state with probability $(n + 2)/(2n + 2) \neq 1/2$.

Six sections compose the paper. In Section 2, we lay out definitions and notation; in Section 3, we establish Theorems 1.1 and 1.2; in Section 4, we prove Theorems 1.3 and 1.4; in Section 5, we deduce Theorems 1.5 and 1.6; in Section 6, we conclude.

2. Preliminaries.

2.1. *Notation.* Throughout the paper, we write x to denote a k-coloring, X a random k-coloring and $\mathbf{X}$ a random collection of k-colorings. We write π to denote a partition, Π a random partition, and $\mathbf{\Pi}$ a random collection of partitions. For terminology and notation pertaining to both k-colorings and partitions, we write λ, Λ, and $\mathbf{\Lambda}$, as appropriate. A collection $\mathbf{\Lambda} = (\Lambda_m, m \geq 0)$ indexed by m evolves in discrete-time, that is, $m = 1, 2, \dots,$ and $\mathbf{\Lambda} = (\Lambda_t, t \geq 0)$ indexed by t evolves in continuous-time, that is, $t \in [0, \infty)$.

2.2. *Partitions and colorings.* For fixed $k \in \mathbb{N}$, a *k-coloring* of $[n] = \{1, \ldots, n\}$ is a $[k]$-valued sequence $x = x^1 \cdots x^n$. A *partition* of $[n]$ is a collection $\pi = \{B_1, \ldots, B_r\}$ of nonempty, disjoint subsets (blocks) satisfying $\bigcup_{j=1}^r B_j = [n]$. We can also regard π as an equivalence relation $\sim_\pi$, where

$$i \sim_\pi j \quad \Longleftrightarrow \quad i \text{ and } j \text{ are in the same block of } \pi.$$

Upon removal of its colors, any k-coloring x projects to a unique partition $\mathcal{B}_n(x)$ of $[n]$, as in (1.7). For $n \in \mathbb{N}$, we write $[k]^{[n]}$ to denote the set of k-colorings of $[n]$, $\mathcal{P}_{[n]}$ to denote the set of partitions of $[n]$, and $\mathcal{P}_{[n]:k}$ to denote the subset of partitions of $[n]$ with at most k blocks.

Any one-to-one mapping $\varphi : [m] \to [n]$, $m \le n$, determines a map $[k]^{[n]} \to [k]^{[m]}$, $x \mapsto x^\varphi$, where

$$x^\varphi = x^{\varphi(1)} \cdots x^{\varphi(m)}. \tag{2.1}$$

We call the image in (2.1) a *composite mapping* because $x \mapsto x^\varphi$ can be obtained by composing the relabeling and restriction operations in (1.1) and (1.2). Let $\mathbf{R}_{m,n}$ denote the restriction map $[k]^{[n]} \to [k]^{[m]}$, that is, $\mathbf{R}_{m,n}x = x^{[m]}$. To any one-to-one map $\varphi : [m] \to [n]$, there exists a permutation $\sigma : [n] \to [n]$ such that $x^\varphi = \mathbf{R}_{m,n}(x^\sigma)$, relabeling by σ followed by restriction to $[k]^{[m]}$.

For a partition $\pi \in \mathcal{P}_{[n]}$, relabeling, restriction and composite operations are defined by $\pi \mapsto \pi^\sigma$, $\pi \mapsto \pi^{[m]}$, and $\pi \mapsto \pi^\varphi$, respectively, where

$$\begin{aligned} i \sim_{\pi^\sigma} j \quad &\Longleftrightarrow \quad \sigma(i) \sim_\pi \sigma(j), \\ i \sim_{\pi^{[m]}} j \quad &\Longleftrightarrow \quad i \sim_\pi j \quad \text{and} \\ i \sim_{\pi^\varphi} j \quad &\Longleftrightarrow \quad \varphi(i) \sim_\pi \varphi(j). \end{aligned}$$

When convenient, we abuse notation and also write $\mathbf{R}_{m,n}$ to denote the restriction $\mathcal{P}_{[n]} \to \mathcal{P}_{[m]}$, that is, $\mathbf{R}_{m,n}\pi = \pi^{[m]}$, so that $\pi^\varphi = \mathbf{R}_{m,n}(\pi^\sigma)$ for some $\sigma : [n] \to [n]$.

Any finite k-coloring can be embedded into a k-coloring of $\mathbb{N}$, and likewise for partitions. A k-coloring of $\mathbb{N}$ is an infinite $[k]$-valued sequence $x = x^1x^2 \cdots$ and is determined by its sequence of finite restrictions $(x^{[1]}, x^{[2]}, \ldots)$. A partition of $\mathbb{N}$ is defined similarly as a sequence of finite partitions $(\pi^{[1]}, \pi^{[2]}, \ldots)$ for which $\pi^{[m]} = \mathbf{R}_{m,n}\pi^{[n]}$, for every $m \le n$. As for finite sets, we denote k-colorings of $\mathbb{N}$ by $[k]^{\mathbb{N}}$, partitions of $\mathbb{N}$ by $\mathcal{P}_{\mathbb{N}}$, and partitions of $\mathbb{N}$ with at most k blocks by $\mathcal{P}_{\mathbb{N}:k}$.

For each $n \in \mathbb{N}$, $\mathbf{R}_n$ denotes the restriction map $[k]^{\mathbb{N}} \to [k]^{[n]}$, or $\mathcal{P}_{\mathbb{N}} \to \mathcal{P}_{[n]}$. The projective nature of both $[k]^{\mathbb{N}}$ and $\mathcal{P}_{\mathbb{N}}$ endows each with a natural product-discrete topology. With λ, λ' denoting objects both in either $[k]^{\mathbb{N}}$ or $\mathcal{P}_{\mathbb{N}}$, we define the ultrametric d by

$$d(\lambda, \lambda') := 2^{-n(\lambda,\lambda')}, \tag{2.2}$$

where $n(\lambda, \lambda') := \max\{n \in \mathbb{N} : \mathbf{R}_n\lambda = \mathbf{R}_n\lambda'\}$. Under (2.2), both $[k]^{\mathbb{N}}$ and $\mathcal{P}_{\mathbb{N}}$ are compact, separable and, therefore, Polish, metric spaces. We equip $[k]^{\mathbb{N}}$ and $\mathcal{P}_{\mathbb{N}:k}$ with their discrete σ-fields, $\sigma\langle\bigcup_{n=1}^\infty [k]^{[n]}\rangle$ and $\sigma\langle\bigcup_{n=1}^\infty \mathcal{P}_{[n]}\rangle$, respectively.

2.3. *Exchangeability.* An infinite sequence $X := (X_1, X_2, \dots)$ of random variables is called *exchangeable* if its law is invariant under finite permutations of its indices, that is, for each $n \in \mathbb{N}$,

$$(X_{\sigma(1)}, \dots, X_{\sigma(n)}) =_{\mathcal{L}} (X_1, \dots, X_n) \qquad \text{for every } \sigma \in \mathscr{S}_n,$$

where $\mathscr{S}_n$ denotes the symmetric group of permutations of $[n]$. By de Finetti's theorem (see, e.g., Aldous [1]), the law of any exchangeable sequence $X \in [k]^{\mathbb{N}}$ is determined by a unique directing probability measure ν on the $(k-1)$-dimensional simplex

$$\Delta_k := \left\{(s_1, \dots, s_k) : s_i \geq 0, \sum_{i=1}^{k} s_i = 1\right\}.$$

In particular, conditional on $s \sim \nu$, $X_1, X_2, \dots$ are independent and identically distributed according to

$$\mathbb{P}_s\{X_1 = j\} = s_j, \qquad j = 1, \dots, k.$$

A random partition Π is exchangeable if $\Pi =_{\mathcal{L}} \Pi^{\sigma}$ for all $\sigma \in \mathscr{S}_{\mathbb{N}}$, where $\mathscr{S}_{\mathbb{N}}$ is the set of *finite* permutations of $\mathbb{N}$, that is, permutations $\sigma : \mathbb{N} \to \mathbb{N}$ that fix all but finitely many elements. Through (1.7), any exchangeable $[k]$-valued sequence X projects to an exchangeable random partition $\Pi := \mathcal{B}(X)$. This construction of Π is a special case of Kingman's paintbox representation for exchangeable random partitions of $\mathbb{N}$ [12]. If X is directed by ν, then we denote the law of $\Pi = \mathcal{B}(X)$ by ϱ_ν, the *paintbox measure* directed by ν.

With $f_i(X)$ defined in (1.4), the asymptotic frequency $|X| = (f_1(X), \dots, f_k(X))$ of any exchangeable k-coloring exists almost surely. Likewise for the asymptotic frequency of an exchangeable partition Π, denoted $|\Pi|^{\downarrow}$, the vector of asymptotic block frequencies listed in decreasing order of size which lives in the ranked k-simplex $\Delta_k^{\downarrow} := \{(s_1, \dots, s_k) : s_1 \geq \cdots \geq s_k \geq 0, \ \sum_i s_i = 1\}$.

REMARK 2.1. To avoid measurability concerns, we can add the point ∂ to both Δ_k and $\Delta_k^{\downarrow}$ and put $|x| = \partial$ (resp., $|\pi|^{\downarrow} = \partial$) whenever the asymptotic frequency of $x \in [k]^{\mathbb{N}}$ (resp., $\pi \in \mathcal{P}_{\mathbb{N}:k}$) does not exist. We equip Δ_k, respectively, $\Delta_k^{\downarrow}$, with the σ-field generated by $|\cdot| : [k]^{\mathbb{N}} \to \Delta_k \cup \{\partial\}$ and $|\cdot|^{\downarrow} : \mathcal{P}_{\mathbb{N}:k} \to \Delta_k^{\downarrow} \cup \{\partial\}$, respectively. Beyond this point, issues of measurability never arise, and so neither does the above formalism.

2.4. *Exchangeable Markov processes.* Let $\mathbf{X} = (X_t, t \in T)$ be a random collection in $[k]^{\mathbb{N}}$, with T either $\mathbb{Z}_+ = \{0, 1, \dots\}$ (discrete-time) or $\mathbb{R}_+ = [0, \infty)$ (continuous-time). We say $\mathbf{X}$ is *Markovian* if, for every $t, t' \geq 0$, the conditional law of $X_{t+t'}$, given $\mathcal{F}_t := \sigma\langle X_s, s \leq t\rangle$, depends only on X_t and t'. Specifically, we distinguish between collections with finitely many jumps in bounded intervals (Markov chains) and those with infinitely many jumps in bounded intervals

(Markov processes). When speaking generally, we use the terminology and notation of Markov processes as a catch-all.

The *Markov semigroup* $\mathbf{P} = (\mathbf{P}_t, t \in T)$ of $\mathbf{X} = (X_t, t \in T)$ is defined for all bounded, measurable functions $g : [k]^{\mathbb{N}} \to \mathbb{R}$ by

$$\mathbf{P}_t g(x) := \mathbb{E}_x g(X_t), \qquad t \in T, \tag{2.3}$$

the conditional expectation of $g(X_t)$ given $X_0 = x$. We say $\mathbf{X}$ *enjoys the Feller property*, or is a *Feller process*, if for every bounded, continuous $g : [k]^{\mathbb{N}} \to \mathbb{R}$, its semigroup $\mathbf{P}$ satisfies:

- $\lim_{t \downarrow 0} \mathbf{P}_t g(x) = g(x)$ for all $x \in [k]^{\mathbb{N}}$ and
- $x \mapsto \mathbf{P}_t g(x)$ is continuous for all $t \in T$.

In general, since each $\mathbf{R}_n : [k]^{\mathbb{N}} \to [k]^{[n]}$ is a many-to-one function, the restriction $\mathbf{X}^{[n]}$ need not be Markovian. Under the product-discrete topology induced by (2.2), exchangeability and consistency are equivalent to exchangeability and the Feller property, and so we use the terms *consistency* and *Feller* interchangeably.

PROPOSITION 2.2. *The following are equivalent for a Markov process $\mathbf{\Lambda}$ on either $[k]^{\mathbb{N}}$ or $\mathcal{P}_{\mathbb{N}:k}$:*

(i) $\mathbf{\Lambda}$ *is exchangeable and consistent under subsampling.*
(ii) $\mathbf{\Lambda}$ *is exchangeable and enjoys the Feller property.*

2.5. *Coset decompositions and associated mappings.* For fixed $k \in \mathbb{N}$, we define the *coset decomposition* of $x \in [k]^{\mathbb{N}}$ by the k-tuple $(x_1, \dots, x_k)$, where

$$x_i = x^i x^{i+k} x^{i+2k} \cdots, \qquad i = 1, \dots, k. \tag{2.4}$$

In words, the *ith coset* of x is the subsequence of x including every kth element, beginning at coordinate i. Through (2.4), the sets $[k]^{\mathbb{N}}$ and $[k]^{\mathbb{N}\otimes k} \cong [k]^{\mathbb{N}} \times \cdots \times [k]^{\mathbb{N}}$ (k times) are in one-to-one correspondence, but we sometimes prefer one representation over the other. To distinguish between representations, we write:

- $x = x^1 x^2 \cdots$ to denote the object in $[k]^{\mathbb{N}}$ and
- $x = (x_1, \dots, x_k)$ to denote the coset representation in $[k]^{\mathbb{N}\otimes k}$, with each coset written

$$x_i = x_i^1 x_i^2 \cdots = x^i x^{i+k} \cdots.$$

We usually write x to denote an object *initially* defined in $[k]^{\mathbb{N}}$ and M to denote an object *initially* defined in $[k]^{\mathbb{N}\otimes k}$. The importance of this decomposition becomes apparent in Section 3.

For $n \in \mathbb{N}$, the restriction of $M \in [k]^{\mathbb{N}\otimes k}$ to $[k]^{[n]\otimes k} \cong [k]^{[n]} \times \cdots \times [k]^{[n]}$ (k times) is defined componentwise by

$$M^{[n]} := (M_1^{[n]}, \dots, M_k^{[n]}). \tag{2.5}$$

Likewise, a k-tuple of finite permutations $\sigma_1,\dots,\sigma_k:\mathbb{N}\to\mathbb{N}$ acts on $M\in[k]^{\mathbb{N}}$ by

$$M^{\sigma_1,\dots,\sigma_k} := \big(M_1^{\sigma_1},\dots,M_k^{\sigma_k}\big). \tag{2.6}$$

Any $M\in[k]^{[n]\otimes k}$ functions as a map $[k]^{[n]}\to[k]^{[n]}$. For each $x\in[k]^{[n]}$, we define the injection $\varphi_x:[n]\to[nk]$ by

$$\varphi_x(j) := x^j + (j-1)k, \qquad j=1,\dots,n. \tag{2.7}$$

For any $M\in[k]^{\mathbb{N}\otimes k}$, its restriction $M^{[n]}$ to $[k]^{[n]\otimes k}$, as in (2.5), is in correspondence with a unique k-coloring $M^1\cdots M^{nk}$ of $[nk]$. Using (2.1), we define $M^{[n]}:[k]^{[n]}\to[k]^{[n]}$ by

$$M^{[n]}(x) := M^{\varphi_x} = M^{x^1}M^{x^2+k}\cdots M^{x^n+(n-1)k}, \qquad x\in[k]^{[n]}. \tag{2.8}$$

The finite maps $(M^{[n]}, n\in\mathbb{N})$ derived from M determine a unique map $M:[k]^{\mathbb{N}}\to[k]^{\mathbb{N}}$.

Importantly, each $M\in[k]^{\mathbb{N}\otimes k}$ determines a Lipschitz continuous map in the metric (2.2). The identity map $\mathrm{id}_k:[k]^{\mathbb{N}}\to[k]^{\mathbb{N}}$ corresponds to the infinite repeating pattern $12\cdots k$,

$$\mathbf{Z}_k = 12\cdots k12\cdots k\cdots, \tag{2.9}$$

for example, $\mathbf{Z}_2 = 121212\cdots$, $\mathbf{Z}_3 = 123123\cdots$, and so on. The coset decomposition of $\mathbf{Z}_k$ is $(\mathbf{1},\mathbf{2},\dots,\mathbf{k})$, where $\mathbf{i}=iii\cdots$ is the infinite sequence of all i's, for each $i=1,2,\dots$. For $n\in\mathbb{N}$, we write $\mathbf{Z}_{k,n}$ to denote the restriction of $\mathbf{Z}_k$ to $[k]^{[n]\otimes k}$ and $\mathrm{id}_{k,n}$ to denote its associated identity map $[k]^{[n]}\to[k]^{[n]}$. By definition (2.6), $\mathbf{Z}_k$, and hence $\mathbf{Z}_{k,n}$, is invariant under relabeling by any k-tuple $\sigma_1,\dots,\sigma_k$ of permutations.

Since any mapping $M:[k]^{\mathbb{N}}\to[k]^{\mathbb{N}}$ is determined by its coset decomposition $(M_1,\dots,M_k)$, we can define the *asymptotic frequency* of M by the k-tuple $(|M_1|,\dots,|M_k|)$, provided each $|M_i|$ exists. We express the asymptotic frequency of M as a stochastic matrix $|M|_k = S = (S_{ii'}, 1\le i,i'\le k)$, where

$$S_{ii'} := \lim_{n\to\infty} n^{-1}\sum_{j=1}^{n}\mathbf{1}\{M_i^j = i'\}, \qquad 1\le i,i'\le k. \tag{2.10}$$

3. Discrete-time cut-and-paste chains. In this section, $\mathbf{X}=(X_m, m\ge 0)$ denotes a discrete-time exchangeable and consistent Markov chain on $[k]^{\mathbb{N}}$, and $\mathbf{X}^{[n]} = (X_m^{[n]}, m\ge 0)$ its restriction to $[k]^{[n]}$, for each $n=1,2,\dots$. By assumptions (A) and (B), each $\mathbf{X}^{[n]}$ is an exchangeable Markov chain with transition probability measure

$$P_n(x,x') := \mathbb{P}\{X_1 = x' \mid X_0 = x\}, \qquad x,x'\in[k]^{[n]}.$$

Exchangeability implies $P_n(x,x') = P_n(x^\sigma, x'^\sigma)$ for all permutations $\sigma:[n]\to[n]$, while consistency relates $(P_n, n\in\mathbb{N})$ through

$$P_m(x,x') = P_n\big(x^*, \mathbf{R}_{m,n}^{-1}(x')\big), \qquad x,x'\in[k]^{[m]},$$

for all $x^* \in \mathbf{R}_{m,n}^{-1}(x) = \{\hat{x} \in [k]^{[n]} : \hat{x}^{[m]} = x\}$. Writing P to denote the transition probability measure of $\mathbf{X}$ on $[k]^{\mathbb{N}}$, we conclude

$$P_n(x, x') = P(x^*, \mathbf{R}_n^{-1}(x')), \qquad x, x' \in [k]^{[n]}, \text{for all } x^* \in \mathbf{R}_n^{-1}(x), \tag{3.1}$$

for every $n \in \mathbb{N}$.

Theorem 1.1 asserts that P is determined by a unique probability measure Σ on $\mathcal{S}_k$. We construct Σ directly from P using the connection between k-colorings and stochastic matrices from Section 2.5. For $\mathbf{Z}_k$ in (2.9), we define a probability measure χ on $[k]^{\mathbb{N} \otimes k}$ by

$$\chi(\cdot) := P(\mathbf{Z}_k, \cdot). \tag{3.2}$$

DEFINITION 3.1 (Coset exchangeability). A random mapping $M = (M_1, \dots, M_k) \in [k]^{\mathbb{N} \otimes k}$ is *coset exchangeable* if

$$(M_1, \dots, M_k) =_{\mathcal{L}} \left(M_1^{\sigma_1}, \dots, M_k^{\sigma_k}\right) \qquad \text{for all } \sigma_1, \dots, \sigma_k \in \mathscr{S}_{\mathbb{N}}. \tag{3.3}$$

For any random mapping M constructed from a random k-coloring through (2.4), exchangeability implies coset exchangeability, but not the reverse. By assumption, P is an exchangeable transition probability on $[k]^{\mathbb{N}}$ and the coset decomposition of $\mathbf{Z}_k$ is invariant under coset relabeling (2.6); hence, χ defined in (3.2) is coset exchangeable and the asymptotic frequency of $M \sim \chi$, as defined in (2.10), exists with probability one. We denote the law of $|M|_k$ by $|\chi|_k$.

We complete the proof of Theorem 1.1 by showing that a random k-coloring X' generated by first drawing $M \sim \chi$ and then putting $X' = M(x)$, for fixed $x \in [k]^{\mathbb{N}}$, is a draw from $P(x, \cdot)$. By consistency, we need only show that $M^{[n]}(x) \sim P_n(x, \cdot)$ for every $x \in [k]^{[n]}$, for every $n \in \mathbb{N}$. We have defined $\mathbf{Z}_k$ so that

$$\mathbf{Z}_k(x) = \mathbf{Z}_{k,n}^{\varphi_x} = \mathbf{Z}_k^{x^1} \cdots \mathbf{Z}_k^{x^n + (n-1)k} = x^1 \cdots x^n = x \qquad \text{for all } x \in [k]^{[n]}.$$

By (3.1) and (3.2), the restriction of $M \sim \chi$ to $[k]^{[n] \otimes k}$ is distributed as

$$M^{[n]} \sim \chi^{(n)}(\cdot) = P_{nk}(\mathbf{Z}_{k,n}, \cdot),$$

which combines with (2.8) to imply $M^{[n]}(x) \sim P_n(x, \cdot)$.

We have proven the following prelude to Theorem 1.1.

THEOREM 3.2. *Let $\mathbf{X} = (X_m, m \geq 0)$ be a discrete-time, exchangeable, consistent Markov chain on $[k]^{\mathbb{N}}$. Then there exists a probability measure χ on $[k]^{\mathbb{N} \otimes k}$ such that $\mathbf{X}^* = (X_m^*, m \geq 0)$ is a version of $\mathbf{X}$, where $X_0^* =_{\mathcal{L}} X_0$ and*

$$X_m^* = (M_m \circ \cdots \circ M_1)(X_0^*), \qquad m \geq 1,$$

for $M_1, M_2, \dots$ drawn i.i.d. from χ.

To establish Theorem 1.1, we must show that χ is determined by a unique probability measure on $\mathcal{S}_k$. By (2.10) and coset exchangeability, χ induces a probability measure $|\chi|_k$ on $\mathcal{S}_k$. By de Finetti's theorem, the components of $M \sim \chi$, given $|M|_k = S$, are conditionally independent with distribution

$$\mathbb{P}_S\{M^{i+(j-1)k} = i'\} = S_{ii'}, \qquad i, i' = 1, \dots, k;\ j = 1, 2, \dots. \tag{3.4}$$

We write μ_S to denote the conditional distribution of M, given $|M|_k = S$, as in (3.4) and

$$\mu_\Sigma(\cdot) := \int_{\mathcal{S}_k} \mu_S(\cdot)\Sigma(dS) \tag{3.5}$$

to denote the mixture of μ_S-measures with respect to Σ. By (3.4), the components $Y^1 Y^2 \cdots$ of $M(x)$ are conditionally independent given $|M|_k = S$ and have distribution

$$\mathbb{P}_S\{Y^j = i' \mid x^j = i\} = S_{ii'}, \qquad j = 1, 2, \dots. \tag{3.6}$$

For every $n \in \mathbb{N}$, the unconditional law of $M^{[n]}(x)$ is thus

$$P_n(x, x') = \int_{\mathcal{S}_k} \prod_{j=1}^{n} S(x^j, x'^j)|\chi|_k(dS), \qquad x' \in [k]^{[n]}.$$

Putting $\Sigma := |\chi|_k$ establishes Theorem 1.1.

REMARK 3.3. We call $\mathbf{X}^*_\Sigma$ in Theorem 1.1 an *(exchangeable) cut-and-paste chain* with directing measure Σ and cut-and-paste measure μ_Σ.

From Theorems 1.1 and 3.2, we can generate a version of $\mathbf{X}$ by drawing X_0 from the initial distribution of $\mathbf{X}$ and $M_1, M_2, \dots$ i.i.d. from μ_Σ. Given $X_0, M_1, M_2, \dots,$ we define

$$X_m := M_m(X_{m-1}) = (M_m \circ \cdots \circ M_1)(X_0), \qquad m \geq 1. \tag{3.7}$$

By de Finetti's theorem, $|X_0| = (f_1(X_0), \dots, f_k(X_0))$ exists almost surely and $|M_1|_k, |M_2|_k, \dots$ is an i.i.d. sequence from Σ. By the construction of $\mathbf{X}$ in (3.7), X_1 is chosen from the conditional transition probability in (3.6), with $S = |M_1|_k$. By the strong law of large numbers, $f_{i'}(X_1)$ exists almost surely for every $i' = 1, \dots, k$ and equals the i'th component of $|X_0|S_1$, that is,

$$f_{i'}(X_1) = \sum_{i=1}^{k} f_i(X_0)S_1(i, i').$$

By induction, the components of $|X_m|$, given $|X_{m-1}|$ and $|M_m|_k$, equal

$$|X_{m-1}||M_m|_k = |X_0||M_1|_k \cdots |M_m|_k \qquad \text{for every } m \geq 1,$$

and Theorem 1.2 follows.

4. Continuous-time cut-and-paste processes. We now let $\mathbf{X} = (X_t, t \geq 0)$ denote an exchangeable, consistent Markov process in continuous-time. We have noted previously that $\mathbf{X}$ can jump infinitely often in bounded intervals, but its finite restrictions can jump only finitely often. To characterize the behavior of $\mathbf{X}$, we use a Poisson point process to build a version sequentially through its finite restrictions. Similar to our discrete-time construction (3.7), we define the intensity measure of the Poisson point process directly from the transition law of $\mathbf{X}$. Dissimilar to the discrete-time case, this intensity need not be finite.

Let χ be a coset exchangeable measure on $[k]^{\mathbb{N}\otimes k}$ satisfying

$$\chi(\{\mathrm{id}_k\}) = 0 \quad \text{and} \quad \chi(\{M \in [k]^{\mathbb{N}\otimes k} : M^{[n]} \neq \mathrm{id}_{k,n}\}) < \infty \qquad \text{for all } n \in \mathbb{N}. \tag{4.1}$$

We construct a process $\mathbf{X}^*_\chi = (X^*_t, t \geq 0)$ through its finite restrictions $(\mathbf{X}^{*[n]}_\chi, n \in \mathbb{N})$ as follows. Let $\mathbf{M} = \{(t, M_t)\} \subseteq \mathbb{R}_+ \times [k]^{\mathbb{N}\otimes k}$ be a Poisson point process with intensity $dt \otimes \chi$, where dt denotes Lebesgue measure on $[0,\infty)$. Given an exchangeable initial state $X_0 \in [k]^{\mathbb{N}}$, we put $X^{*[n]}_0 = X^{[n]}_0$ and, for each $t > 0$:

- if $t > 0$ is an atom time of $\mathbf{M}$ for which $M^{[n]}_t \neq \mathrm{id}_{k,n}$, we put $X^{*[n]}_t := M^{[n]}_t(X^{*[n]}_{t-})$,
- otherwise, we put $X^{*[n]}_t = X^{*[n]}_{t-}$.

This construction of each $\mathbf{X}^{*[n]}_\chi$ is a continuous-time analog to the discrete-time construction in (3.7); it differs only in the random time between jumps and the possibility of infinitely many jumps in the limiting process. We have constructed each $\mathbf{X}^{*[n]}_\chi$ from the same Poisson process so that $(\mathbf{X}^{*[n]}_\chi, n \in \mathbb{N})$ is compatible, that is, $X^{*[m]}_t = \mathbf{R}_{m,n} X^{*[n]}_t$ for all $t \geq 0$ and $m \leq n$, and determines a unique $[k]^{\mathbb{N}}$-valued process $\mathbf{X}^*_\chi$.

PROPOSITION 4.1. *Let χ be a coset exchangeable measure on $[k]^{\mathbb{N}\otimes k}$ that satisfies* (4.1), *and let $\mathbf{X}^*_\chi$ be as constructed from the Poisson point process $\mathbf{M}$ with intensity $dt \otimes \chi$. Then $\mathbf{X}^*_\chi$ is an exchangeable*, *consistent Markov process on $[k]^{\mathbb{N}}$.*

PROOF. For each $n \in \mathbb{N}$, $\mathbf{X}^{*[n]}_\chi$ is a Markov chain by assumption (4.1) and its Poisson point process construction. Moreover, $\mathbf{R}_{m,n} X^{*[n]}_t = X^{*[m]}_t$ for all $t \geq 0$, for all $m \leq n$, and so $(\mathbf{X}^{*[n]}_\chi, n \in \mathbb{N})$ determines a unique Markov process $\mathbf{X}^*_\chi$ on $[k]^{\mathbb{N}}$. Exchangeability of $\mathbf{X}^*_\chi$ follows by coset exchangeability of χ, since all of its finite restrictions to $[k]^{[n]\otimes k}$ are finite, coset exchangeable measures. $\square$

COROLLARY 4.2. *Every coset exchangeable measure χ on $[k]^{\mathbb{N}\otimes k}$ satisfying* (4.1) *determines the jump rates of an exchangeable Feller process on $[k]^{\mathbb{N}}$.*

A measure satisfying (4.1) can be constructed directly from the transition rates of $\mathbf{X}$. By assumption, each finite restriction $\mathbf{X}^{[n]} = (X_t^{[n]}, t \geq 0)$ is a càdlàg, exchangeable Markov process on $[k]^{[n]}$. Since $[k]^{[n]}$ is finite, the evolution of $\mathbf{X}^{[n]}$ is characterized by its jump rates

$$Q_n(x, x') := \lim_{t \downarrow 0} \frac{1}{t} \mathbb{P}(X_t^{[n]} = x' \mid X_0^{[n]} = x), \qquad x \neq x' \in [k]^{[n]}, \tag{4.2}$$

which satisfy

$$Q_n(x, [k]^{[n]} \setminus \{x\}) < \infty \qquad \text{for all } x \in [k]^{[n]}, \tag{4.3}$$

are exchangeable in the sense that, for every $\sigma \in \mathscr{S}_n$,

$$Q_n(x, x') = Q_n(x^\sigma, x'^\sigma), \qquad x \neq x' \in [k]^{[n]}, \tag{4.4}$$

and are consistent,

$$\begin{aligned} Q_m(x, x') = Q_n(x^*, \mathbf{R}_{m,n}^{-1}(x')),& \\ x \neq x' \in [k]^{[m]}, \text{ for all } x^* \in \mathbf{R}_{m,n}^{-1}(x).& \end{aligned} \tag{4.5}$$

For each $n \in \mathbb{N}$, we define

$$\chi_n(M) := Q_n(\mathbf{Z}_{k,n}, M), \qquad M \in [k]^{[n]\otimes k} \setminus \{\mathrm{id}_{k,n}\}. \tag{4.6}$$

LEMMA 4.3. *The collection* $(\chi_n, n \in \mathbb{N})$ *in* (4.6) *is coset exchangeable and satisfies*

$$\chi_m(M) = \chi_n(\{M^* \in [k]^{[n]\otimes k} : M^{*[m]} = M\}) \qquad \textit{for all } M \in [k]^{[m]\otimes k},$$

for all $m \leq n$.

PROOF. This follows from the definition of χ_n in (4.6), the correspondence $[k]^{\mathbb{N}} \leftrightarrow [k]^{\mathbb{N}\otimes k}$ in (2.4), and conditions (4.3), (4.4) and (4.5). $\square$

PROPOSITION 4.4. *Let* $(\chi_n, n \in \mathbb{N})$ *be defined in* (4.6). *Then there exists a unique coset exchangeable measure* χ *on* $[k]^{\mathbb{N}\otimes k}$ *satisfying* (4.1) *and*

$$\begin{aligned} \chi(\{M^* \in [k]^{\mathbb{N}\otimes k} : M^{*[n]} = M\}) = \chi_n(M),& \\ M \in [k]^{[n]\otimes k} \setminus \{\mathrm{id}_{k,n}\}, \textit{ for every } n \in \mathbb{N}.& \end{aligned}$$

PROOF. Because $\bigcup_{n=1}^\infty [k]^{[n]\otimes k}$ is a generating π-system of the product σ-field over $[k]^{\mathbb{N}\otimes k}$, we need only determine χ on subsets of the form

$$\{M^* \in [k]^{\mathbb{N}\otimes k} : M^{*[n]} = M\},$$

for every $n \in \mathbb{N}$ and $M \in [k]^{[n]\otimes k}$. Lemma 4.3 implies

$$\chi_m(M) = \chi_n(\{M^* \in [k]^{[n]\otimes k} : M^{*[m]} = M\}) = \sum_{M^* \in [k]^{[n]\otimes k} : M^{*[m]} = M} \chi_n(M^*),$$

for all $m \le n$ and $M \in [k]^{[m]\otimes k}$. Therefore, χ defined by

$$\chi\big(\{M^* \in [k]^{\mathbb{N}\otimes k} : M^{*[n]} = M\}\big) = \chi_n(M), \qquad M \in [k]^{[n]\otimes k} \setminus \{\mathrm{id}_{k,n}\}, \tag{4.7}$$

is additive, and Caratheodory's extension theorem implies χ has a unique extension to a measure on $[k]^{\mathbb{N}\otimes k} \setminus \{\mathrm{id}_k\}$.

To satisfy the first half of (4.1), we simply put $\chi(\{\mathrm{id}_k\}) = 0$. For the second half, (4.3) implies

$$\begin{aligned}\chi\big(\{M \in [k]^{\mathbb{N}\otimes k} : M^{[n]} \neq \mathrm{id}_{k,n}\}\big) &= \chi_n\big([k]^{[n]\otimes k} \setminus \{\mathrm{id}_{k,n}\}\big) \\ &= Q_{nk}\big(\mathbf{Z}_{k,n}, [k]^{[nk]} \setminus \{\mathbf{Z}_{k,n}\}\big) < \infty.\end{aligned}$$

This completes the proof. $\square$

The measure χ in Proposition 4.4 ties the Poissonian construction of $\mathbf{X}^*_\chi$ to $\mathbf{X}$, as the next theorem shows.

THEOREM 4.5. *Let $\mathbf{X}$ be a continuous-time, exchangeable, consistent Markov process on $[k]^{\mathbb{N}}$. Then there exists a coset exchangeable measure χ on $[k]^{\mathbb{N}\otimes k}$ satisfying* (4.1) *such that $\mathbf{X}^*_\chi$ is a version of $\mathbf{X}$.*

PROOF. Let χ be the coset exchangeable measure with finite-dimensional distributions (4.6). By Proposition 4.4, χ satisfies (4.1).

Let $\mathbf{X}^*_\chi$ be the Markov process constructed from $\mathbf{M}$ with intensity $dt \otimes \chi$. The total intensity at which events occur in $\mathbf{M}$ is $\chi([k]^{\mathbb{N}\otimes k})$. For $n \in \mathbb{N}$, the atom times of $\mathbf{X}^{*[n]}_\chi$ are a thinned version of the atom times of $\mathbf{M}$. In the construction of $\mathbf{X}^{*[n]}_\chi$, an atom $(t, M_t) \in \mathbf{M}$ results in a jump in $\mathbf{X}^{*[n]}_\chi$ if and only if $M_t^{[n]} \neq \mathrm{id}_{k,n}$ and $M_t^{[n]}(X^{*[n]}_{t-}) \neq X^{*[n]}_{t-}$. By the thinning property of Poisson processes, given $X^{*[n]}_{t-} = x \in [k]^{[n]}$, the total intensity at which $\mathbf{X}^{*[n]}_\chi$ jumps from state x to $x' \neq x$ is $\chi_n(\{M \in [k]^{[n]\otimes k} : M(x) = x'\})$. And by (4.4) and (4.5),

$$\begin{aligned}\chi_n\big(\{M \in [k]^{[n]\otimes k} : M(x) = x'\}\big) &= \sum_{M : M(x) = x'} Q_{nk}(\mathbf{Z}_{k,n}, M) \\ &= Q_{nk}\big(\mathbf{Z}_{k,n}, \{z \in [k]^{[nk]} : z^{\varphi_x} = x'\}\big) \\ &= Q_n\big(x, x'\big).\end{aligned}$$

It follows that the total intensity of jumps out of x is

$$\chi_n\big(\{M \in [k]^{[n]\otimes k} : M(x) \neq x\}\big) = Q_n\big(x, [k]^{[n]} \setminus \{x\}\big) < \infty,$$

and, for each $n \in \mathbb{N}$, $\mathbf{X}^{*[n]}_\chi$ is an exchangeable Markov process with jump rates $Q_n(\cdot, \cdot)$. Kolmogorov's extension theorem implies $\mathbf{X}^*_\chi$ is a version of $\mathbf{X}$. $\square$

4.1. *Lévy–Itô representation.* Our entire discussion climaxes in Theorem 1.3, the Lévy–Itô representation. For any exchangeable, consistent Markov process on $[k]^{\mathbb{N}}$, its characteristic measure χ has two unique components: a measure Σ on $k \times k$ stochastic matrices for which

$$\Sigma(\{I_k\}) = 0 \quad \text{and} \quad \int_{\mathcal{S}_k} (1 - S_*)\Sigma(dS) < \infty, \tag{4.8}$$

where $S_* := \min(S_{11}, \ldots, S_{kk})$, and a collection $\mathbf{c} = (\mathbf{c}_{ii'}, 1 \le i \ne i' \le k)$ of nonnegative constants.

For $1 \le i \ne i' \le k$ and $n \in \mathbb{N}$, we define $\rho_{ii'}^{(n)}$ as the point mass at $\kappa_{ii'}^{(n)} = (z_1, \ldots, z_k) \in [k]^{\mathbb{N} \otimes k}$, where

$$z_j^{j'} = \begin{cases} i', & j = i, j' = n, \\ j, & \text{otherwise.} \end{cases}$$

In words, $\rho_{ii'}^{(n)}$ charges only the map $\kappa_{ii'}^{(n)}$ that fixes all but the nth coordinate of every $x \in [k]^{\mathbb{N}}$: if $x^n = i$, then the nth coordinate of $\kappa_{ii'}^{(n)}(x)$ is i'; otherwise, the nth coordinate is also unchanged. We call each $\kappa_{ii'}^{(n)}$ a *single-index flip*. For example, with $k = 3$, $\rho_{12}^{(3)}$ puts unit mass at $\kappa_{12}^{(3)} = (1121\cdots, 2222\cdots, 3333\cdots)$. The measure

$$\rho_{ii'}(\cdot) := \sum_{n=1}^{\infty} \rho_{ii'}^{(n)}(\cdot), \qquad 1 \le i \ne i' \le k,$$

puts unit mass at every single-index flip from i to i'.

For any Σ satisfying (4.8) and any collection $(\mathbf{c}_{ii'}, 1 \le i \ne i' \le k)$ of nonnegative constants, we define

$$\chi_{\Sigma,\mathbf{c}} := \mu_\Sigma + \sum_{1 \le i \ne i' \le k} \mathbf{c}_{ii'} \rho_{ii'}, \tag{4.9}$$

where μ_Σ was defined in (3.5).

PROPOSITION 4.6. *Let Σ satisfy (4.8) and $\mathbf{c} = (\mathbf{c}_{ii'}, 1 \le i \ne i' \le k)$ be nonnegative constants. Then $\chi_{\Sigma,\mathbf{c}}$ defined in (4.9) is a coset exchangeable measure satisfying (4.1).*

PROOF. We treat each term of $\chi_{\Sigma,\mathbf{c}}$ separately.

Clearly, $\mu_\Sigma(\{\mathrm{id}_k\}) = 0$ by the first half of (4.8) and the strong law of large numbers. Now, for every $n \in \mathbb{N}$ and $S \in \mathcal{S}_k$, we have

$$\mu_S(\{M : M^{[n]} \ne \mathrm{id}_{k,n}\}) \le \sum_{j=1}^{k} \mu_S(\{M : M_j^{[n]} \ne \mathbf{j}^{[n]}\}) \le k(1 - S_*^n) \le nk(1 - S_*),$$

where $\mathbf{j} = jj \cdots \in [k]^{\mathbb{N}}$ and $\mathbf{j}^{[n]} := j \cdots j$ is its restriction to $[k]^{[n]}$. By (4.8),

$$\mu_\Sigma(\{M : M^{[n]} \neq \mathrm{id}_{k,n}\}) \leq nk \int_{\mathcal{S}_k} (1 - S_*)\Sigma(dS) < \infty.$$

The first half of (4.1) is satisfied by $\sum_{i \neq i'} \mathbf{c}_{ii'} \rho_{ii'}$ because each $\rho_{ii'}$ charges only single-index flips. Furthermore, with $c^* := \max_{1 \leq i \neq i' \leq k} \mathbf{c}_{ii'} < \infty$,

$$\begin{aligned}\sum_{1 \leq i \neq i' \leq k} \mathbf{c}_{ii'} \rho_{ii'}(\{M : M^{[n]} \neq \mathrm{id}_{k,n}\}) &\leq c^* \sum_{1 \leq i \neq i' \leq k} \sum_{j=1}^{n} \rho_{ii'}^{(j)}([k]^{\mathbb{N} \otimes k}) \\ &= nk(k-1)c^* < \infty.\end{aligned}$$

Thus, $\chi_{\Sigma,\mathbf{c}}$ satisfies (4.1).

Coset exchangeability of $\chi_{\Sigma,\mathbf{c}}$ follows since it is the sum of coset exchangeable measures. $\square$

Now, the denouement.

PROOF OF THEOREM 1.3. By Theorem 4.5, every exchangeable Feller process on $[k]^{\mathbb{N}}$ admits a version $\mathbf{X}^*_\chi$, for χ satisfying (4.1). In Theorem 1.3, we assert that χ can be decomposed as in (4.9). To prove this, we proceed in three steps:

(i) χ-almost every $M \in [k]^{\mathbb{N} \otimes k}$ possesses asymptotic frequency $|M|_k \in \mathcal{S}_k$,

(ii) there exists a unique measure Σ satisfying (4.8) such that the restriction of χ to $\{M \in [k]^{\mathbb{N} \otimes k} : |M|_k \neq I_k\}$ is a cut-and-paste measure,

$$\mathbf{1}_{\{|M|_k \neq I_k\}} \chi(dM) = \mu_\Sigma(dM) \quad \text{and}$$

(iii) there exist unique nonnegative constants $\mathbf{c} = (\mathbf{c}_{ii'}, 1 \leq i \neq i' \leq k)$ such that the restriction of χ to $\{M \in [k]^{\mathbb{N} \otimes k} : |M|_k = I_k\}$ is a single-index flip measure,

$$\mathbf{1}_{\{|M|_k = I_k\}} \chi(dM) = \sum_{1 \leq i \neq i' \leq k} \mathbf{c}_{ii'} \rho_{ii'}.$$

For (i), we let χ be the exchangeable characteristic measure of $\mathbf{X}$ from Theorem 4.5. Then χ satisfies (4.1) and we can write χ_n to denote the restriction of χ to the event $\{M \in [k]^{\mathbb{N} \otimes k} : M^{[n]} \neq \mathrm{id}_{k,n}\}$, for each $n \in \mathbb{N}$. By (4.1), each χ_n is a finite measure on $[k]^{\mathbb{N} \otimes k}$ and, by coset exchangeability, it is invariant under action by k-tuples of permutations $\sigma = (\sigma_1, \ldots, \sigma_k) : \mathbb{N}^k \to \mathbb{N}^k$ that fix $[n]^k$. As a result, we define the n-shift $\overleftarrow{M}_{[n]}$ of $M \in [k]^{\mathbb{N} \otimes k}$ as follows: for $M := (M_1, \ldots, M_k)$, we put $\overleftarrow{M}_{[n]} := (\overleftarrow{M}_{1,[n]}, \ldots, \overleftarrow{M}_{k,[n]})$, where

$$\overleftarrow{M}_{i,[n]} := M_i^{n+1} M_i^{n+2} \cdots, \qquad i = 1, \ldots, k.$$

(The n-shift of M is the coset decomposition of $M' = M^{nk+1} M^{nk+2} \cdots$, the k-coloring obtained by removing the first nk coordinates of M.) The image $\overleftarrow{\chi}_n$ of χ_n by the n-shift is a finite, coset exchangeable measure on $[k]^{\mathbb{N} \otimes k}$ that satisfies (4.1).

By corollary to Theorem 1.1, $\overleftarrow{\chi}_n$-almost every $M \in [k]^{\mathbb{N}\otimes k}$ possesses asymptotic frequency $|M|_k \in \mathcal{S}_k$. Since the asymptotic frequency of any $M \in [k]^{\mathbb{N}\otimes k}$ depends only on its n-shift, for every $n \in \mathbb{N}$, χ_n-almost every $M \in [k]^{\mathbb{N}\otimes k}$ possesses asymptotic frequency and, by Theorem 1.1, we may write

$$\chi_n(dM) = \int_{\mathcal{S}_k} \mu_S(dM)\chi_n\big(|M|_k \in dS\big). \tag{4.10}$$

Since $\chi_n \uparrow \chi$ as $n \uparrow \infty$, the monotone convergence theorem implies that χ-almost every $M \in [k]^{\mathbb{N}\otimes k}$ possesses asymptotic frequencies.

To establish (ii), we consider the event that $\{M \in [k]^{\mathbb{N}\otimes k} : \overleftarrow{M}^{[2]}_{[n]} \neq \mathrm{id}_{k,2}\}$ under χ_n. (Here, $\overleftarrow{M}^{[m]}_{[n]}$ denotes the restriction to $[k]^{[m]\otimes k}$ of the n-shift of $\overleftarrow{M}_{[n]}$.) We define the n-shift measure by

$$\overleftarrow{\chi}_n(dM) = \int_{\mathcal{S}_k} \mu_S(dM)\overleftarrow{\chi}_n\big(|M|_k \in dS\big), \tag{4.11}$$

from which, for every $S \in \mathcal{S}_k$,

$$\begin{aligned}
\chi_n\big(\{\overleftarrow{M}^{[2]}_{[n]} \neq \mathrm{id}_{k,2}\} \mid |M|_k = S\big) &= \overleftarrow{\chi}_n\big(M^{[2]} \neq \mathrm{id}_{k,2} \mid |M|_k = S\big) \\
&= \mu_S\big(\{M^{[2]} \neq \mathrm{id}_{k,2}\}\big) \\
&\geq 1 - S_*^2 \\
&\geq 1 - S_*.
\end{aligned}$$

Writing $\Sigma_n(dS) := \mathbf{1}_{\{|M|_k \neq I_k\}}|\chi_n|_k(dS)$, we obtain the inequality

$$\chi_n\big(\{\overleftarrow{M}^{[2]}_{[n]} \neq \mathrm{id}_{k,2}\}\big) \geq \int_{\mathcal{S}_k} (1 - S_*)\Sigma_n(dS). \tag{4.12}$$

By definition of χ_n and Σ_n, Σ_n increases to $\mathbf{1}_{\{|M|_k \neq I_k\}}|\chi|_k =: \Sigma$ as $n \to \infty$, the right-hand side above converges to

$$\int_{\mathcal{S}_k} (1 - S_*)\Sigma(dS),$$

and $\Sigma(\{I_k\}) = 0$. On the other hand, the left-hand side in (4.12) satisfies

$$\chi_n\big(\{\overleftarrow{M}^{[2]}_{[n]} \neq \mathrm{id}_{k,2}\}\big) \leq \chi\big(\{\overleftarrow{M}^{[2]}_{[n]} \neq \mathrm{id}_{k,2}\}\big) = \chi\big(\{M^{[2]} \neq \mathrm{id}_{k,2}\}\big) < \infty,$$

by coset exchangeability and (4.1). We conclude that

$$\int_{\mathcal{S}_k} (1 - S_*)\chi\big(|M|_k \in dS\big) = \int_{\mathcal{S}_k} (1 - S_*)\Sigma(dS) \leq \chi\big(\{\overleftarrow{M}^{[2]}_{[n]} \neq \mathrm{id}_{k,2}\}\big) < \infty;$$

and Σ satisfies (4.8).

Finally, we must establish $\mathbf{1}_{\{|M|_k \neq I_k\}}\chi = \mu_\Sigma$. Indeed, for every $n \in \mathbb{N}$ and fixed $M^* \neq \mathrm{id}_{k,n}$, the monotone convergence theorem implies

$$\chi\big(\{M^{[n]} = M^*, |M|_k \neq I_k\}\big) = \lim_{m\uparrow\infty} \chi\big(\{M^{[n]} = M^*, \overleftarrow{M}^{[m]}_{[n]} \neq \mathrm{id}_{k,m}, |M|_k \neq I_k\}\big).$$

By coset exchangeability, we can write

$$\chi\big(\{M^{[n]} = M^*, \overleftarrow{M}_{[n]}^{[m]} \neq \mathrm{id}_{k,m}, |M|_k \neq I_k\}\big) = \overleftarrow{\chi}_m\big(\{M^{[n]} = M^*, |M|_k \neq I_k\}\big),$$

and (4.11) implies

$$\overleftarrow{\chi}_m\big(\{M^{[n]} = M^*, |M|_k \neq I_k\}\big) = \int_{\mathcal{S}_k} \mu_S\big(\{M^{[n]} = M^*\}\big)\overleftarrow{\chi}_m\big(|M|_k \in dS\big),$$

which converges to

$$\int_{\mathcal{S}_k} \mu_S\big(\{M^{[n]} = M^*\}\big)\Sigma(dS) = \mu_\Sigma\big(\{M \in [k]^{\mathbb{N}\otimes k} : M^{[n]} = M^*\}\big).$$

As n was chosen arbitrarily and the restriction $|M|_k \neq I_k$ forbids $M = \mathrm{id}_k$, we conclude (ii).

To establish (iii), let χ^* be the restriction of χ to the event $\{M \in [k]^{\mathbb{N}\otimes k} : M^{[2]} \neq \mathrm{id}_{k,2}, |M|_k = I_k\}$. By (4.1) and corollary to Theorem 1.1, χ^* is finite and its image $\overleftarrow{\chi}_n^*$ by the n-shift is coset exchangeable; thus, $\overleftarrow{\chi}_n^*$-almost every $M \in [k]^{\mathbb{N}\otimes k}$ has asymptotic frequency $|M|_k = I_k$ and $\overleftarrow{\chi}_n^*$ is proportional to the unit mass at id_k. So, we may restrict our attention to the event $E := \{M^{[2]} \neq \mathrm{id}_{k,2}, \overleftarrow{M}_{[3]} = \mathrm{id}_k\}$ consisting of maps $[k]^{\mathbb{N}} \to [k]^{\mathbb{N}}$ that fix coordinates $n \geq 3$.

Any $M = (M_1, \ldots, M_k) \in E$ is specified by a k-tuple $((j_{11}, j_{12}), (j_{21}, j_{22}), \ldots, (j_{k1}, j_{k2}))$, that is, the ith coset of M [as in (2.4)] is

$$M_i = j_{i1} j_{i2} iii \cdots, \qquad i = 1, \ldots, k. \tag{4.13}$$

With $I = ((j_{11}, j_{12}), (j_{21}, j_{22}), \ldots, (j_{k1}, j_{k2}))$, we write $M_I \in [k]^{\mathbb{N}}$ to denote the map in (4.13). Let $K := \{((j_{11}, j_{12}), \ldots, (j_{k1}, j_{k2}))\}$ be the set of all k-tuples and $K^* := K \setminus \{I^*\}$, where $I^* \in K$ is defined as

$$I^* := \big((1,1), (2,2), \ldots, (k,k)\big).$$

Then $E := \bigcup_{I \in K^*} M_I$, which includes all single-index flip maps $\kappa_{ii'}^{(n)}$ for $n = 1, 2$.

Now, since $\overleftarrow{\chi}_n^*$ is proportional to the point mass at id_k, χ^* is the sum

$$\chi^*(\cdot) = \sum_{I \in K^*} c_I \delta_{M_I}(\cdot),$$

where $\delta_{M_I}(\cdot)$ is the Dirac point mass at M_I. By exchangeability, the requirement $\chi(\{M : M^{[2]} \neq \mathrm{id}_{k,2}\}) < \infty$ forces $c_I = 0$ unless M_I is a single-index flip map. By extension of the above argument, any $M \in [k]^{\mathbb{N}\otimes k}$ for which $|M|_k = I_k$ and $\mathbf{c}_M > 0$ must be a single-index flip map; otherwise, by exchangeability, each index changes states at an infinite rate and the finite restrictions cannot have càdlàg paths. This establishes (iii) and completes the proof. □

4.2. *Projection into the simplex.* By exchangeability of $\mathbf{X}$, the asymptotic frequency $|X_t|$ exists almost surely for any fixed $t \geq 0$. In discrete-time, this and countable additivity of probability measures imply the almost sure existence of $|\mathbf{X}| = (|X_m|, m \geq 0)$. In continuous-time, however, $\mathbf{X} = (X_t, t \geq 0)$ is uncountable and the corresponding conclusion does not follow immediately. Nevertheless, Theorem 1.3 harnesses the behavior of $\mathbf{X}$ to a fruitful outcome: $|\mathbf{X}| = (|X_t|, t \geq 0)$ exists and is a Feller process.

To show this, we work on the compact metric space $(\Delta_k, \tilde{d})$, where

$$\tilde{d}(s, s') := \frac{1}{2} \sum_{j=1}^{k} |s_j - s'_j|, \qquad s, s' \in \Delta_k.$$

Under this metric, any $S \in \mathcal{S}_k$ determines a Lipschitz continuous map $\Delta_k \to \Delta_k$, that is, for all $D, D' \in \Delta_k$ and any $S \in \mathcal{S}_k$,

$$\tilde{d}(DS, D'S) \leq \tilde{d}(D, D').$$

We further exploit an alternative description of $\mathbf{X}^*_{\Sigma,\mathbf{c}}$ by an associated Markov process on $[k]^{\mathbb{N}\otimes k}$.

Let $\mathbf{M}$ be the Poisson point process with intensity $dt \otimes \chi_{\Sigma,\mathbf{c}}$, as above. For each $n \in \mathbb{N}$, we define $\mathbf{F}^{[n]} := (F^{[n]}_t, t \geq 0)$ on $[k]^{[n]\otimes k}$ by $F^{[n]}_0 = \mathrm{id}_{k,n}$ and:

- if $t > 0$ is an atom time of $\mathbf{M}$ for which $M^{[n]}_t \neq \mathrm{id}_{k,n}$, we put $F^{[n]}_t = M^{[n]}_t(F^{[n]}_{t-})$,
- otherwise, we put $F^{[n]}_t = F^{[n]}_{t-}$.

We define $\mathbf{F}$ as the limit of $(\mathbf{F}^{[n]}, n \in \mathbb{N})$, which is a coset exchangeable, consistent Markov process on $[k]^{\mathbb{N}\otimes k}$. By its construction, $\mathbf{F}$ is closely tied to $\mathbf{X}^*_{\Sigma,\mathbf{c}} = (X^*_t, t \geq 0)$ by the relations:

- $|F_0|_k = I_k$ and
- $X^*_t = F_t(X^*_0)$ for all $t \geq 0$.

PROOF OF THEOREM 1.4. Let $(\mathcal{F}_t, t \geq 0)$ denote the natural filtration of $\mathbf{X}$ and, independently of $(\mathcal{F}_t, t \geq 0)$, let $\mathbf{F} := (F_t, t \geq 0)$ be the process on $[k]^{\mathbb{N}\otimes k}$ constructed above. By Theorem 1.3, the conditional law of X_{t+s} given $\mathcal{F}_t$ is that of $F_s(X_t)$. By (4.1) and exchangeability of X_0, X_t possesses asymptotic frequencies almost surely for every $t \geq 0$. In fact, $|X_t|$ exists simultaneously for all $t \geq 0$ with probability one.

From Theorem 1.3, a version of $\mathbf{X}$ can be constructed as $\mathbf{X}^*_{\Sigma,\mathbf{c}} = (X^*_t, t \geq 0)$, whose discontinuities are of Types-(I) and (II) in Section 1.2. In the projection $|\mathbf{X}^*_{\Sigma,\mathbf{c}}|$, discontinuities only occur at the times of Type-(I) discontinuities, of which there are at most countably many. In between jumps, the trajectory of $|\mathbf{X}^*_{\Sigma,\mathbf{c}}|$ is deterministic and continuous in Δ_k. As a result, $|\mathbf{X}^*_{\Sigma,\mathbf{c}}|$ exists and is càdlàg almost surely. By corollary to Theorem 1.2, $|X^*_{t+s}| =_{\mathcal{L}} |F_s(X^*_t)| = |F_s|_k |X^*_t|$, given $\mathcal{F}_t$.

Since permutation does not affect the asymptotic frequency of either F_s or X_t^*, $|\mathbf{X}^*_{\Sigma,\mathbf{c}}|$ has the Markov property.

Lipschitz continuity of every $S : \Delta_k \to \Delta_k$, $S \in \mathcal{S}_k$, implies the Feller property. By compactness of Δ_k, any continuous $g : \Delta_k \to \mathbb{R}$ is uniformly continuous and, therefore, bounded. By the dominated convergence theorem, continuity of the map defined by $S \in \mathcal{S}_k$, and Theorem 1.3, the maps $D \mapsto \mathbf{P}_t g(D)$ are continuous for all $t > 0$. By (4.1), $F_t \to \mathrm{id}_k$ in probability as $t \downarrow 0$; whence, $|F_t|_k \to I_k$ and $|F_t(X_0^*)| = |F_t|_k |X_0^*| \to |X_0^*|$, both in probability as $t \downarrow 0$. We conclude that $\lim_{t\downarrow 0} \mathbf{P}_t g(D) = g(D)$ for every continuous function $g : \Delta_k \to \mathbb{R}$, from which follows the Feller property. $\square$

5. Homogeneous cut-and-paste processes. Theorems 1.1–1.4 extend to partition-valued processes with minor modifications. Let $\mathbf{\Pi} = (\Pi_t, t \geq 0)$ be a continuous-time exchangeable, consistent Markov process on $\mathcal{P}_{\mathbb{N}:k}$. Specifically, $\mathbf{\Pi}$ is a Markov process such that

(A) $\mathbf{\Pi}^\sigma = (\Pi_t^\sigma, t \geq 0)$ is a version of $\mathbf{\Pi}$ for all $\sigma \in \mathscr{S}_{\mathbb{N}}$ and

(B) $\mathbf{\Pi}^{[n]} = (\Pi_t^{[n]}, t \geq 0)$ is a Markov chain on $\mathcal{P}_{[n]:k}$, for every $n = 1, 2, \ldots.$

By Proposition 2.2, $\mathbf{\Pi}$ is a Feller process, and thus, its evolution is determined by the finite jump rates

$$Q_n(\pi, \pi') := \lim_{t\downarrow 0} \frac{1}{t} \mathbb{P}\{\Pi_t^{[n]} = \pi' \mid \Pi_0^{[n]} = \pi\}, \qquad \pi \neq \pi' \in \mathcal{P}_{[n]:k}, \text{ for each } n \in \mathbb{N}, \tag{5.1}$$

which satisfy (4.3), (4.4) and (4.5).

For any $\pi \in \mathcal{P}_{\mathbb{N}:k}$, we obtain its *symmetric associate* $\tilde{x} \in [k]^{\mathbb{N}}$ by labeling the blocks of π uniformly and without replacement in $[k]$. In particular, for $\pi = (B_1, \ldots, B_r) \in \mathcal{P}_{\mathbb{N}:k}$ (listed in order of least element), $\tilde{x}$ is a random k-coloring of $\mathbb{N}$ obtained by drawing labels $(l_1, \ldots, l_r)$ without replacement from $[k]$ and putting $\tilde{x} = \tilde{x}^1\tilde{x}^2\cdots$, where

$$\tilde{x}^j = l_i \quad \Longleftrightarrow \quad j \in B_i.$$

Thus, $\mathcal{B}(\tilde{x}) = \pi$ with probability one and each element in the set $\mathcal{B}^{-1}(\pi)$ has equal probability. For each $n \in \mathbb{N}$, we define the *symmetric associate transition rate* $\tilde{Q}_n$ on $[k]^{[n]}$ by

$$\tilde{Q}_n(x, x') := Q_n(\mathcal{B}_n(x), \mathcal{B}_n(x'))/k^{\downarrow \#\mathcal{B}_n(x')}, \qquad x \neq x' \in [k]^{[n]}, \tag{5.2}$$

where $\#\pi$ denotes the number of blocks of $\pi \in \mathcal{P}_{\mathbb{N}}$ and $k^{\downarrow j} := k(k-1)\cdots(k-j+1)$. Under $\tilde{Q}_n$, a transition from $x \in [k]^{[n]}$ is obtained by projecting $x \mapsto \mathcal{B}_n(x) = \pi$, generating a transition $\Pi' \sim Q_n(\pi, \cdot)$, and randomly coloring the blocks of Π' to obtain a symmetric associate $\tilde{X}' \in [k]^{[n]}$. The next proposition follows from definition (5.2) and properties (4.3), (4.4) and (4.5) of $(Q_n, n \in \mathbb{N})$ in (5.1).

PROPOSITION 5.1. *The collection* $(\tilde{Q}_n, n \in \mathbb{N})$ *defined in* (5.2) *determines a unique exchangeable transition rate measure* $\tilde{Q}$ *on* $[k]^{\mathbb{N}}$.

From $\tilde{Q}$, we construct $\tilde{\mathbf{X}} = (\tilde{X}_t, t \geq 0)$, the *symmetric associate* of $\mathbf{\Pi}$, by first generating $\tilde{X}_0$ as the symmetric associate of a partition from the initial distribution of $\mathbf{\Pi}$ and, given $\tilde{X}_0$, letting $\tilde{\mathbf{X}}$ evolve as a Markov process with initial state $\tilde{X}_0$ and transition rate measure $\tilde{Q}$.

PROPOSITION 5.2. *The symmetric associate* $\tilde{\mathbf{X}}$ *of* $\mathbf{\Pi}$ *is an exchangeable, consistent Markov process on* $[k]^{\mathbb{N}}$ *and* $\mathcal{B}(\tilde{\mathbf{X}}) = (\mathcal{B}(\tilde{X}_t), t \geq 0)$ *is a version of* $\mathbf{\Pi}$.

PROOF. We have constructed $\tilde{\mathbf{X}}$ so that it projects to and respects the structure of $\mathbf{\Pi}$. To wit, $\mathbf{\Pi}$ is exchangeable and consistent, and so is $\tilde{\mathbf{X}}$. $\square$

For any permutation $\gamma : [k] \to [k]$, we define the *recoloring* of $x \in [k]^{\mathbb{N}}$ by

$$\gamma x := \gamma(x^1)\gamma(x^2)\cdots. \tag{5.3}$$

Since $\mathcal{B}(x)$ is the projection of x into $\mathcal{P}_{\mathbb{N}:k}$ by removing colors, recoloring does not affect $x \mapsto \mathcal{B}(x)$, that is, $\mathcal{B}(x) = \mathcal{B}(\gamma x)$ for all $x \in [k]^{\mathbb{N}}$ and $\gamma \in \mathscr{S}_k$. Thus, by definition (5.2), $\tilde{Q}$ is invariant under arbitrary recoloring of its arguments,

$$\tilde{Q}(\gamma x, \gamma' A) = \tilde{Q}(x, A), \qquad x \in [k]^{\mathbb{N}}, A \subseteq [k]^{\mathbb{N}}, \tag{5.4}$$

for all $\gamma, \gamma' \in \mathscr{S}_k$, where $\gamma' A := \{\gamma' x' : x' \in A\}$ is the image of A under recoloring by γ'. By Theorem 4.5, $\tilde{Q}$ is characterized by a coset exchangeable measure $\tilde{\chi}$ which, by condition (5.4), is invariant under the action of *left- and right-recoloring*, which we now define.

For $M \in [k]^{\mathbb{N} \otimes k}$ and $\gamma, \gamma' \in \mathscr{S}_k$, we define the *left–right recoloring of* M *by* (γ, γ') by $M' := \gamma M \gamma'$, where

$$M'(x) := \gamma' M(\gamma^{-1} x), \qquad x \in [k]^{\mathbb{N}}, \tag{5.5}$$

the k-coloring obtained by first recoloring x by γ^{-1}, then applying M, and finally recoloring by γ'. We call a coset exchangeable measure *row–column exchangeable* if it is invariant under left–right recoloring by all pairs $(\gamma, \gamma') \in \mathscr{S}_k \times \mathscr{S}_k$.

LEMMA 5.3. *Let* $\tilde{\chi}$ *be the coset exchangeable measure that determines* $\tilde{Q}$. *Then* $\tilde{\chi}$ *is row–column exchangeable.*

PROOF. Fix $x \in [k]^{\mathbb{N}}$ and $A \subseteq [k]^{\mathbb{N}}$. By (5.4) and Theorem 1.3,

$$\begin{aligned}
\tilde{\chi}(\{M : M(x) \in A\}) &= \tilde{Q}(x, A) \\
&= \tilde{Q}(\gamma x, \gamma' A) \\
&= \tilde{\chi}(\{M : M(\gamma x) \in \gamma' A\}) \\
&= \tilde{\chi}(\{M : \gamma^{-1} M \gamma'^{-1}(x) \in A\}) \\
&= \tilde{\chi}(\{\gamma M \gamma' : M(x) \in A\}),
\end{aligned}$$

implying $\tilde{\chi}$ is row–column exchangeable. $\square$

As a corollary to Theorem 1.3 and Proposition 5.2, $\tilde{\chi}$ is determined by a unique pair $(\tilde{\Sigma}, \tilde{\mathbf{c}})$, where $\tilde{\Sigma}$ is a measure satisfying (1.6) and $\tilde{\mathbf{c}} = (\tilde{\mathbf{c}}_{ii'}, 1 \le i \ne i' \le k)$ is a collection of nonnegative constants, that is,

$$\tilde{\chi} = \mu_{\tilde{\Sigma}} + \sum_{1\le i\ne i'\le k} \tilde{\mathbf{c}}_{ii'}\rho_{ii'}. \tag{5.6}$$

On $\mathcal{S}_k$, we call a measure Σ *row–column exchangeable* if it is invariant under arbitrary permutation of rows and columns, $S \mapsto \gamma S \gamma'^{-1} := (S_{\gamma(i)\gamma'(i')}, 1 \le i, i' \le k)$ for all $\gamma, \gamma' \in \mathscr{S}_k$.

PROPOSITION 5.4. *Let $\tilde{\chi}$ be as defined in* (5.6). *Then $\tilde{\Sigma}$ is row–column exchangeable and there exists a unique $c \ge 0$ such that $\tilde{\mathbf{c}}_{ii'} = c$ for all $1 \le i \ne i' \le k$.*

PROOF. In (5.6), $\tilde{\chi}$ is expressed as the sum of mutually singular measures, and we treat $\sum_{1\le i\ne i'\le k} \tilde{\mathbf{c}}_{ii'}\rho_{ii'}$ first.

For $1 \le i \ne i' \le k$ and $n \in \mathbb{N}$, we define

$$A_{ii'}(n) := \{\kappa_{ii'}^{(1)}, \dots, \kappa_{ii'}^{(n)}\},$$

the subset of $[k]^{\mathbb{N}\otimes k}$ containing all single-index flips from i to i' for indices in $[n]$. By Lemma 5.3, $\tilde{\chi}$ is invariant under arbitrary left- and right-recoloring as in (5.5); whence,

$$n\tilde{\mathbf{c}}_{ii'} = \tilde{\chi}(A_{ii'}(n)) = \tilde{\chi}(A_{\gamma(i)\gamma(i')}(n)) = n\tilde{\mathbf{c}}_{\gamma(i)\gamma(i')}$$

for all $n \in \mathbb{N}$ and $\gamma \in \mathscr{S}_k$, implying $\tilde{\mathbf{c}}_{ii'} = \tilde{\mathbf{c}}_{jj'} = c$ for all $i \ne i'$ and $j \ne j'$.

Restricted to the event $\{M \in [k]^{\mathbb{N}\otimes k} : |M|_k \ne I_k\}$, $\tilde{\chi}$ induces a measure $\tilde{\Sigma}$ satisfying (1.6) through the map $M \mapsto |M|_k$. Row–column exchangeability follows by row–column exchangeability of $\tilde{\chi}$ and definition of $M \mapsto |M|_k$ in (2.10). $\square$

PROOF OF THEOREM 1.5. For $\mathbf{\Pi}$ in continuous-time, Theorem 1.5 is a corollary of Theorem 1.3 and Propositions 5.1, 5.2 and 5.4. The discrete-time conclusion follows since single-index flips are not permitted (forcing $c = 0$) and Markov processes with finite jump rates can be treated as discrete-time chains with exponentially distributed hold times between jumps. $\square$

According to Theorem 1.4, the projection into Δ_k of an exchangeable $[k]^{\mathbb{N}}$-valued Feller process exists and is also a Feller process. The analogous projection of $\mathbf{\Pi}$ into $\Delta_k^{\downarrow}$ by $|\cdot|^{\downarrow}$ also exists and is Feller.

PROOF OF THEOREM 1.6. Almost sure existence of $|\mathbf{\Pi}|^{\downarrow}$ follows from Theorem 1.5 and the existence of $|\mathbf{X}|$ for any exchangeable Feller process on $[k]^{\mathbb{N}}$ (Theorem 1.4). By Proposition 5.4, the characteristic measure χ induces a row–column exchangeable measure $|\chi|_k$ on $\mathcal{S}_k$, and so $|\mathbf{\Pi}|^{\downarrow}$ is Markovian. Theorem 1.4 implies the Feller property since $|\mathbf{X}|$ is Feller and any continuous $g : \Delta_k^{\downarrow} \to \mathbb{R}$ induces a continuous function $g' : \Delta_k \to \mathbb{R}$ which is symmetric in its arguments. $\square$

By the description in Theorem 1.5, $\mathbf{\Pi}$ is characterized by its symmetric associate $\tilde{\mathbf{X}}$, whose transition law treats colors homogeneously. We commingle terms and call both $\tilde{\mathbf{X}}$ and $\mathbf{\Pi}$ a *homogeneous cut-and-paste process* with parameter $(\tilde{\Sigma}, \tilde{c})$.

5.1. *Self-similar cut-and-paste processes.* In [6], we introduced a family of cut-and-paste chains, which we now call self-similar homogeneous cut-and-paste chains. We showed an instance of these chains in Example 1.7.

For a self-similar cut-and-paste process, the measure Σ is the k-fold product of some σ-finite measure on Δ_k, that is, $\Sigma = \nu \otimes \cdots \otimes \nu$, for ν symmetric and satisfying

$$\nu\bigl(\{(1, 0, \dots, 0)\}\bigr) = 0 \quad \text{and} \quad \int_{\Delta_k^{\downarrow}} (1 - s_*)\nu(ds) < \infty, \tag{5.7}$$

where $s_* := \min\{s_1, \dots, s_k\}$. By symmetry of ν, Σ is row–column exchangeable.

The processes studied in [6] were *pure-jump* in that they did not admit single-index flips. By letting single-index flips occur at rate $c \geq 0$, we obtain the class of *self-similar homogeneous cut-and-paste processes* with characteristic measure

$$\chi = \mu_{\nu \otimes \cdots \otimes \nu} + c\rho,$$

where $\rho := \sum_{1 \leq i \neq i' \leq k} \rho_{ii'}$. The special case $c = 0$ and $\nu = \mathrm{PD}(-\alpha/k, \alpha)$ plays a role in clustering applications [4].

6. Concluding remarks.

6.1. *Relation to exchangeable coalescent and fragmentation processes.* In spirit, our main theorems resemble previous results for exchangeable coalescent and fragmentation processes. In substance, our processes differ in fundamental ways.

6.1.1. *Bounded number of blocks.* All processes studied in this paper evolve on either $[k]^{\mathbb{N}}$ or $\mathcal{P}_{\mathbb{N}:k}$ for fixed $k \in \mathbb{N}$. Bounding the number of blocks is necessary to characterize the jump probabilities/rates by a measure on stochastic matrices. Without an upper bound on the number of blocks, an exchangeable partition need not admit proper asymptotic frequencies. In general, for $\pi = \{B_1, B_2, \dots\} \in \mathcal{P}_{\mathbb{N}}$, the sum of its asymptotic block frequencies may be strictly less than one, in which case, it is common to write $s_0 := 1 - \sum_i |B_i|$ to denote the amount of *dust* in $|\pi|^{\downarrow}$. For an exchangeable partition of $\mathbb{N}$, the dust is the totality of its singleton blocks. Furthermore, Theorem 1.5 requires the cut-and-paste measure Σ to treat all blocks symmetrically. Without a uniform distribution on a countable set, we cannot specify such a measure on $[k]^{\mathbb{N} \otimes k}$ with k unbounded.

6.1.2. *Coalescent processes with finite initial state*. The representation in (5.6) covers a special subclass of exchangeable coalescent processes whose initial state has a finite number of blocks. In this case, we let k be the number of blocks of the initial state Π_0, $c = 0$, and Σ a σ-finite row–column exchangeable measure concentrated on $\{0, 1\}$-valued stochastic matrices. In this case, the homogeneous cut-and-paste process with initial state Π_0 and characteristic measure $\chi = \mu_\Sigma$ is an exchangeable coalescent.

On the other hand, no class of fragmentation processes corresponds to a cut-and-paste process. Fragmentation processes eventually fragment into the state of all singletons, for which the number of blocks is infinite.

6.1.3. *Poissonian structure*, *coset mappings and* COAG–FRAG *operators*. Exchangeable coalescent and fragmentation processes admit Poisson point process constructions akin to our construction of $\mathbf{X}$ from the Poisson point process $\mathbf{M}$ on $\mathbb{R}_+ \times [k]^{\mathbb{N}\otimes k}$. For a coalescent process, $\mathbf{B} = \{(t, B_t)\}$ is a random subset of $\mathbb{R}_+ \times \mathcal{P}_\mathbb{N}$ and $\mathbf{\Pi} = (\Pi_t, t \geq 0)$ is constructed (informally) by putting $\Pi_t = \mathrm{COAG}(\Pi_{t-}, B_t)$, for each atom time t. For $\pi, \pi' \in \mathcal{P}_\mathbb{N}$, $\mathrm{COAG}(\pi, \pi')$ is the *coagulation of* π *by* π', which determines a Lipschitz continuous mapping $\mathcal{P}_\mathbb{N} \to \mathcal{P}_\mathbb{N}$. Fragmentation processes have a similar construction in terms of the FRAG-operator, which is also Lipschitz continuous.

The coset mappings, essential to our construction of cut-and-paste processes, are also Lipschitz continuous. To mimic the above constructions by the COAG and FRAG operators, we can define an operation $\mathrm{CUT\text{-}PASTE} : [k]^{\mathbb{N}\otimes k} \times \mathcal{P}_{\mathbb{N}:k} \to \mathcal{P}_{\mathbb{N}:k}$ by

$$\mathrm{CUT\text{-}PASTE}(M, \pi) := \mathcal{B}\big(M(\tilde{x})\big), \qquad \tilde{x} \text{ the symmetric associate of } \pi.$$

From a Poisson point process $\mathbf{M}$ with intensity $dt \otimes \tilde{\chi}$, we generate $\mathbf{\Pi} = (\Pi_t, t \geq 0)$ (informally) by putting $\Pi_t = \mathrm{CUT\text{-}PASTE}(M_t, \Pi_{t-})$, for each atom time of $\mathbf{M}$. The CUT-PASTE operator differs from COAG and FRAG because it maps $[k]^{\mathbb{N}\otimes k} \times \mathcal{P}_{\mathbb{N}:k} \to \mathcal{P}_{\mathbb{N}:k}$, rather than $\mathcal{P}_\mathbb{N} \times \mathcal{P}_\mathbb{N} \to \mathcal{P}_\mathbb{N}$.

We spare the details. See [5] for more on the interplay between Poissonian structure, the Feller property and Lipschitz continuous mappings.

6.2. *Equilibrium measures of cut-and-paste processes*. The process in Example 1.7 is a self-similar homogeneous cut-and-paste chain which is also reversible with respect to the Poisson–Dirichlet distribution. The process in Example 1.8 evolves in continuous-time and converges to a distribution whose projection to the simplex is degenerate at $(1/2, 1/2)$. By Kingman's paintbox correspondence, these are the only possibilities. In particular, the unique equilibrium measure of an exchangeable cut-and-paste process, if it exists, is one of Kingman's paintbox measures. The cut-and-paste representation is a powerful tool for studying equilibrium measures of these chains, evinced by Crane and Lalley [7].

Acknowledgements. I thank Steve Lalley for his helpful comments on an early version and Chris Burdzy for his kind encouragement.

REFERENCES

[1] ALDOUS, D. J. (1985). Exchangeability and related topics. In *École D'été de Probabilités de Saint-Flour, XIII—1983*. *Lecture Notes in Math*. **1117** 1–198. Springer, Berlin. MR0883646

[2] BERTOIN, J. (2001). Homogeneous fragmentation processes. *Probab. Theory Related Fields* **121** 301–318. MR1867425

[3] BERTOIN, J. (2006). *Random Fragmentation and Coagulation Processes*. *Cambridge Studies in Advanced Mathematics* **102**. Cambridge Univ. Press, Cambridge. MR2253162

[4] CRANE, H. (2014). Clustering from partition data. Unpublished manuscript.

[5] CRANE, H. (2014). Lipschitz partition processes. *Bernoulli*. To appear.

[6] CRANE, H. (2011). A consistent Markov partition process generated from the paintbox process. *J. Appl. Probab.* **48** 778–791. MR2884815

[7] CRANE, H. and LALLEY, S. P. (2013). Convergence rates of Markov chains on spaces of partitions. *Electron. J. Probab.* **18** 23. MR3078020

[8] DURRETT, R., GRANOVSKY, B. L. and GUERON, S. (1999). The equilibrium behavior of reversible coagulation-fragmentation processes. *J. Theoret. Probab.* **12** 447–474. MR1684753

[9] EWENS, W. J. (1972). The sampling theory of selectively neutral alleles. *Theoret. Population Biology* **3** 87–112; erratum, ibid. **3** (1972), 240; erratum, ibid. **3** (1972), 376. MR0325177

[10] FURSTENBERG, H. and KESTEN, H. (1960). Products of random matrices. *Ann. Math. Statist.* **31** 457–469. MR0121828

[11] KINGMAN, J. F. C. (1978). Random partitions in population genetics. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **361** 1–20. MR0526801

[12] KINGMAN, J. F. C. (1978). The representation of partition structures. *J. Lond. Math. Soc.* (2) **18** 374–380. MR0509954

[13] KINGMAN, J. F. C. (1982). The coalescent. *Stochastic Process. Appl.* **13** 235–248. MR0671034

[14] PITMAN, J. (1999). Coalescents with multiple collisions. *Ann. Probab.* **27** 1870–1902. MR1742892

[15] PITMAN, J. (2002). Poisson–Dirichlet and GEM invariant distributions for split-and-merge transformation of an interval partition. *Combin. Probab. Comput.* **11** 501–514. MR1930355

[16] PITMAN, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Math*. **1875**. Springer, Berlin. MR2245368

DEPARTMENT OF STATISTICS
RUTGERS UNIVERSITY
110 FRELINGHUYSEN ROAD
PISCATAWAY, NEW JERSEY 08854
USA
E-MAIL: hcrane@stat.rutgers.edu