The Annals of Probability
2004, Vol. 32, No. 1B, 915–938

THE POISSON–DIRICHLET LAW IS THE UNIQUE INVARIANT DISTRIBUTION FOR UNIFORM SPLIT–MERGE TRANSFORMATIONS

BY PERSI DIACONIS, EDDY MAYER-WOLF,[1] OFER ZEITOUNI[2] AND MARTIN P. W. ZERNER[3]

Stanford University, Technion, Technion and University of Minnesota, and Stanford University

Dedicated to the memory of Bob Brooks (1952–2002)

We consider a Markov chain on the space of (countable) partitions of the interval $[0, 1]$, obtained first by size-biased sampling twice (allowing repetitions) and then merging the parts (if the sampled parts are distinct) or splitting the part uniformly (if the same part was sampled twice). We prove a conjecture of Vershik stating that the Poisson–Dirichlet law with parameter $\theta = 1$ is the unique invariant distribution for this Markov chain. Our proof uses a combination of probabilistic, combinatoric and representation-theoretic arguments.

1. Introduction. Let Ω_1 denote the space of (ordered) partitions of $[0, 1]$, that is,

$$\Omega_1 := \{p \in [0,1]^{\mathbb{N}} : p_1 \geq p_2 \geq \cdots \geq 0, |p|_1 = 1\},$$

where $|x|_1 = \sum_i |x_i|$ for any finite or countable sequence (x_i). By *size-biased* sampling according to a point $p \in \Omega_1$ we mean picking the jth part p_j with probability p_j. Our interest in this paper is in the following Markov chain on Ω_1, which we call a *continuous coagulation–fragmentation* process (CCF): size-bias sample (with replacement) two parts from p; if the same part was picked twice, split it (uniformly) and reorder the partition; if different parts were picked, merge them and reorder the partition.

We denote by $\mathrm{DCF}^{(n)}$ (*discrete coagulation–fragmentation*) the Markov chain describing the evolution of the cycle lengths of permutations of $\{1, \dots, n\}$ under random transpositions. The CCF process appears in a variety of contexts, but of particular relevance to us is its occurrence as a natural limit of $\mathrm{DCF}^{(n)}$, when

Received July 2002; revised January 2003.

[1]Supported in part by the S. Faust research fund.

[2]Supported in part by the fund for promotion of research at the Technion, and by a US–Israel BSF grant.

[3]Supported in part by an Aly Kaufman Fellowship at the Technion.

AMS 2000 subject classifications. Primary 60K35; secondary 60J27, 60G55.

Key words and phrases. Partitions, coagulation, fragmentation, invariant measures, Poisson–Dirichlet.

n increases; see [16] for a discussion of this and its link with the space of "virtual permutations."

For any $n \in \mathbb{N}$ denote

$$\mathcal{P}_n := \{\ell = (\ell_i)_{i\geq 1} \in \{0, 1, \ldots, n\}^{\mathbb{N}} : \ell_1 \geq \ell_2 \geq \cdots \geq 0, |\ell|_1 = n\} \subset n\Omega_1.$$

(Elements in $\mathcal{P}_n$ may be thought of as being of length n; the remaining entries are necessarily zero.)

A sequence $\ell \in \mathcal{P}_n$ is uniquely determined by its *type* $(N_\ell(k) = \sharp\{i : \ell_i = k\})_{k=1}^n$, with $N_\ell = \sum_{k=1}^n N_\ell(k)$ denoting ℓ's total number of parts.

The long-time behavior of the $\mathrm{DCF}^{(n)}$, viewed as an evolution in $\mathcal{P}_n$, is well understood. In particular (see, e.g., [4]), it possesses a unique stationary distribution given by the Ewens formula:

$$\pi_S^{(n)}(\ell) = \left(\prod_{k=1}^{n} k^{N_\ell(k)} N_\ell(k)!\right)^{-1} = \left(\prod_{i=1}^{n} \ell_i \prod_{k=1}^{n} N_\ell(k)!\right)^{-1}, \qquad \ell \in \mathcal{P}_n. \tag{1.1}$$

It is well known, at least since [11, 12, 18], that the measures $\pi_S^{(n)}(n\cdot)$ on Ω_1 converge weakly to the Poisson–Dirichlet distribution $\widehat{\mu}_1$ with parameter $\theta = 1$ (a precise definition of $\widehat{\mu}_1$ is given in Section 2.1). It has been shown in more than one way (cf. [8, 15, 16]) that $\widehat{\mu}_1$ is invariant for the CCF transition. This fact, and hints coming from the theory of virtual permutations, led Vershik (see [16]) to the following conjecture.

CONJECTURE 1.1 (Vershik). *$\widehat{\mu}_1$ is the unique invariant distribution for the CCF.*

Our goal in this article is to prove Vershik's conjecture. A naive approach toward the proof would be to use the link with the $\mathrm{DCF}^{(n)}$ and the fact that the latter converges to the distribution $\pi_S^{(n)}$ exponentially fast. However, the rate of that convergence deteriorates with n. To overcome this difficulty, our strategy consists of the following steps:

1. We provide a priori estimates (Proposition 2.1) showing that every invariant distribution for the CCF leads to a good control on the number of "small parts."
2. We couple the $\mathrm{DCF}^{(n)}$ and the CCF in such a way that whenever they start from initial distributions with such control on the tails, the decoupling time is roughly $\sqrt{n}$ (Theorem 3.1).
3. For initial conditions as above, and for an appropriate class of test functions, we show by using some harmonic analysis on the symmetric group that the $\mathrm{DCF}^{(n)}$ achieves near equilibrium before the decoupling time (Theorem 4.1).

These steps are then combined in Theorem 5.1 to yield the proof of Vershik's conjecture.

Our work began from discussions with Bob Brooks on various models for "random Riemann surfaces." Brooks and Makover [2, 3] studied Riemann surfaces via a dense set of "Belyi surfaces" associated with three-regular graphs on n vertices with an orientation at each vertex. Their construction gives a complete Riemann surface with finite area πn for each graph. Uniformly choosing a random three-regular graph gives a probability distribution on Riemann surfaces; see [7] for an accessible account of this model. Practical choice of a random three-regular graph is not so easy when n is large. Brooks proposed a Markov chain method which involved splitting and joining cycles; investigating properties of his algorithm gave rise to the present paper.

We next review some of the literature on this question. Tsilevich [16] proves that $\widehat{\mu}_1$ is the only CCF-invariant measure that is also invariant under additional symmetry conditions. Pitman [15] proves that $\widehat{\mu}_1$ is the only CCF-invariant measure which is also invariant under size-biased sampling. Related results appear in [9]. In another direction, it is shown in [14] that $\widehat{\mu}_1$ is the only CCF-invariant measure that is analytic in the sense that, for any k, the law of an independently size-biased sample (with replacement) possesses an analytic density. Finally, Tsilevich [17] shows that the law of the CCF, initialized at $p = (1, 0, \dots)$ and stopped at a Binomial(n, $1/2$) random time, converges to $\widehat{\mu}_1$.

We conclude this Introduction by noting that, in [14], we have introduced a slightly more general model of split–merge transformations, by allowing either the split or the merge operation to be rejected with a certain probability. An invariant measure for these generalizations is the Poisson–Dirichlet law of parameter $\theta > 0$. The discrete counterpart of this chain has been analyzed in [5], Section 4. While it is plausible that the techniques of the current paper can be adapted to that setup using the results of [5], we do not pursue this generalization here.

2. Continuous and discrete coagulation–fragmentation.

2.1. *Preliminaries and CCF.* Given a topological space W, its Borel σ-algebra will be denoted by $\mathcal{B}_W$, and the space of probability measures on $(W, \mathcal{B}_W)$ by $\mathcal{M}_1(W)$. By a slight abuse of notation, $\mathcal{M}_1(V)$ will also be $\mathcal{M}_1(W)$'s subspace of probability measures whose support is contained in a given closed subset V of W. The total variation of a measure ν is denoted by $\|\nu\|_{\mathrm{var}}$.

We equip Ω_1 with its relative $|\cdot|_1$-topology which, on Ω_1, coincides with the weak (coordinatewise convergence) topology.

On Ω_1 we consider the Markov chain CCF in which two segments p_i and p_j of a given partition p are size-bias sampled with replacement and then, if $i \neq j$, they merge into one of length $p_i + p_j$ (coagulation), while if $i = j$, p_i splits into two new parts up_i, $(1-u)p_i$ with $u \sim U[0, 1]$ independent of all the rest (fragmentation). In either case the new partition is then rearranged nonincreasingly.

Recall that the Poisson–Dirichlet law $\widehat{\mu}_1$ is invariant for the CCF transition. Indeed, $\widehat{\mu}_1$ itself has been defined in a variety of manners [1, 11] which are well known to be equivalent. Perhaps the simplest is the GEM description in which segments are successively and uniformly removed from whatever remains of $[0,1]$ and then rearranged nonincreasingly. Namely, let $Y_1 = 1$ and for $n \in \mathbb{N}$ define $X_n = U_n Y_n$ (the removed part at stage n) and $Y_{n+1} = Y_n - X_n$ [the remaining segment from which the $(n+1)$st part is to be removed], where the U_n's are independent $U[0,1]$ variables. Since $Y_{n+1} = (1 - U_n)Y_n$ it follows that $1 - Y_{n+1} = \sum_{i=1}^n X_i$ increases almost surely to 1 as $n \to \infty$. The distribution on Ω_1 of the nonincreasing rearrangement $(p_i)_i$ of $(X_n)_n$ is called the Poisson–Dirichlet law (with parameter $\theta = 1$) and denoted $\widehat{\mu}_1$.

As has been mentioned in the Introduction, it is the ultimate goal of this work to show that the Poisson–Dirichlet law is the *only* CCF-invariant probability distribution. It will be crucial for the main argument to establish in advance that any such invariant distribution does not put too much weight on very small parts:

PROPOSITION 2.1. *Let $\mu \in \mathcal{M}_1(\Omega_1)$ be CCF-invariant. Then*

$$\int \sum_{i \geq 1} p_i^\alpha \, d\mu < \infty \qquad \text{for all } \alpha > 2/5. \tag{2.1}$$

The proof is deferred to the Appendix.

2.2. *DCF.* In this section we formally introduce the coagulation–fragmentation chain on the discrete version of Ω_1, in which the partition points lie on a finite equidistant grid in $[0,1]$, or its equivalent state space $\mathcal{P}_n$, the set of integer partitions of a fixed $n \in \mathbb{N}$ defined in the Introduction. It will be helpful to view $\mathcal{P}_n$ as the conjugacy classes of the permutation group S_n.

The DCF$^{(n)}$ Markov chain on $\mathcal{P}_n$ is defined similarly to the CCF chain on Ω_1. Identify each $\ell \in \mathcal{P}_n$ with a partition $\bigcup_i A_i$ of $\{1, 2, \dots, n\}$, where for each i, ℓ_i denotes the cardinality of A_i, and sample two independent integers x, y uniformly from $\{1, \dots, n\}$ and without replacement, say $x \in A_i$ and $y \in A_j$. If $i \neq j$, replace A_i and A_j by $A_i \cup A_j$, while if $i = j$ (in which case $\ell_i \geq 2$ since $x \neq y \in A_i$), replace A_i by two of its subsets, consisting respectively of A_i's k smallest elements and of the $\ell_i - k$ remaining ones, where k is uniformly sampled from $\{1, \dots, \ell_i - 1\}$ independently of x and y. In either case relabel and rearrange the new A_i's if necessary.

The transition matrix $K^{(n)}$ of DCF$^{(n)}$ is described as follows: To split into or merge two parts of different sizes j and k ($1 \leq j < k \leq n$), let $\ell, \ell' \in \mathcal{P}_n$ be such that $N_{\ell'}(j) = N_\ell(j) - 1$, $N_{\ell'}(k) = N_\ell(k) - 1$, $N_{\ell'}(j+k) = N_\ell(j+k) + 1$ and $N_{\ell'}(q) = N_\ell(q)$ for all $q \notin \{j, k, j+k\}$. Then

$$\begin{aligned} K^{(n)}(\ell, \ell') &= \frac{2jk}{n(n-1)} N_\ell(j) N_\ell(k) \qquad \text{merge}, \\ K^{(n)}(\ell', \ell) &= \frac{2(j+k)}{n(n-1)} N_{\ell'}(j+k) \qquad \text{split}. \end{aligned} \tag{2.2}$$

To split into or merge two parts of the same size j with $2 \le 2j \le n$ let $\ell, \ell' \in \mathcal{P}_n$ and $0 \le N_{\ell'}(j) = N_\ell(j) - 2$, $N_{\ell'}(2j) = N_\ell(2j) + 1$ and $N_{\ell'}(q) = N_\ell(q)$ for all $q \notin \{j, 2j\}$. Then

$$
\begin{aligned}
K^{(n)}(\ell, \ell') &= \frac{j^2}{n(n-1)} N_\ell(j)\big(N_\ell(j) - 1\big) && \text{merge,}\\
K^{(n)}(\ell', \ell) &= \frac{2j}{n(n-1)} N_{\ell'}(2j) && \text{split.}
\end{aligned}
\tag{2.3}
$$

All other entries of the transition kernel are zero.

It is customary to think of the representation $\{1, 2, \ldots, n\} = \bigcup A_i$ above as the notation for the conjugacy class of a permutation $\sigma \in S_n$. Seen this way, the DCF$^{(n)}$ transition is nothing but the action of a random transposition on S_n's conjugacy classes. Since the random transposition's unique stationary probability measure is the uniform law on S_n (being a finite group convolution), one concludes that the DCF$^{(n)}$'s unique stationary probability measure is the one induced on S_n's conjugacy classes by the uniform law, namely (1.1) (the Ewens sampling formula). In fact, DCF$^{(n)}$ is reversible with respect to $\pi_S^{(n)}$, which can also be checked directly by using (2.2), (2.3) and (1.1) to verify the detailed balance equation $K^{(n)}(\ell, \ell')\pi_S^{(n)}(\ell) = K^{(n)}(\ell', \ell)\pi_S^{(n)}(\ell')$.

3. Coupling of CCF and DCF. To successfully approximate a CCF chain by DCF$^{(n)}$ chains as $n \to \infty$ it is necessary to couple them on a common probability space.

THEOREM 3.1. *For all $\mu \in \mathcal{M}_1(\Omega_1)$ and $\alpha < 1/2$ satisfying*

$$
\int \sum_{i \ge 1} p_i^\alpha \, d\mu < \infty, \tag{3.1}
$$

it is possible to define for all $n \ge 1$ a CCF Markov chain $p(k)$, $k \ge 0$, with initial distribution μ and a $DCF^{(n)}$ Markov chain $\ell(k)$, $k \ge 0$, on the same probability space with probability measure $Q_\mu^{(n)}$ and expectation $E_\mu^{(n)}$, in such a way that

$$
\lim_{n\to\infty} Q_\mu^{(n)}\big[N_{\ell(0)} \ge n^\beta\big] = 0 \qquad \textit{for all } \alpha < \beta \tag{3.2}
$$

and

$$
\lim_{n\to\infty} E_\mu^{(n)}\left[\left| p(\lfloor n^\beta \rfloor) - \frac{\ell(\lfloor n^\beta \rfloor)}{n}\right|_1\right] = 0 \qquad \textit{for all } \beta < 1/2. \tag{3.3}
$$

PROOF. Fix $n \ge 1$. We shall construct a Markov chain (c_k, d_k, e_k), $k \ge 0$, on the state space

$$
\begin{aligned}
\Omega_{cde}^{(n)} := \big\{(c, d, e) | c : [0, n) \to \mathbb{Z} \text{ measurable}, \mathbb{Z} \setminus c[[0, n)] \text{ infinite},\\
d : \{1, \ldots, n\} \to \mathbb{Z}, e \in \{0, 1\}\big\}.
\end{aligned}
$$

Here c and d describe a continuous partition of $[0,n)$ and a discrete partition of $\{1,\dots,n\}$, respectively. The interpretation of c and d in terms of elements of Ω_1 and $\mathcal{P}_n$ is given by the functions $\pi_c:\Omega_{cde}^{(n)}\to\Omega_1$ and $\pi_d:\Omega_{cde}^{(n)}\to\mathcal{P}_n$, respectively, defined by

$$\pi_c(c,d,e):=\operatorname{sort}\left(\left(\frac{\operatorname{Leb}(c^{-1}(\{i\}))}{n}\right)_{i\in\mathbb{Z}}\right) \tag{3.4}$$

and

$$\pi_d(c,d,e):=\operatorname{sort}\big(\big(\sharp d^{-1}(\{i\})\big)_{i\in\mathbb{Z}}\big). \tag{3.5}$$

Here Leb denotes the Lebesgue measure and $\operatorname{sort}((x_i)_i)$ is the sequence obtained by arranging the x_i's in decreasing order, ignoring the 0's if there are infinitely many positive x_i's. Thus two points $x,y\in[0,n)$ belong to the same set in the partition of $[0,n)$ which is described by c iff $c(x)=c(y)$. Analogously, $x,y\in\{1,\dots,n\}$ belong to the same set in the partition of $\{1,\dots,n\}$ described by d iff $d(x)=d(y)$. The CCF Markov chain $p(k)$ and the $\mathrm{DCF}^{(n)}$ Markov chain $\ell(k)$ will be realized as

$$p(k):=\pi_c(c_k,d_k,e_k)\quad\text{and}\quad\ell(k):=\pi_d(c_k,d_k,e_k). \tag{3.6}$$

The flag e_k indicates whether the coupling between the two processes $p(k)$ and $\ell(k)$ is considered to be still in force ($e=0$) or to have already broken down ($e=1$).

The distribution of (c_0,d_0,e_0), that is, the initial distribution of the Markov chain, is defined as the image of μ under the function $\Phi^{(n)}=(\Phi_1^{(n)},\Phi_2^{(n)},0):\Omega_1\longrightarrow\Omega_{cde}^{(n)}$ which assigns to each element of Ω_1 an equivalent function c and an approximating function d as follows (see Figure 1):

$$\Phi_1^{(n)}(p)(x):=\sum_{j\ge1} j\mathbf{1}\left\{x\in[0,np_j)+n\sum_{i=1}^{j-1}p_i\right\},\qquad x\in[0,n),$$

$$\Phi_2^{(n)}(p)(m):=\Phi_1^{(n)}(p)(m-1),\qquad m\in\{1,\dots,n\}.$$

Thus $p(0)=p$ and $\ell(0)=\pi_d(\Phi^{(n)}(p))$ are the initial continuous and discrete partitions generated by $p\in\Omega_1$.

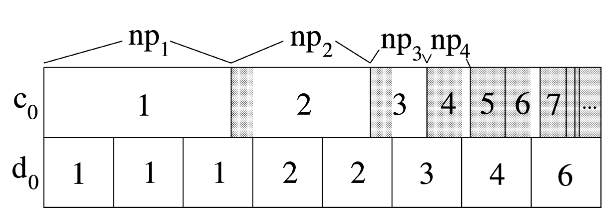


FIG. 1. *Constructing a continuous partition of* $[0,n)$ *and a discrete partition of* $\{1,\dots,n\}$ *from a partition* $p\in\Omega_1$. *Here* $n=8$. *The shaded area indicates the region where the continuous and the discrete numbering disagree.*

Proof of (3.2). To bound $N_{\ell(0)}$, the number of parts in $\ell(0)$, observe that all the pieces in p of size less than $1/n$ can give rise to at most $\sum_i np_i \mathbf{1}_{np_i<1}$ parts (singletons) in $\ell(0)$. Therefore,

$$
\begin{aligned}
N_{\ell(0)} &= \#\Phi_2^{(n)}(p)[\{1,\dots,n\}] = \sum_i \mathbf{1}_{np_i\geq 1} + np_i \mathbf{1}_{np_i<1} \\
&\leq \sum_i (np_i)^\alpha \mathbf{1}_{np_i\geq 1} + (np_i)^\alpha \mathbf{1}_{np_i<1} = n^\alpha \sum_i p_i^\alpha.
\end{aligned} \tag{3.7}
$$

Consequently, due to assumption (3.1),

$$
E_\mu^{(n)}[N_{\ell(0)}] = O(n^\alpha) \tag{3.8}
$$

and hence, for all $\beta > \alpha$,

$$
Q_\mu^{(n)}[N_{\ell(0)} > n^\beta] \leq n^{-\beta} E_\mu^{(n)}[N_{\ell(0)}] = O(n^{\alpha-\beta}),
$$

thus proving (3.2).

We now define informally the kernel of the Markov chain (c_k, d_k, e_k) with state space $\Omega_{cde}^{(n)}$. Assume that the current state of the Markov chain is (c, d, e). To compute the state $(\bar{c}, \bar{d}, \bar{e})$ to which the Markov chain is going to jump in the next step we generate four random variables ξ_1, ξ_2, ζ_1 and ζ_2 such that ξ_1 and ξ_2 and (ζ_1, ζ_2) are independent of each other and of everything else and such that the ξ_i are uniformly distributed on $[0, n)$ and (ζ_1, ζ_2) is uniformly distributed on $[0, n)^2 \setminus \bigcup_{j=1}^n [j-1, j)^2$. The ξ_i will serve to sample uniformly with replacement from $[0, n)$ whereas the ζ_i will be used to sample uniformly without replacement from $\{1, \dots, n\}$ in case the ξ_i have chosen the same atom in d twice. The new continuous partition $\bar{c}$ is then defined as follows:

$$
\text{If} \quad c(\xi_1) \neq c(\xi_2),
\qquad
\bar{c}(x) = \begin{cases} c(\xi_1), & \text{if } c(x) = c(\xi_2), \\ c(x), & \text{else.} \end{cases} \tag{3.9}
$$

$$
\text{If} \quad c(\xi_1) = c(\xi_2),
\qquad
\bar{c}(x) = \begin{cases} \text{new}(c, d), & \text{if } c(x) = c(\xi_1) \text{ and } x > \xi_1, \\ c(x), & \text{else.} \end{cases} \tag{3.10}
$$

We see that the two parts are indeed chosen with probabilities given by their size. In (3.9) two different sets, of sizes $\text{Leb}(c^{-1}(c(\{\xi_i\})))$, $i = 1, 2$, have been selected and are merged by assigning the set $c^{-1}(c(\{\xi_2\})$, hit by ξ_2, the number $c(\{\xi_1\})$ of the set $c^{-1}(c(\{\xi_1\})$, selected by ξ_1. This creates a new set $\bar{c}^{-1}(c(\xi_1))$ with Lebesgue measure $\text{Leb}(c^{-1}(c(\xi_1))) + \text{Leb}(c^{-1}(c(\xi_2)))$.

In (3.10) the set $c^{-1}(c(\{\xi_1\}) = c^{-1}(c(\{\xi_2\})$ is chosen twice, so it has to be split. Since ξ_1 is conditionally uniformly distributed on this set we can reuse it as splitting point for that set: the part to the left of ξ_1 retains its old number $c(\xi_1) = c(\xi_2)$ whereas the part to its right gets a new number $\text{new}(c, d)$, which is not in the range of c or d. Note that it is always possible to find such a new number since $\mathbb{Z}\backslash c[[0, n)]$ is assumed to be infinite. By comparing this with the definition of CCF given at the beginning of the Introduction we see that $p(k)$ defined in (3.6) is a CCF Markov chain.

In the discrete case, the two parts chosen are the ones containing the numbers $\lceil \xi_1 \rceil$ and $\lceil \xi_2 \rceil$, which ensures that the parts are chosen size biased. The rule for merges in the discrete partition is analogous to (3.9):

$$\text{If} \quad d(\lceil \xi_1 \rceil) \neq d(\lceil \xi_2 \rceil),$$
$$\bar{d}(m) = \begin{cases} d(\lceil \xi_1 \rceil), & \text{if } d(m) = d(\lceil \xi_2 \rceil), \\ d(m), & \text{else.} \end{cases} \tag{3.11}$$

Here two different parts with numbers $d(\lceil \xi_1 \rceil)$ and $d(\lceil \xi_2 \rceil)$ have been chosen. They are merged by giving both of them the number $d(\lceil \xi_1 \rceil)$.

The rule for splitting is slightly more complicated. If the same part (but not the same atom) is sampled twice by the ξ_i, then again, as in the continuous setting, ξ_1 determines the point at which the set $d^{-1}(\{\lceil \xi_1 \rceil\})$ is going to be split: the points to the left of $\lceil \xi_1 \rceil$ and the points to the right of $\lceil \xi_1 \rceil$ will constitute the two new fragments; the point $\lceil \xi_1 \rceil$ itself will be attached to the left or the right part in such a way that the splitting rule for $\text{DCF}^{(n)}$, given in (2.2) and (2.3), is imitated. This is done as follows:

If $d(\lceil \xi_1 \rceil) = d(\lceil \xi_2 \rceil)$ and $\lceil \xi_1 \rceil \neq \lceil \xi_2 \rceil$,

$$\bar{d}(m) = \begin{cases} \text{new}(c, d), & \text{if } d(m) = d(\lceil \xi_1 \rceil) \text{and } m > \lceil \xi_1 \rceil, \\ & \quad \text{or } m = \lceil \xi_1 \rceil \text{ and } \xi_1 < \lfloor \xi_1 \rfloor + \dfrac{\# d^{-1}(\{d(\lceil \xi_1 \rceil)\}) \cap [0, \xi_1]}{\# d^{-1}(\{d(\lceil \xi_1 \rceil)\}) - 1}, \\ d(m), & \text{else.} \end{cases} \tag{3.12}$$

Indeed, consider for simplicity the case that the atoms of the set $d^{-1}(\{d(\lceil \xi_1 \rceil)\})$ are not scattered around the whole set $\{1, \dots, n\}$, which they typically will be, but are collected at the bottom: $d^{-1}(\{d(\lceil \xi_1 \rceil)\}) = \{1, \dots, a\}$, where $a := \sharp d^{-1}(\{d(\lceil \xi_1 \rceil)\})$. Definition (3.12) tells us that this set is split into $\{1, \dots, j\}$ and $\{j + 1, \dots, a\}$ if

$$j - 1 + \frac{j - 1}{a - 1} \leq \xi_1 < j + \frac{j}{a - 1}.$$

Conditioned on $\xi_1 \in [0, a)$, the probability for this to happen is $1/(a - 1)$. This means that the discrete set $\{1, \dots, a\}$ is indeed split as described at the beginning of Section 2.2.

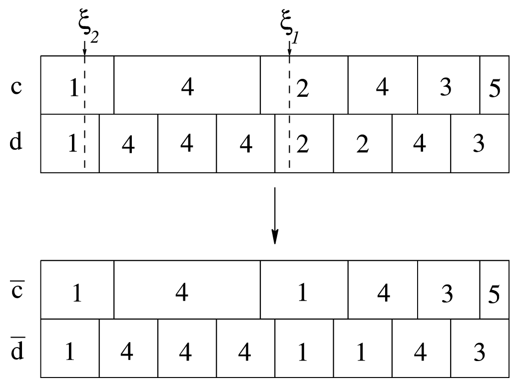


FIG. 2. *Merging the parts with numbers* 1 *and* 2.

If, however, the same atom in d has been sampled twice by the ξ_i's, that is, $\lceil \xi_1 \rceil = \lceil \xi_2 \rceil$, then ξ_1 and ξ_2 are disregarded and $\bar{d}(n)$ is defined as in (3.11) and (3.12) but with (ξ_1, ξ_2) replaced by (ζ_1, ζ_2) in order to sample without replacement. The process $\ell(k)$ defined in (3.6) is a $\mathrm{DCF}^{(n)}$ Markov chain.

It remains to define $\bar{e}$:

$$\bar{e} = \begin{cases} 1, & \text{if } \lceil \xi_1 \rceil = \lceil \xi_2 \rceil \text{ or } c(\xi_1) \neq d(\xi_1) \text{ or } c(\xi_2) \neq d(\xi_2), \\ e, & \text{else.} \end{cases}$$

In the case $\bar{e} = 1$ the coupling has broken down: either the same atom in the discrete partition has been sampled twice by the ξ_i's or at least one of the ξ_i's belongs to noncorresponding sets in the continuous and the discrete partition. The time $\tau := \inf\{k \geq 1 : e_k = 1\}$ is regarded as the decoupling time of the chains $p(k)$ and $\ell(k)$.

The definition of the transition kernel for the Markov chain on $\Omega_{cde}^{(n)}$ is now complete. It is summarized in Figures 2–4.

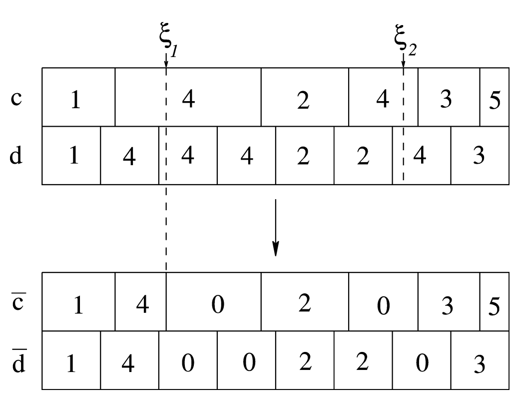


FIG. 3. *Splitting the part with number* 4 *into a part with number* 4 *and a part with number* $\mathrm{new}(c, d) = 0$.

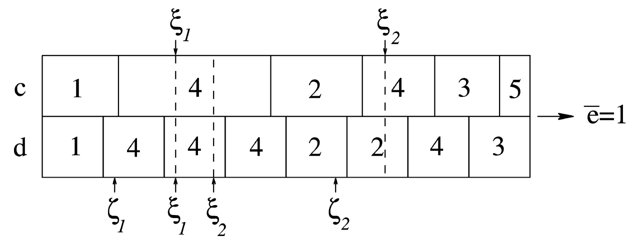


FIG. 4. *Two ways to decouple the chains*: *sampling from the region where c and d disagree* (ξ_2, *top*) *or sampling with replacement from d* (ξ_1 *and* ξ_2, *bottom*).

Proof of (3.3). We denote by

$$\rho_k := \mathrm{Leb}\big(\{x \in [0,n) : c_k(x) \neq d_k(\lceil x \rceil)\}\big)$$

the discrepancy between c_k and d_k. For $k = 0$, this is the roundoff error caused by the approximation of c_0 by d_0; its size is the length of the shaded area in Figure 1. Note that

$$\rho_0 \le N_{\ell(0)} \tag{3.13}$$

because any part in d_0 might disagree with c_0 at most in its rightmost atom. Moreover, ρ_k can increase in each step by at most 1 as long as $k < \tau$: indeed, if two parts are merged, ρ_k does not increase at all (it might even decrease) whereas it might increase by at most $\mathrm{Leb}((\lfloor \xi_1 \rfloor, \lceil \xi_1 \rceil]) = 1$ in case of splitting. Hence, $\rho_{k+1} \le \rho_k + 1$ if $k < \tau$ and therefore

$$\rho_k \le \rho_0 + k \qquad \text{on the event } \{k < \tau\}. \tag{3.14}$$

Since the $|\cdot|_1$-diameter of Ω_1 is at most 2 we have

$$\begin{aligned} &E_\mu^{(n)}\left[\left|p(\lfloor n^\beta \rfloor) - \frac{\ell(\lfloor n^\beta \rfloor)}{n}\right|_1\right] \\ &\quad \le E_\mu^{(n)}\left[\left|p(\lfloor n^\beta \rfloor) - \frac{\ell(\lfloor n^\beta \rfloor)}{n}\right|_1, \lfloor n^\beta \rfloor \le \tau\right] + 2Q_\mu^{(n)}[\tau < \lfloor n^\beta \rfloor]. \end{aligned} \tag{3.15}$$

We are going to bound the first term in (3.15) first. It is easy to see that $|p - q|_1 \ge |\operatorname{sort}(p) - \operatorname{sort}(q)|_1$ for any two summable sequences $p = (p_i)_i$ and $q = (q_i)_i$ of nonnegative numbers. Indeed, if $p_i > p_j$ and $q_i < q_j$, then swapping q_i and q_j would not increase $|p - q|_1$. Therefore, by definitions (3.4)–(3.6), on the event $\{\lfloor n^\beta \rfloor < \tau\}$,

$$\begin{aligned} &\left|p(\lfloor n^\beta \rfloor) - \frac{\ell(\lfloor n^\beta \rfloor)}{n}\right|_1 \\ &\quad \le \frac{1}{n} \sum_{i \ge 1} \left|\mathrm{Leb}\big(c^{-1}_{\lfloor n^\beta \rfloor}(\{i\})\big) - \# d^{-1}_{\lfloor n^\beta \rfloor}(\{i\})\right| \end{aligned}$$

$$\le \frac{1}{n}\sum_{i\ge 1}\mathrm{Leb}\big(\{x : i \in \{c_{\lfloor n^\beta\rfloor}(x), d_{\lfloor n^\beta\rfloor}(x)\}, c_{\lfloor n^\beta\rfloor}(x) \ne d_{\lfloor n^\beta\rfloor}(x)\}\big)$$

$$\le \frac{2}{n}\rho_{\lfloor n^\beta\rfloor} \le \frac{2}{n}(\rho_0 + \lfloor n^\beta\rfloor) \le \frac{2}{n}\big(N_{\ell(0)} + \lfloor n^\beta\rfloor\big)$$

by (3.14) and (3.13). Consequently, due to (3.8), the first term in (3.15) is of order $O(n^{\alpha-1} + n^{\beta-1})$, thus going to 0 as $n \to \infty$.

To show that the second term in (3.15) goes to 0 as well we assume without loss of generality that $\alpha < \beta < 1/2$. Consider the probability that a chain which has not decoupled until the kth step will decouple in the $(k+1)$st step. Given $\rho_0, \dots, \rho_k$, the event that ξ_1 samples two different parts in c_k and d_k has probability ρ_k/n. The same holds for ξ_2. Moreover, the event that one atom in d_k is sampled twice, that is, that $\lceil\xi_1\rceil = \lceil\xi_2\rceil$ has probability $1/n$. Therefore, the probability that either of these events occurs and the chain decouples is at most $(2\rho_k + 1)/n$. On the event $\{\tau > k, \rho_0 < n^\beta\}$ this can be bounded from above, due to (3.14), by $(2(n^\beta + k) + 1)/n$, which is less than $5n^{\beta-1}$ if $k \le n^\beta$. Thus we, get by induction over k,

$$Q_\mu^{(n)}[\tau > k, \rho_0 < n^\beta] \ge (1 - 5n^{\beta-1})^k Q_\mu^{(n)}[\rho_0 < n^\beta]$$

for all $k \le n^\beta$ and hence

$$Q_\mu^{(n)}\big[\tau \ge \lfloor n^\beta\rfloor\big] \ge \big((1 - 5n^{\beta-1})^{n^{1-\beta}}\big)^{n^{2\beta-1}} Q_\mu^{(n)}[\rho_0 < n^\beta]. \tag{3.16}$$

Due to $2\beta - 1 < 0$, the first factor in (3.16) converges to one as $n \to \infty$. The same holds for the second factor due to (3.13) and (3.2). Consequently, also the second term in (3.15) goes to 0, which completes the proof of (3.3). □

4. DCF$^{(n)}$ convergence.

It was mentioned in the Introduction that the *uniform* rate of convergence to $\pi_S^{(n)}$ is too weak to combine properly with $n \to \infty$. However, according to the following theorem (to be proved in Section 4.2), the situation is better when starting off from partitions with relatively few parts and restricting our attention to a certain family $\mathcal{C}$ of Ω_1-neighborhoods to be defined below. Thus, for every $n \in \mathbb{N}$ and $\beta \in (0, 1]$, denote accordingly

$$\mathcal{P}_{n,\beta} = \{\ell \in \mathcal{P}_n : N_\ell < n^\beta\} = \{\ell \in \mathcal{P}_n : \ell_{\lceil n^\beta\rceil} = 0\}.$$

As for the definition of $\mathcal{C}$, for each $k \in \mathbb{N}$ let

$$I_k = \left\{(\mathbf{a}, \mathbf{b}) = (a_i, b_i)_{i=1}^k : 0 < a_i < b_i < 1,\ \sum_{i=1}^k b_i < 1,\ a_k > 1 - \sum_{i=1}^k a_i\right\} \tag{4.1}$$

and denote $\delta_{\mathbf{a},\mathbf{b}} = \min\{1 - \sum_{i=1}^k b_i, a_k - (1 - \sum_{i=1}^k a_i)\}$. Then, for each $(\mathbf{a}, \mathbf{b}) \in I_k$, define

$$C_{\mathbf{a},\mathbf{b}} = \{x \in \Omega_1 : x_i \in (a_i, b_i) \text{ for } i = 1, \dots, k\}, \tag{4.2}$$

which is nonempty if and only if $0 < a_i < \min_{1\le j\le i} b_j$ for $i = 1, \dots, k$, in which case the conditions on $(\mathbf{a}, \mathbf{b}) \in I_k$ guarantee that

$$C_{\mathbf{a},\mathbf{b}} = \{x = (x', x'') : x' \in G_{\mathbf{a},\mathbf{b}}, x'' \in (1 - |x'|_1)\Omega_1\}. \tag{4.3}$$

[Here $(\cdot,\cdot)$ denotes concatenation and $G_{\mathbf{a},\mathbf{b}}$ is the (nonempty) subset of points in $\prod_{i=1}^k (a_i, b_i)$ whose coordinates are nonincreasing.] Moreover, I_k's definition (4.1) implies that

$$\delta_{\mathbf{a},\mathbf{b}} < |x''|_1 < x'_k - \delta_{\mathbf{a},\mathbf{b}} \qquad \forall (x', x'') \in C_{\mathbf{a},\mathbf{b}}. \tag{4.4}$$

Finally,

$$\mathcal{C} = \{C_{\mathbf{a},\mathbf{b}} : (\mathbf{a}, \mathbf{b}) \in I_k,\ k \ge 1\}. \tag{4.5}$$

The family $\mathcal{C}$ of Ω_1-neighborhoods will be shown in Section 5 to be sufficiently rich to characterize $\widehat{\mu}_1$ uniquely. At the same time, and as a result of their special features (4.3) and (4.4), the convergence of the DCF$^{(n)}$ to its equilibrium is fast on the sets in $\mathcal{C}$:

THEOREM 4.1. *Fix $\beta \in (0, \frac{1}{2})$. For each $n \in \mathbb{N}$ let $(X^{(n)}(k))_{k\ge 0}$ be a DCF$^{(n)}$ Markov chain with underlying probability measure $P^{(n)}$ and initial distribution $\mu_0^{(n)} \in \mathcal{M}_1(\mathcal{P}_{n,\beta})$. Then, for any $C \in \mathcal{C}$, $\beta' > \beta$ and integer sequence $k = k_n \ge n^{\beta'}$,*

$$\Delta_C^{(n)}(k) := P^{(n)}\big(X^{(n)}(k) \in nC\big) - \pi_S^{(n)}(nC) \underset{n\to\infty}{\longrightarrow} 0.$$

4.1. *Characters in S_n—background.* Recall that the partition space $\mathcal{P}_n$ can be viewed as the quotient of the permutation group S_n under conjugacy. Thus the natural inner product on $F_n := \{f : \mathcal{P}_n \to \mathbb{R}\}$ is

$$\langle f, g\rangle = \langle f, g\rangle_n = \sum_{\gamma\in\mathcal{P}_n} f(\gamma)g(\gamma)\pi_S^{(n)}(\gamma).$$

The fact mentioned earlier that $\pi_S^{(n)}$ is a reversing measure for the DCF$^{(n)}$ means precisely that $K^{(n)}$ is self-adjoint with respect to this inner product.

The following basic facts regarding the character theory of S_n, as well as the full theory, can be found, for example, in [10], and their relevance to random group actions (such as transpositions in our case) in [4, 6]. The characters $\{\chi\}$ of S_n (traces of the irreducible representations) are functions on S_n, constant on conjugacy classes, and as such can be seen to be functions on $\mathcal{P}_n$. They are orthonormal under $\langle\cdot,\cdot\rangle$ and since there are $\#\mathcal{P}_n$ of them, they are indexed by the partitions $((\chi_\lambda)_{\lambda\in\mathcal{P}_n})$ and form an orthonormal base of F_n.

Since $K^{(n)}$ represents a random transposition, its dual $K^{(n)^*}$ acts on $\mathcal{M}_1(S_n)$ as a convolution

$$K^{(n)^*}\mu = \kappa^{(n)} \star \mu \qquad \left(\kappa^{(n)}(\text{transposition}) = \frac{2}{n(n-1)} \quad \text{and} \quad 0 \quad \text{otherwise}\right)$$

as a result of which, and of a corollary ([4], Chapter 2, Proposition 6) of Schur's lemma, the following hold:

(a) $K^{(n)}$'s eigenfunctions are the characters $(\chi_\lambda)_{\lambda\in\mathcal{P}_n}$;

(b) the eigenvalue $\theta_\lambda^{(n)}$ corresponding to χ_λ is given by $\chi_\lambda(\text{transposition})/\chi_\lambda(\text{identity})$.

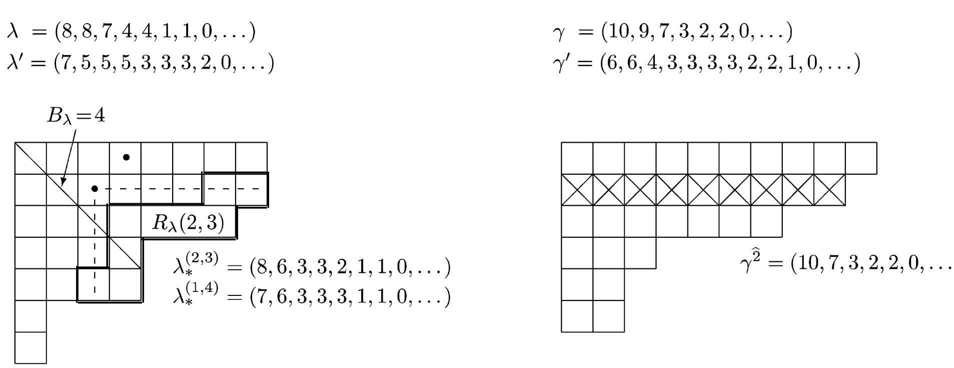
A result of Frobenius in principle provides formulae for all characters. Although in general they can be intractable, this is not so at transpositions and at the identity, thus yielding ([4], D-2, page 40)

$$\theta_\lambda^{(n)} = \frac{1}{n(n-1)}\sum_j \lambda_j(\lambda_j - 2j + 1) = \frac{1}{n(n-1)}\left(\sum_{i=1}^{n}\lambda_i^2 - \sum_{j=1}^{\lambda_1}{\lambda'_j}^2\right). \tag{4.6}$$

(λ's adjoint partition λ' is defined below). In particular $\theta_{(n,0,\dots)}^{(n)} = 1$ and $\chi_{(n,0,\dots)} \equiv 1$.

For many purposes, a partition $\lambda \in \mathcal{P}_n$ can be best described by its Young diagram Υ_λ (Figure 5), consisting of N_λ rows of $\lambda_1, \dots, \lambda_{N_\lambda}$ cells, respectively, in terms of which some useful features of λ can be defined [the jth cell in row i is denoted (i, j)]:

1. $\lambda' \in \mathcal{P}_n$ is the partition whose Young diagram is obtained from λ's by transposition: $\Upsilon_{\lambda'} = \Upsilon_\lambda^T$;
2. $B_\lambda = \max\{i : (i,i) \in \Upsilon_\lambda\} = \max\{i : \lambda_i \geq i\}$ (λ's diagonal length);
3. $R_\lambda(i, j) = \{(u, v) : i \leq u \leq \lambda'_j,\ j \vee \lambda_{u+1} \leq v \leq \lambda_u\}$ [Υ_λ's rim segment straddled by (i, j)];
4. $\Upsilon_{\lambda_*^{(i,j)}} = \Upsilon_\lambda \setminus R_\lambda(i, j)$ defines $\lambda_*^{(i,j)}$ (a diagram obtained from λ's by removing a rim segment is a Young diagram; this defines the partition $\lambda_*^{(i,j)}$).



Murnaghan–Nakayama rule: $\chi_\lambda(\gamma) = \chi_{\lambda_*^{(1,4)}}(\gamma^{\widehat{2}}) - \chi_{\lambda_*^{(2,3)}}(\gamma^{\widehat{2}})$

FIG. 5. *Young diagrams of* $\lambda, \gamma \in \mathcal{P}_{33}$. *Two* λ*-cells*, $(1,4)$ *and* $(2,3)$, *generate rim segments of size* 9, *the latter shown explicitly, which the Murnaghan–Nakayama rule "peels off" together with the deletion of* γ_2.

In addition, for any $\gamma \in \mathcal{P}_n$, define $\gamma^{\widehat{r}} = (\gamma_1, \dots, \widehat{\gamma_r}, \dots) \in \mathcal{P}_{n-\gamma_r}$, the partition obtained from γ by removing its rth part. The following Murnaghan–Nakayama rule (see [6], Theorem 3.4) provides a way of recursively evaluating characters: for all $\lambda, \gamma \in \mathcal{P}_n$ and $1 \le r \le N_\gamma$

$$\chi_\lambda(\gamma) = \sum_{(i,j):\#R_\lambda(i,j)=\gamma_r} (-1)^{\lambda'_j - i} \chi_{\lambda_*^{(i,j)}}(\gamma^{\widehat{r}}) \tag{4.7}$$

in the sense that the sum is zero if its index set is empty, and $\chi_\varnothing(\varnothing) = 1$. Thus, for a fixed order in which γ's parts are chosen, $\chi_\lambda(\gamma)$ can be calculated by covering all possible ways of successively stripping off γ_r-sized rim segments from λ's diagram, and $\chi_\lambda(\gamma) = 0$ if it is impossible to exhaust Υ_λ entirely in this way. In particular

$$N_\gamma < B_\lambda \quad \Longrightarrow \quad \chi_\lambda(\gamma) = 0 \tag{4.8}$$

since any rim segment of λ contains at most one diagonal cell (i, i).

4.2. *Proof of Theorem* 4.1. Before proceeding with the proof itself, it will be helpful to characterize the $\gamma \in \mathcal{P}_n$ which belong to $nC = nC_{\mathbf{a},\mathbf{b}}$ for given $k \in \mathbb{N}$ and $(\mathbf{a}, \mathbf{b}) \in I_k$ (assuming $C \neq \varnothing$). It follows from $C_{\mathbf{a},\mathbf{b}}$'s description (4.3) that any such γ can be expressed as a concatenation (γ', γ''), where $\gamma' \in G^{(n)}_{\mathbf{a},\mathbf{b}}$ and $\gamma'' \in \mathcal{P}_{n-|\gamma'|_1}$, and where $G^{(n)}_{\mathbf{a},\mathbf{b}}$ consists of nonincreasing integer-valued k-sequences γ' which by virtue of (4.4) satisfy

$$\begin{aligned} &\text{(i)} \quad |\gamma'|_1 < n; \\ &\text{(ii)} \quad \exists\, \delta = \delta(C) > 0 \quad \text{such that} \quad \gamma'_k > (n - |\gamma'|_1) + \delta n. \end{aligned} \tag{4.9}$$

This state of affairs is illustrated in Figure 6.

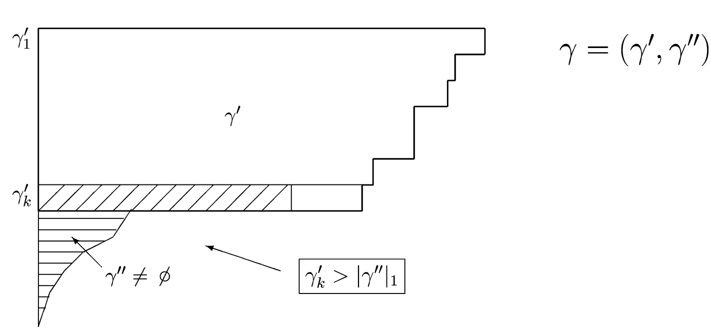


FIG. 6. *A partition γ in $nC_{\mathbf{a},\mathbf{b}}$ splits into its first k rows γ' and the remainder γ'' which is nonempty but smaller in size than γ''s last row.*

PROOF OF THEOREM 4.1. Fix $C \in \mathcal{C}$ and define $f_n = \mathbf{1}_{nC}$. Then, in terms of $\mu_0^{(n)}$'s density $g_0^{(n)}(\gamma) = \mu_0^{(n)}(\gamma)/\pi_S^{(n)}(\gamma)$:

$$P^{(n)}\big(X^{(n)}(k) \in nC\big) = \sum_{\gamma} \mu_0^{(n)}(\gamma) K^{(n)^k} f_n(\gamma) = \langle g_0^{(n)}, K^{(n)^k} f_n\rangle$$
$$= \sum_{\lambda \in \mathcal{P}_n} \theta_\lambda^{(n)^k} \langle g_0^{(n)}, \chi_\lambda\rangle \langle f_n, \chi_\lambda\rangle$$

and, since $\theta_{(n,0,\dots)}^{(n)} = 1$ and $\chi_{(n,0,\dots)} \equiv 1$,

$$\pi_S^{(n)}(nC) = \langle f_n, 1\rangle = \theta_{(n,0,\dots)}^{(n)^k} \langle g_0^{(n)}, \chi_{(n,0,\dots)}\rangle \langle f_n, \chi_{(n,0,\dots)}\rangle$$

so that

$$\Delta_C^{(n)}(k) = \sum_{(n,0,\dots) \neq \lambda \in \mathcal{P}_n} \theta_\lambda^{(n)^k} \langle g_0^{(n)}, \chi_\lambda\rangle \langle f_n, \chi_\lambda\rangle. \tag{4.10}$$

By assumption, $g_0^{(n)}(\gamma) = 0$ whenever $N_\gamma > n^\beta$. On the other hand, $\chi_\lambda(\gamma) = 0$ whenever $B_\lambda > n^\beta$ and $N_\gamma \le n^\beta$ by the consequence (4.8) of Murnaghan–Nakayama's rule. Thus (4.10) becomes

$$\Delta_C^{(n)}(k) = \sum_{(n,0,\dots) \neq \lambda \in \mathcal{P}_n, B_\lambda \le n^\beta} \theta_\lambda^{(n)^k} \langle g_0^{(n)}, \chi_\lambda\rangle \langle f_n, \chi_\lambda\rangle.$$

Now choose an η such that $1 - (\beta' - \beta) < \eta < 1$ and let $n_0 = 5^{1/(1-\eta)}$. Then, for all $n \ge n_0$,

$$\Delta_C^{(n)}(k) = \left(\sum_{\lambda \in \mathcal{P}'_n} + \sum_{\lambda \in \mathcal{P}''_n} + \sum_{\lambda \in \mathcal{P}'''_n}\right) \theta_\lambda^{(n)^k} \langle g_0^{(n)}, \chi_\lambda\rangle \langle f_n, \chi_\lambda\rangle, \tag{4.11}$$

where

$$\mathcal{P}'_n = \mathcal{P}'_n(\eta, \beta) = \{\lambda \in \mathcal{P}_n : B_\lambda \le n^\beta, \lambda_1 \le n - 2n^\eta, N_\lambda \le n - 2n^\eta\},$$
$$\mathcal{P}''_n = \mathcal{P}''_n(\eta, \beta) = \{\lambda \in \mathcal{P}_n : B_\lambda \le n^\beta, n - 2n^\eta < \lambda_1 < n\},$$
$$\mathcal{P}'''_n = \mathcal{P}'''_n(\eta, \beta) = \{\lambda \in \mathcal{P}_n : B_\lambda \le n^\beta, n - 2n^\eta < N_\lambda\}.$$

[Our choice of n_0 ensures that $\mathcal{P}''_n$ and $\mathcal{P}'''_n$ are disjoint and that $(n, 0, \dots) \notin \mathcal{P}'''_n$.] It turns out that, for n large enough, the terms in (4.11) vanish for all $\lambda \in \mathcal{P}''_n \cup \mathcal{P}'''_n$, whereas when $\lambda \in \mathcal{P}'_n$ the factor $|\theta_\lambda^{(n)}|$ is sufficiently separated from 1:

LEMMA 4.2. *$\exists n_1 = n_1(\beta, C)$ such that $\langle f_n, \chi_\lambda\rangle = 0$ $\forall n \ge n_1$, $\forall \lambda \in \mathcal{P}''_n \cup \mathcal{P}'''_n$.*

LEMMA 4.3. *For all* $\lambda \in \mathcal{P}_n$, $|\theta_\lambda^{(n)}| \le \frac{\lambda_1 \vee N_\lambda}{n}$ *and thus* $\exists\, n_2 = n_2(\eta)$ *such that*

$$|\theta_\lambda^{(n)}| \le e^{-n^{\eta-1}} \qquad \forall\, n \ge n_2, \forall\, \lambda \in \mathcal{P}'_n. \tag{4.12}$$

PROOF OF LEMMA 4.2. Consider first $\lambda \in \mathcal{P}''_n$. Now, $C = C_{\mathbf{a},\mathbf{b}}$ for some $k \in \mathbb{N}$ and $(\mathbf{a}, \mathbf{b}) \in I_k$, so that, as discussed at the beginning of the section and illustrated in Figure 6, γ can be split into (γ', γ'') and

$$\langle f_n, \chi_\lambda \rangle = \sum_{\gamma' \in G_{\mathbf{a},\mathbf{b}}^{(n)}} \sum_{\gamma'' \in \mathcal{P}_{n-|\gamma'|_1}} \pi_S^{(n)}(\gamma', \gamma'')\chi_\lambda(\gamma', \gamma''). \tag{4.13}$$

[Note that property (4.9)(i) guarantees that the inner sum is not vacuous, i.e., $|\gamma''|_1 > 0$.]

We shall show that for every fixed $\gamma' \in G_{\mathbf{a},\mathbf{b}}^{(n)}$ the inner sum in (4.13) equals zero. First apply the Murnaghan–Nakayama rule (4.7) k times to $\chi_\lambda(\gamma', \gamma'')$ by successively stripping rim segments from λ, of lengths γ'_i at each stage i, $i = 1, \dots, k$. On the one hand $\lambda_1 > (1-\delta)n$ for $n \ge n''_1 = n''_1(\beta, \eta, C)$ (since $\lambda \in \mathcal{P}''_n$), and on the other $\gamma'_i > \delta n$, $i = 1, \dots, k$ [by (4.9)(ii)]. This implies that at each of these k reduction stages precisely one rim segment can be stripped off, namely the last γ'_i cells of whatever remains of λ_1, $i = 1, \dots, k$. Summing up,

$$\chi_\lambda(\gamma', \gamma'') = \chi_{\lambda^*}(\gamma''), \tag{4.14}$$

where $\lambda^* \in \mathcal{P}_{n-|\gamma'|_1}$ is defined by $\lambda_1^* = \lambda_1 - |\gamma'|_1$ and $\lambda_j^* = \lambda_j$, $j \ge 2$. As for the first factor of the summand in (4.13), note that (4.9)(ii) implies $\gamma'_k > \gamma''_1$ (see Figure 6) and thus

$$\frac{\pi_S^{(n)}(\gamma', \gamma'')}{\pi_S^{(n-|\gamma'|_1)}(\gamma'')} = \frac{1}{\prod_{i=1}^k \gamma'_i \prod_j N_{\gamma'}(j)!} =: R(\gamma'). \tag{4.15}$$

Inserting (4.14) and (4.15) in the inner sum of (4.13) we obtain

$$\sum_{\gamma'' \in \mathcal{P}_{n-|\gamma'|_1}} \pi_S^{(n)}(\gamma', \gamma'')\chi_\lambda(\gamma', \gamma'') = R(\gamma')\langle \chi_{\lambda^*}, 1\rangle_{n-|\gamma'|_1} = 0$$

since λ^* is not the trivial partition, that is, $\lambda^* \ne (n - |\gamma'|_1, 0, \dots)$ [because $\lambda \ne (n, 0, \dots)$], and thus χ_{λ^*} is orthogonal to $\chi_{(n-|\gamma'|_1, 0, \dots)} \equiv 1$.

The proof for $\lambda \in \mathcal{P}'''_n$ is similar, with $n \ge n'''_1 = n'''_1(\beta, \eta, C)$, where now the only rim segments which can be stripped off from λ are from its first column. It remains to define $n_1 = n''_1 \vee n'''_1$. $\square$

PROOF OF LEMMA 4.3. Using the formula for $\theta_\lambda^{(n)}$ given in (4.6),

$$\theta_\lambda^{(n)} = \frac{1}{n(n-1)}\left(\sum_{i=1}^n \lambda_i^2 - \sum_{j=1}^{\lambda_1} {\lambda'_j}^2\right) \le \frac{1}{n(n-1)}(\lambda_1 n - \lambda_1) = \frac{\lambda_1}{n},$$

whereas, by duality, $-\theta_\lambda^{(n)} = \theta_{\lambda'}^{(n)} \le \lambda_1'/n = N_\lambda/n$. Moreover, for $\lambda \in \mathcal{P}_n'$,

$$|\theta_\lambda^{(n)}| \le \left(1 - \frac{2}{n^{1-\eta}}\right) = \left(1 - \frac{2}{n^{1-\eta}}\right)^{n^{1-\eta}\, n^{\eta-1}} \le e^{-n^{\eta-1}}$$

as soon as $(1 - \frac{2}{n^{1-\eta}})^{n^{1-\eta}} \le \frac{1}{e}$. □

We now continue with the estimation of (4.11). As a result of Lemmas 4.2 and 4.3, and recalling that $k \ge n^{\beta'}$, it holds for any $n \ge n_1 \vee n_2$, that

$$|\Delta_C^{(n)}(k)| \le \sum_{\lambda \in \mathcal{P}_n'} e^{-n^{\beta'+\eta-1}} |\langle g_0^{(n)}, \chi_\lambda\rangle||\langle f_n, \chi_\lambda\rangle|. \tag{4.16}$$

To estimate the number of terms in (4.16), note that the Young diagram Υ_λ of any $\lambda \in \mathcal{P}_n$ with $B_\lambda = s$ consists of an $s \times s$ square of cells, with (certainly no more than $n-1$) cells added to each one of the square's s rows and s columns. Ignoring the various additional constraints, there are n^{2s} ways of making such additions, and thus for any $t > 0$, $\#\{\lambda \in \mathcal{P}_n : B_\lambda \le t\} \le tn^{2t}$, so that

$$\#\mathcal{P}_n' \le \#\{\lambda \in \mathcal{P}_n : B_\lambda \le n^\beta\} \le n^\beta n^{2n^\beta} \le e^{3n^\beta \log n}.$$

As for the terms in (4.16), $|\langle f_n, \chi_\lambda\rangle| \le 1$ by the Cauchy–Schwarz inequality, and

$$\sup_{\lambda \in \mathcal{P}_n'} |\langle g_0^{(n)}, \chi_\lambda\rangle| \le \sup_{\lambda \in \mathcal{P}_n'\ N_\gamma \le n^\beta} |\chi_\lambda(\gamma)| \le n^{n^\beta} = e^{n^\beta \log n},$$

where the second inequality follows from applying Murnaghan–Nakayama's rule at most n^β times, each time with not more that n terms in the sum (4.7).

The above and (4.16) imply that, for all $n \ge \max\{n_0, n_1, n_2\}$,

$$|\Delta_C^{(n)}(k)| \le \exp\{-n^\beta\bigl(n^{(\beta'-\beta)-(1-\eta)} - 4\log n\bigr)\}.$$

Eventually, thus, $|\Delta_C^{(n)}(k)| \le e^{-n^\beta/2}$, which concludes the proof of Theorem 4.1. □

5. Proof of Vershik's conjecture. This section is devoted to the proof of Conjecture 1.1, which we restate as follows.

THEOREM 5.1. *If $\mu \in \mathcal{M}_1(\Omega_1)$ is CCF-invariant, then μ is the Poisson–Dirichlet measure $\widehat{\mu}_1$.*

The main ingredients in its proof have been established in Sections 2–4 and are, respectively, the a priori finite moment estimate Proposition 2.1, the couplings with approximating DCF$^{(n)}$'s of Theorem 3.1, and the fast convergence to equilibrium of the DCF$^{(n)}$ chains in the sense of Theorem 4.1.

PROOF OF THEOREM 5.1. Let $\Omega_1' = \{p \in \Omega_1 : \exists$ infinitely many $n \in \mathbb{N}$ such that $p_n > \sum_{j>n} p_j\}$. We shall show that

$$\widehat{\mu}_1(\Omega_1') = 1, \tag{5.1}$$

$$\{C \cap \Omega_1' : C \in \mathcal{C}\} \subset \mathcal{B}_{\Omega_1'} \qquad \text{is measure determining on } (\Omega_1', \mathcal{B}_{\Omega_1'}), \tag{5.2}$$

$$\mu(C) = \widehat{\mu}_1(C) \qquad \forall\, C \in \mathcal{C}, \tag{5.3}$$

which together imply in particular that $\mu(\Omega_1') = 1$, and indeed the theorem's statement as well.

Proof of (5.1). Recall $\widehat{\mu}_1$'s description as the law of the nonincreasing rearrangement of the uniform stickbreaking process X_n (with $Y_n = 1 - \sum_{j<n} X_j$ the remaining stick length prior to the nth break and $X_n = U_n Y_n$) and define $\tau_1 = 1$, $\tau_{k+1} = \min\{n > \tau_k : X_n \wedge (Y_n - X_n) < X_j, \forall\, j \le \tau_k\}$ for $k \ge 1$. Since a.s. $X_n \searrow 0$, each τ_k is finite. We claim that

$$\begin{aligned} &A_k := \{U_{\tau_k} > \tfrac{1}{2}\} = \Big\{X_{\tau_k} > \sum_{j>\tau_k} X_j\Big\} \qquad \text{are independent,} \\ &P(A_k) = \tfrac{1}{2} \qquad \forall\, k. \end{aligned} \tag{5.4}$$

This implies that a.s. $U_{\tau_k} > \frac{1}{2}$ infinitely often, and these $n = \tau_k$ will be the ones alluded to in Ω_1''s definition. Indeed, on A_k, $\sum_{j>\tau_k} X_j < X_i\ \forall\, i \le \tau_k$, so that the nondecreasing permutation of the X_i's decouples on $[1, \tau_k]$ and (τ_k, ∞) and thus $p_{\tau_k} = \min_{i \le \tau_k} X_i > \sum_{j>\tau_k} X_j = \sum_{j>\tau_k} p_j$.

To prove (5.4), represent the splitting variables as

$$U_n = \begin{cases} V_n, & \text{if } \eta_n = 1, \\ 1 - V_n, & \text{if } \eta_n = 0, \end{cases}$$

where $V_n \sim U[0, 1]$ and $\eta_n \sim$ Bernoulli(0.5) are independent of each other, and write $A_k = (B_k \cap C_k) \cup (B_k^C \cap C_k^C)$, with $B_k = \{V_{\tau_k} > \frac{1}{2}\}$ and $C_k = \{\eta_{\tau_k} = 1\}$. The τ_k are $\mathcal{F}$-stopping times, where $\mathcal{F}_n = \sigma(V_1, \ldots, V_n, \eta_0, \ldots, \eta_{n-1})$ (arbitrarily set $\eta_0 = 1$) so that $B_k \in \mathcal{F}_{\tau_k}$ and C_k is independent of $\mathcal{F}_{\tau_k}$ [in particular $P(C_k) = 0.5$] for all k. For any $D \in \mathcal{F}_{\tau_k}$,

$$\begin{aligned} P(D \cap A_k) &= P\big((D \cap B_k) \cap C_k\big) + P\big((D \cap B_k^C) \cap C_k^C\big) \\ &= P(D \cap B_k) P(C_k) + P(D \cap B_k^C) P(C_k^C) \\ &= \tfrac{1}{2} P(D). \end{aligned}$$

Choosing first $D = \Omega_1$ and then $D = \bigcap_{j \in J} A_j$ with $J \subset \{1, \ldots, k-1\}$ (indeed, $A_j \in \mathcal{F}_{\tau_j + 1} \subset \mathcal{F}_{\tau_k}$ for $j < k$), we, respectively, obtain $P(A_k) = 0.5$ and the independence of the A_n's. We have proved (5.4) and thus (5.1).

Proof of (5.2). Fix $\varepsilon > 0$, $p \in \Omega_1'$, and choose k large enough so that $0 < q := \sum_{j>k} p_j < \min(p_k, \frac{\varepsilon}{4})$. Then let $\delta = \frac{q \wedge (p_k - q)}{k+2}$ and $a_i = p_i - \delta$, $b_i = p_i + \delta$ for $i = 1, \dots, k$. We claim that $(\mathbf{a}, \mathbf{b}) \in I_k$. Indeed,

$$\sum_{i \le k} b_i = \sum_{i \le k} (p_i + \delta) \le 1 - q + \frac{k}{k+2} q < 1,$$

whereas

$$\begin{aligned} a_k + \sum_{i \le k} a_i &= (p_k - \delta) + \sum_{i \le k} (p_i - \delta) \\ &= (p_k + 1 - q) - (k+1)\delta > 1. \end{aligned}$$

By definition $p \in C_{\mathbf{a},\mathbf{b}}$. Moreover, for any $x \in C_{\mathbf{a},\mathbf{b}}$,

$$\begin{aligned} |x - p|_1 &\le 2 \sum_{j=1}^{k} |x_j - p_j| + 2 \sum_{j=k+1}^{\infty} p_j \le 2k\delta + 2q \\ &\le 2\left(1 + \frac{k}{k+1}\right) q < \varepsilon, \end{aligned}$$

which shows that for any open l_1-ball $B_\varepsilon(p)$ in Ω_1' there is some $C \in \mathcal{C}$ such that $p \in C \cap \Omega_1' \subset B_\varepsilon(p)$. In other words, $\{C \cap \Omega_1', C \in \mathcal{C}\}$ generates Ω_1''s topology.

To conclude the proof of (5.2) we need to check that $\mathcal{C}$ is closed under intersections. For any $j \le k$ then, let $(\mathbf{a}_1, \mathbf{b}_1) \in I_j$ and $(\mathbf{a}_2, \mathbf{b}_2) \in I_k$, and if $j < k$ denote $a_{1_i} = 0$ and $b_{1_i} = 1$ for $i = j+1, \dots, k$. It follows immediately that $(\mathbf{a}, \mathbf{b})$ defined by $a_i = a_{1_i} \vee a_{2_i}$ and $b_i = b_{1_i} \wedge b_{2_i}$ for $i = 1, \dots, k$ belongs to I_k, and $C_{\mathbf{a}_1,\mathbf{b}_1} \cap C_{\mathbf{a}_2,\mathbf{b}_2} = C_{\mathbf{a},\mathbf{b}}$.

Proof of (5.3). First note that if $(\mathbf{a}, \mathbf{b}) \in I_k$, then $((1+\varepsilon)\mathbf{a}, (1-\varepsilon)\mathbf{b}) \in I_k$ for all ε in some neighborhood of 0, and if $C := C_{\mathbf{a},\mathbf{b}} \neq \varnothing$, then so is $C^{(\varepsilon)} := C_{(1+\varepsilon)\mathbf{a},(1-\varepsilon)\mathbf{b}}$ for all ε in a neighborhood of 0.

Once we show that, for all $\varepsilon > 0$ small enough,

$$\widehat{\mu}_1(C^{(-\varepsilon)}) \ge \mu(C) \ge \widehat{\mu}_1(C^{(\varepsilon)}),$$

let $\varepsilon \searrow 0$ and use $\widehat{\mu}_1(\partial C) = 0$ to obtain $\mu(C) = \widehat{\mu}_1(C)$ for every $C \in \mathcal{C}$, thus proving (5.3) and with it the theorem.

Let $\frac{2}{5} < \alpha < \beta < \gamma < \frac{1}{2}$ be three otherwise arbitrary numbers. Since, by Proposition 2.1, μ satisfies (3.1), we consider for every $n \in \mathbb{N}$ the probability measure $Q^{(n)} = Q_\mu^{(n)}$ introduced in Proposition 3.1 which is defined on a space which supports both a CCF Markov chain $p(\cdot)$ with μ as its stationary marginal and a DCF$^{(n)}$ Markov chain $\ell^{(n)}(\cdot)$ which "emulates" $p(\cdot)$ in terms of its initial law [cf. (3.2)] and in the sense that they remain close after n^γ units of time [cf. (3.3)]. For any $n \in \mathbb{N}$,

$$\begin{aligned}
\mu(C) - \widehat{\mu}_1\bigl(C^{(\varepsilon)}\bigr) &= Q^{(n)}\bigl(p(\lfloor n^\gamma \rfloor) \in C\bigr) - \widehat{\mu}_1\bigl(C^{(\varepsilon)}\bigr) \\
&\ge Q^{(n)}\bigl(p(\lfloor n^\gamma \rfloor) \in C\bigr) - Q^{(n)}\biggl(\frac{1}{n}\ell^{(n)}(\lfloor n^\gamma \rfloor) \in C^{(\varepsilon)}\biggr) \\
&\quad - \biggl| Q^{(n)}\biggl(\frac{1}{n}\ell^{(n)}(\lfloor n^\gamma \rfloor) \in C^{(\varepsilon)}\biggr) - \pi^{(n)}\bigl(nC^{(\varepsilon)}\bigr)\biggr| \\
&\quad + \bigl(\pi^{(n)}\bigl(nC^{(\varepsilon)}\bigr) - \widehat{\mu}_1\bigl(C^{(\varepsilon)}\bigr)\bigr) \\
&=: D_1^{(\varepsilon)} - D_2^{(\varepsilon)} + D_3^{(\varepsilon)}.
\end{aligned}$$

The first term is estimated using a simple union bound with $\varepsilon' := \varepsilon \min\{a_i : 1 \le i \le k\}$, and (3.3):

$$\begin{aligned}
D_1^{(\varepsilon)} &\ge -Q^{(n)}\biggl(\biggl| p(\lfloor n^\gamma \rfloor) - \frac{1}{n}\ell^{(n)}(\lfloor n^\gamma \rfloor)\biggr|_1 > \varepsilon'\biggr) \\
&\ge \frac{-1}{\varepsilon'} E_{Q^{(n)}} \biggl| p(\lfloor n^\gamma \rfloor) - \frac{1}{n}\ell^{(n)}(\lfloor n^\gamma \rfloor)\biggr|_1 \xrightarrow[n\to\infty]{} 0.
\end{aligned}$$

To estimate $D_2^{(\varepsilon)}$ we would like to apply Theorem 4.1 to the sequence of discrete processes $\ell^{(n)}(\cdot)$. Their initial laws, however, are guaranteed by (3.2) to be only *nearly* supported on $\mathcal{P}_{n,\beta}$, respectively, but not totally as required by Theorem 4.1. Define thus $\widetilde{Q}^{(n)}(\cdot) := Q^{(n)}(\cdot | \ell^{(n)}(0) \in \mathcal{P}_{n,\beta})$; obviously $\widetilde{Q}^{(n)}(\ell^{(n)}(0) \in \mathcal{P}_{n,\beta}) = 1$, and under $\widetilde{Q}^{(n)}$, $\ell^{(n)}(\cdot)$ remains a DCF$^{(n)}$ chain. Then

$$\begin{aligned}
D_2^{(\varepsilon)} &\le \bigl|\widetilde{Q}^{(n)}\bigl(\ell^{(n)}(\lfloor n^\gamma \rfloor) \in nC^{(\varepsilon)}\bigr) - \pi^{(n)}\bigl(nC^{(\varepsilon)}\bigr)\bigr| \\
&\quad + \|\widetilde{Q}^{(n)} - Q^{(n)}\|_{\mathrm{var}} \xrightarrow[n\to\infty]{} 0.
\end{aligned}$$

Here we applied Theorem 4.1 for the first term, while

$$\|\widetilde{Q}^{(n)} - Q^{(n)}\|_{\mathrm{var}} \le \frac{Q^{(n)}(\ell^{(n)}(0) \notin \mathcal{P}_{n,\beta})}{Q^{(n)}(\ell^{(n)}(0) \in \mathcal{P}_{n,\beta})} \to 0 \qquad \text{by (3.2).}$$

Finally, recall that $\pi_S^{(n)}(n\cdot) \to \widehat{\mu}_1$ weakly [12, 18], and since $C^{(\varepsilon)}$ satisfies $\widehat{\mu}_1(\partial C^{(\varepsilon)}) = 0$, it follows that $\lim_{n\to\infty} \pi_S^{(n)}(nC^{(\varepsilon)}) = \widehat{\mu}_1(C^{(\varepsilon)})$. Thus $\lim_{n\to\infty} D_3^{(\varepsilon)} = 0$. Consequently

$$\mu(C) - \widehat{\mu}_1\bigl(C^{(\varepsilon)}\bigr) \ge \liminf_{n\to\infty} D_1^{(\varepsilon)} - \limsup_{n\to\infty} D_2^{(\varepsilon)} + \liminf_{n\to\infty} D_3^{(\varepsilon)} \ge 0.$$

The reverse inequality $\widehat{\mu}_1(C^{(-\varepsilon)}) \ge \mu(C)$ is obtained similarly from

$$\widehat{\mu}_1\bigl(C^{(-\varepsilon)}\bigr) - \mu(C) \ge -D_1^{(-\varepsilon)} - D_2^{(-\varepsilon)} - D_3^{(-\varepsilon)}. \qquad \square$$

APPENDIX

PROOF OF PROPOSITION 2.1. Consider the partition of $(0,1]$ by $J_n := (2^{-n-1}, 2^{-n}]$, $n \geq 0$, and define on Ω_1 the random variables

$$W_n := \sum_{i \geq 1} p_i \mathbf{1}_{p_i > 2^{-n}}, \qquad n \geq 1.$$

Fix $n \geq 1$. If two intervals are merged, then W_n can only increase; if some interval is split, then W_n can only decrease. We call the increment in the case of merging $\Delta_+ \geq 0$ and the loss in the case of splitting $\Delta_- \geq 0$. Given p, we can bound Δ_+ by

$$\begin{aligned}
\Delta_+ &\geq \sum_{i \neq j} p_i^2 p_j \mathbf{1}_{p_i \in J_n} \mathbf{1}_{p_j > 2^{-n-1}} \\
&= \left(\sum_i p_i^2 \mathbf{1}_{p_i \in J_n}\right)\left(\sum_j p_j \mathbf{1}_{p_j > 2^{-n-1}}\right) - \sum_i p_i^3 \mathbf{1}_{p_i \in J_n} \\
&\geq 2^{-2n-2}\left(\sum_i \mathbf{1}_{p_i \in J_n}\right)\left(\sum_j p_j \mathbf{1}_{p_j > 2^{-n-1}}\right) - 2^{-2n} \sum_i p_i \mathbf{1}_{p_i \in J_n},
\end{aligned}$$

and compute Δ_- as

$$\begin{aligned}
\Delta_- &= \sum_i p_i^2 \int_0^1 x p_i \mathbf{1}_{x p_i \leq 2^{-n} < p_i} + (1-x) p_i \mathbf{1}_{(1-x) p_i \leq 2^{-n} < p_i} \, dx \\
&= 2 \sum_i p_i^3 \mathbf{1}_{2^{-n} < p_i} \int_0^1 x \mathbf{1}_{x \leq 2^{-n}/p_i} \, dx = 2 \sum_i p_i^3 \mathbf{1}_{2^{-n} < p_i} \left[\frac{x^2}{2}\right]_0^{2^{-n}/p_i} \\
&= 2^{-2n} \sum_i p_i \mathbf{1}_{2^{-n} < p_i}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\Delta_+ - \Delta_- &\geq 2^{-2n}\left(\tfrac{1}{4}\left(\sum_i \mathbf{1}_{p_i \in J_n}\right)\left(\sum_j p_j \mathbf{1}_{p_j > 2^{-n-1}}\right) - \sum_i p_i \mathbf{1}_{2^{-n-1} < p_i}\right) \\
&\geq 2^{-2n}\left(\tfrac{1}{4}\left(\sum_i \mathbf{1}_{p_i \in J_n}\right)\left(\sum_j p_j \mathbf{1}_{p_j > 2^{-n-1}}\right) - 1\right).
\end{aligned}$$

Since $\int \Delta_+ - \Delta_- \, d\mu = 0$ due to stationarity this implies

$$\begin{aligned}
4 &\geq \int \left(\sum_i \mathbf{1}_{p_i \in J_n}\right)\left(\sum_j p_j \mathbf{1}_{p_j > 2^{-n-1}}\right) d\mu \\
&\geq 2^{-n-1} \int \left(\sum_i \mathbf{1}_{p_i \in J_n}\right)^2 d\mu \geq 2^{-n-1}\left(\int \sum_i \mathbf{1}_{p_i \in J_n} \, d\mu\right)^2.
\end{aligned} \tag{A.1}$$

Since this holds for any $n \geq 1$ we get

$$\int \sum_i \mathbf{1}_{p_i \in J_n} \, d\mu = O(2^{\beta n}), \qquad n \to \infty \tag{A.2}$$

with $\beta = 1/2$. Therefore, for any $\alpha > \beta = 1/2$,

$$\int \sum_i p_i^\alpha \, d\mu \leq \sum_{n \geq 0} 2^{-\alpha n} \int \sum_i \mathbf{1}_{p_i \in J_n} \, d\mu \leq c \sum_{n \geq 0} 2^{n(\beta - \alpha)} < \infty \tag{A.3}$$

for some constant $c > 0$, thus proving (2.1) for $\alpha > 1/2$. We shall now use this result to extend it to all $2/5 < \alpha < 1$, as required. To this end, observe that we have due to (A.1), for arbitrary $0 < \beta < 1$,

$$\begin{aligned}
4 &\geq \int \left(\sum_i \mathbf{1}_{p_i \in J_n} \right) \left(\sum_j p_j \mathbf{1}_{p_j > 2^{-n-1}} \right) \mathbf{1}_{\{\sum_j p_j \mathbf{1}_{p_j > 2^{-n-1}} > 2^{-n\beta}\}} \, d\mu \\
&\geq 2^{-n\beta} \int \sum_i \mathbf{1}_{p_i \in J_n} \mathbf{1}_{\{\sum_j p_j \mathbf{1}_{p_j > 2^{-n-1}} > 2^{-n\beta}\}} \, d\mu \\
&= 2^{-n\beta} \left(\int \sum_i \mathbf{1}_{p_i \in J_n} \, d\mu - \int \sum_i \mathbf{1}_{p_i \in J_n} \mathbf{1}_{\{\sum_j p_j \mathbf{1}_{p_j > 2^{-n-1}} \leq 2^{-n\beta}\}} \, d\mu \right) \qquad \text{(A.4)} \\
&\geq 2^{-n\beta} \left(\int \sum_i \mathbf{1}_{p_i \in J_n} \, d\mu - \frac{2^{-n\beta}}{2^{-n-1}} \mu \left[\sum_j p_j \mathbf{1}_{p_j > 2^{-n-1}} \leq 2^{-n\beta} \right] \right) \\
&\geq 2^{-n\beta} \left(\int \sum_i \mathbf{1}_{p_i \in J_n} \, d\mu - 2^{-n\beta + n + 1} \mu[\forall i : p_i \leq 2^{-n\beta}] \right).
\end{aligned}$$

To bound the μ-probability in the last expression we recall from (A.3) that, for $1/2 < \alpha < 1$,

$$\infty > \int \sum_i p_i^\alpha \, d\mu \geq \int \left(\sum_i p_i^\alpha \right) \mathbf{1}_{\{\forall i : p_i \leq 2^{-n\beta}\}} \, d\mu \tag{A.5}$$

for any $n \geq 0$. On the event $\{\forall i : p_i \leq 2^{-n\beta}\}$, by Jensen's inequality,

$$\sum_i p_i^\alpha = \sum_i p_i \cdot p_i^{\alpha - 1} \geq \left(\sum_i p_i \cdot p_i \right)^{\alpha - 1} \geq 2^{n\beta(1-\alpha)}.$$

Therefore, for all $1/2 < \alpha < 1$ due to (A.5), $\mu[\forall i : p_i \leq 2^{-n\beta}] = O(2^{-n\beta(1-\alpha)})$ as $n \to \infty$. Substituting this into (A.4) we get that, for any $1/2 < \alpha < 1$,

$$\int \sum_i \mathbf{1}_{p_i \in J_n} \, d\mu = O\big(2^{n\beta} + 2^{(-n\beta + n) - n\beta(1-\alpha)}\big) = O\big(2^{n \max\{\beta, 1 - \beta(2-\alpha)\}}\big).$$

The choice $\beta = (3-\alpha)^{-1}$ minimizes $\max\{\beta, 1-\beta(2-\alpha)\}$ and therefore yields that (A.2) and consequently also (A.3) and (2.1) hold for any α, $\beta > (3 - 1/2)^{-1} = 2/5$. $\square$

REMARK. A posteriori, once it has been established that μ must be the Poisson–Dirichlet law, Proposition 2.1 holds for all $\alpha > 0$ since by [11], equation (20)

$$\int \sum_{i\geq 1} p_i^\alpha \, d\,\widehat{\mu}_1 = \int_0^1 x^{\alpha-1}\, dx = \frac{1}{\alpha}.$$

NOTE ADDED IN PROOF. In [14], the question was raised as to whether the state $s := (1, 0, \dots) \in \Omega_1$ is recurrent for the CCF. Our techniques allow one to respond affirmatively to this question. Indeed, let $X^{(n)}(k)$ denote the state, at time k, of a DCF$^{(n)}$ initialized at $X^{(n)}(0) = (n, 0, \dots) =: s_n$. The recurrence of s for the CCF then follows, by the coupling introduced in Theorem 3.1, from the existence of a constant C independent of n and $k < n$ such that $P^{(n)}(X^{(n)}(2k) = s_n) \geq C/k$. To see the last estimate, note that by the character decomposition at the beginning of the proof of Theorem 4.1, it holds that

$$\begin{aligned} P^{(n)}\big(X^{(n)}(2k) = s_n\big) &= \sum_{\lambda\in\mathcal{P}} \theta_\lambda^{(n)^{2k}} \big\langle \big(\pi_S^{(n)}(s_n)\big)^{-1} \mathbf{1}_{s_n,\chi_\lambda}\big\rangle\big\langle \mathbf{1}_{s_n,\chi_\lambda}\big\rangle \\ &= \sum_{\lambda\in\mathcal{P}_n} \theta_\lambda^{(n)^{2k}} \pi_S^{(n)}(s_n)\chi_\lambda^2(s_n). \end{aligned}$$

From (1.1), $\pi_S^{(n)}(s_n) = 1/n$. By the Murnaghan–Nakayama rule, $\chi_\lambda(s_n) = 0$ unless $\lambda = (i, 1, 1, \dots)$ for some $i \in \{1, \dots, n\}$, in which case $|\chi_\lambda(s_n)| = 1$. Using (4.6), one has that for such λ, $\theta_\lambda^{(n)} = (2i - n - 1)/(n - 1)$. Thus,

$$\begin{aligned} P^{(n)}\big(X^{(n)}(2k) = s_n\big) &= \frac{1}{n}\sum_{i=1}^{n}\left(1 - \frac{2i}{n-1}\right)^{2k} \\ &\geq \frac{1}{n}\sum_{i=1}^{n/4}\left(1 - \frac{2i}{n-1}\right)^{2k} \geq \frac{1}{2n}\sum_{i=1}^{n/4} e^{-4ki/(n-1)}, \end{aligned}$$

where we used that $(1-x) \geq e^{-x}$ for $x < 1/\sqrt{2}$. This yields the claim.

REFERENCES

[1] ARRATIA, R., BARBOUR, A. D. and TAVARÉ, S. (2004). Logarithmic combinatorial structures: A probabilistic approach. Unpublished manuscript.

[2] BROOKS, R. and MAKOVER, E. (1997). Random construction of Riemann surfaces. Preprint, Dept. Mathematics, Technion.

[3] BROOKS, R. and MAKOVER, E. (2001). Belyi surfaces. In *Entire Functions in Modern Analysis. Israel Math. Conf. Proc.* **15** 37–46. Bar Ilan Univ.

[4] DIACONIS, P. (1988). *Group Representations in Probability and Statistics*. IMS, Hayward, CA.

[5] DIACONIS, P. and HANLON, P. (1992). Eigen analysis for some examples of the Metropolis algorithm. In *Hypergeometric Functions on Domains of Positivity*, *Jack Polynomials*, *and Applications* (D. St. P. Richards, ed.) 99–117. AMS, Providence, RI.

[6] FLATTO, L., ODLYZKO, A. M. and WALES, D. B. (1985). Random shuffles and group representations. *Ann. Probab.* **13** 154–178.

[7] GAMBURD, A. and MAKOVER, E. (2002). On the genus of a random Riemann surface. In *Complex Manifolds and Hyperbolic Geometry* (C. J. Earle, W. J. Harvey and S. Recillas-Pishmash, eds.) 133–140. Amer. Math. Soc., Providence, RI.

[8] GNEDIN, A. and KEROV, S. (2001). A characterization of GEM distributions. *Combin. Probab. Comput.* **10** 213–217.

[9] GNEDIN, A. and KEROV, S. (2002). Fibonacci solitaire. *Random Struct. Alg.* **20** 71–88.

[10] JAMES, G. and KERBER, A. (1981). *The Representation Theory of the Symmetric Group*. Addison-Wesley, Reading, MA.

[11] KINGMAN, J. F. C. (1975). Random discrete distributions. *J. Roy. Statist. Soc. Ser. B* **37** 1–22.

[12] KINGMAN, J. F. C. (1977). The population structure associated with the Ewens sampling formula. *Theoret. Population Biol.* **11** 274–283.

[13] KINGMAN, J. F. C. (1993). *Poisson Processes*. Oxford Univ. Press.

[14] MAYER-WOLF, E., ZEITOUNI, O. and ZERNER, M. P. W. (2002). Asymptotics of certain coagulation–fragmentation processes and invariant Poisson–Dirichlet measures. *Electron. J. Probab.* **7** 1–25.

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

[16] TSILEVICH, N. V. (2000). Stationary random partitions of positive integers. *Theory Probab. Appl.* **44** 60–74.

[17] TSILEVICH, N. V. (2001). On the simplest split–merge operator on the infinite-dimensional simplex. PDMI Preprint 03/2001. Available at ftp://ftp.pdmi.ras.ru/pub/publicat/preprint/2001/03-01.ps.gz.

[18] VERSHIK, A. M. and SHMIDT, A. A. (1977). Limit theorems arising in the asymptotic theory of symmetric groups, I. *Theory Probab. Appl.* **22** 70–85.

P. DIACONIS
DEPARTMENT OF MATHEMATICS
AND DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305
USA

E. MAYER-WOLF
DEPARTMENT OF MATHEMATICS
TECHNION
HAIFA 32000
ISRAEL
E-MAIL: emw@tx.technion.ac.il

O. ZEITOUNI
DEPARTMENT OF MATHEMATICS
AND DEPARTMENT OF ELECTRICAL ENGINEERING
TECHNION
HAIFA 32000
ISRAEL
AND
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MINNESOTA 55455
USA
E-MAIL: zeitouni@math.umn.edu
URL: www-ee.technion.ac.il/~zeitouni

M. ZERNER
DEPARTMENT OF MATHEMATICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305
USA
E-MAIL: zerner@math.stanford.edu
URL: math.stanford.edu/~zerner