Electron. Commun. Probab. **25** (2020), no. 4, 1–12.
https://doi.org/10.1214/19-ECP283
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Characterising random partitions by random colouring

Jakob E. Björnberg* Cécile Mailler† Peter Mörters‡
Daniel Ueltschi§

Abstract

Let $(X_1, X_2, ...)$ be a random partition of the unit interval $[0,1]$, i.e. $X_i \geq 0$ and $\sum_{i\geq 1} X_i = 1$, and let $(\varepsilon_1, \varepsilon_2, ...)$ be i.i.d. Bernoulli random variables of parameter $p \in (0,1)$. The *Bernoulli convolution* of the partition is the random variable $Z = \sum_{i\geq 1} \varepsilon_i X_i$. The question addressed in this article is: Knowing the distribution of Z for some fixed $p \in (0,1)$, what can we infer about the random partition $(X_1, X_2, ...)$? We consider random partitions formed by residual allocation and prove that their distributions are fully characterised by their Bernoulli convolution if and only if the parameter p is not equal to $1/2$.

Keywords: partition structures; residual allocation.
AMS MSC 2010: 60E10; 60G57; 60K35.
Submitted to ECP on July 13, 2019, final version accepted on December 21, 2019.
Supersedes arXiv:1907.05960.

1 Introduction

Random partitions appear in the mathematical description of many natural systems, such as particle clustering and condensation in physics [3]; dynamics of gene populations in biology [8]; wealth distribution in economics [18]; etc. There is a vast amount of possible probability laws of random partitions, but one often encounters convergence to one of a few universal laws, most notably the Poisson–Dirichlet distribution with parameter $\theta > 0$, henceforth denoted $\mathtt{PD}(\theta)$ and defined below after Eq. (1.5).

To show convergence of a tight sequence of random partitions it is often feasible to show convergence of a derived quantity like the Bernoulli convolutions studied in this paper. If the limit of the derived quantity characterises the law of the underlying random partition among the class of possible limits, convergence is shown. It is therefore an important question whether the distribution of a random partition can be identified from its Bernoulli convolution, and in this paper we contribute to this problem.

We describe two scenarios that motivate this study in Sections 1.1 and 1.2. We introduce the precise setting and our results in Section 1.3 — the definition of the

*Department of Mathematics, Chalmers University of Technology and the University of Gothenburg, Sweden.
E-mail: jakob.bjornberg@gu.se
†Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom.
E-mail: c.mailler@bath.ac.uk
‡Mathematisches Institut, Universität zu Köln, Weyertal 86–90, 50931 Köln, Germany.
E-mail: moerters@math.uni-koeln.de
§Department of Mathematics, University of Warwick, Coventry, CV4 7AL, United Kingdom.
E-mail: daniel@ueltschi.org

Bernoulli convolution can be found around Eq. (1.4). Sections 2 and 3 contain the proofs of our two theorems. We make further comments in Section 4; it includes a counterexample due to A. Holroyd, that sheds much light on these questions.

1.1 Random interchange model and quantum spin systems

The random interchange model is a process on permutations constructed as products of random transpositions. Namely, given integers n and k, we pick k pairs of distinct integers $(x_1, y_1), \dots, (x_k, y_k)$ from $\{1, \dots, n\}$ uniformly at random, and consider the permutation

$$\sigma = \tau_k \circ \cdots \circ \tau_1. \tag{1.1}$$

Here, $\tau_i = (x_i, y_i)$ denotes the transposition of x_i and y_i. The cycle structure (i.e. the lengths of the permutation cycles) of σ gives an integer partition of n; dividing by n gives a partition of $[0, 1]$.

Schramm [17] studied this model in the case where $k = \lfloor cn \rfloor$ with $c > 1$. He proved that, with high probability as $n \to \infty$, there are cycles whose lengths are of order n. Let L_i denote the length of the ith largest cycle. The sum of cycles of length of order n is $\kappa n(1 + o(1))$ with $\kappa = \kappa(c)$ fixed (and $\kappa \to 1$ when $c \to \infty$); and the sequence $(\frac{L_1}{\kappa n}, \frac{L_2}{\kappa n}, \dots)$ converges (weakly) to $\mathtt{PD}(1)$, the Poisson–Dirichlet distribution with parameter 1.

One motivation for the random interchange model, pointed out and exploited by Tóth [20], is that it provides a probabilistic representation of the Heisenberg model of quantum spins. For this representation the density of the random interchange model gets an extra weight $2^{\#\text{cycles}}$, which leads to a conjectured limit which is the Poisson–Dirichlet distribution $\mathtt{PD}(2)$, see [10]. In this case the number of transpositions k is random, chosen to be $\mathtt{Poisson}(cn)$. Recently, it was proved in [5] that, in the model with weight $\theta^{\#\text{cycles}}$, $\theta = 2, 3, 4, \dots$, we have

$$\lim_{n\to\infty} \mathbb{E}_n \Big[\prod_{i\geq 1} \tfrac{1}{\theta} \big(\mathrm{e}^{hL_i/n} + \theta - 1 \big) \Big] = \mathrm{e}^{\frac{h}{\theta}(1-\kappa)}\, \mathbb{E}_{\mathtt{PD}(\theta)} \Big[\prod_{i\geq 1} \tfrac{1}{\theta} \big(\mathrm{e}^{h\kappa X_i} + \theta - 1 \big) \Big], \tag{1.2}$$

for some (deterministic) $\kappa \in [0, 1]$ which depends on c and θ and is positive for c large enough; the above identity holds for all $h \in \mathbb{C}$. The last expectation in (1.2) is equal to the moment generating function at $h\kappa$ of the Bernoulli convolution of $\mathtt{PD}(\theta)$ with parameter $p = 1/\theta$. The interpretation is that the system displays small (order 1) and large (order n) cycles, and that the joint distribution of the lengths of large cycles is $\mathtt{PD}(\theta)$; see [5] for more details. But is Eq. (1.2) enough to guarantee that the limiting sequence of renormalised cycle lengths be equal to $\mathtt{PD}(\theta)$? We prove here that, among the residual allocation distributions, the answer is yes for $\theta = 3, 4, \dots$, but no for $\theta = 2$.

There are related loop models that include 'double bars' as well as the transposition 'crosses', that represent further quantum spin systems [1, 21]. Without weights, it was proved in [6] that the joint distribution of the lengths of long loops is $\mathtt{PD}(1/2)$. With weights $2^{\#\text{loops}}$, the result of [5] is that

$$\lim_{n\to\infty} \mathbb{E}_n \Big[\prod_{i\geq 1} \cos(hL_i/n) \Big] = \mathbb{E}_{\mathtt{PD}(1)} \Big[\prod_{i\geq 1} \cos(h\kappa X_i) \Big], \tag{1.3}$$

for all $h \in \mathbb{C}$. The latter expectation is closely related to the moment generating function of the Bernoulli convolution of $\mathtt{PD}(1)$ with parameter $p = 1/2$. Results of the present article show that the above claim is not enough to guarantee that the limiting distribution is $\mathtt{PD}(1)$, even if one assumes that the limiting distribution is a residual allocation.

1.2 Exchangeable divide-and-color models

In a recent paper by Steif and Tykesson [19], the authors introduce *generalized divide-and-color models* as follows. Given a countable set S and $p \in (0, 1)$, one starts by forming a random partition Π of S according to some rule; one then assigns to each part of Π a 'color' 0 or 1, independently and with probability p for 1. Letting each element of S take the color of the part it belongs to and then forgetting about the original partition Π, one ends up with a random element $\omega \in \{0,1\}^S$. This construction is motivated by the Fortuin–Kasteleyn representation of the Ising model, among other examples.

A particular case is when $S = \mathbb{N}$ and when the random partition Π is *exchangeable*, i.e. its distribution is invariant under all finite permutations of $\mathbb{N}$. By Kingman's famous theorem [12], such a random partition of $\mathbb{N}$ is uniquely encoded by a random vector $(X_i)_{i\geq 1}$ satisfying $X_i \geq X_{i+1} \geq 0$ for all $i \geq 1$ and $\sum_{i\geq 1} X_i \leq 1$; note that < 1 is allowed in this case. On the other hand, the resulting color process $\omega \in \{0,1\}^{\mathbb{N}}$ is also exchangeable; by de Finetti's theorem, this means that there is some random variable $\xi \in [0,1]$ such that, conditional on ξ, the ω_i are i.i.d. $\texttt{Bernoulli}(\xi)$. It is not hard to see that (when $\sum_{i\geq 1} X_i = 1$) ξ equals the Bernoulli convolution of $(X_i)_{i\geq 1}$, see [19, Lemma 3.12]. Steif and Tykesson ask whether the law of the random partition Π can be recovered from the law of ω when $p \neq 1/2$. This is equivalent to asking whether the law of $(X_i)_{i\geq 1}$ can be recovered from the law of its Bernoulli convolution. Our results on residual allocation models show that the answer can be *yes* under additional assumptions on $(X_i)_{i\geq 1}$.

1.3 Framework and results

We define a Bernoulli convolution as follows.

Definition 1.1 (Bernoulli convolution)**.** *Let $(X_i)_{i\geq 1}$ be a random partition of $[0,1]$, i.e. $X_i \geq 0$ for all $i \geq 1$ and $\sum_{i\geq 1} X_i = 1$. Let $(\varepsilon_i)_{i\geq 1}$ be a sequence of i.i.d. Bernoulli random variables of parameter $p \in (0,1)$, independent of $(X_i)_{i\geq 1}$. Set*

$$Z = \sum_{i\geq 1} \varepsilon_i X_i. \tag{1.4}$$

The law of Z, and sometimes the random variable Z itself, is called the Bernoulli(p) convolution of the random partition $(X_i)_{i\geq 1}$.

We restrict our setting to random partitions obtained from *residual allocation*. Namely, we consider the interval $[0,1]$ with the Borel σ-algebra. Given a probability measure μ on $[0,1]$, let $(Y_i)_{i\geq 1}$ be i.i.d. random variables distributed according to μ, and consider the sequence $(X_i)_{i\geq 1}$ defined by

$$\begin{aligned} &X_1 = Y_1, \\ &X_2 = (1-Y_1)Y_2, \\ &X_3 = (1-Y_1)(1-Y_2)Y_3, \\ &\text{etc...} \end{aligned} \tag{1.5}$$

Assuming that $\mu(\{0\}) < 1$, it is not hard to prove that $X_i \to 0$ as $i \to \infty$ and that $\sum_{i\geq 1} X_i = 1$, almost surely. It is possible to rearrange the sequence $(X_i)_{i\geq 1}$ in decreasing order if one wants an ordered partition, but this is not necessary here.

An important example of this construction is the Griffiths, Engen and McCloskey distribution, $\texttt{GEM}(\theta)$, obtained when $\mu = \texttt{Beta}(1,\theta)$. If one orders the entries of a $\texttt{GEM}(\theta)$ sample by decreasing size, one obtains the famous Poisson–Dirichlet distribution $\texttt{PD}(\theta)$, see [11]. Another important example is the 'classical' Bernoulli convolution $\sum_{i\geq 1} \pm\lambda^i$ with i.i.d. random signs; see the review [13]. This falls into our framework (take $\mu = \delta_{1-\lambda}$ for some fixed $\lambda \in (0,1)$ so that $X_i = (1-\lambda)\lambda^{i-1}$), except that our Bernoulli coefficients take value in $\{0,1\}$ instead of $\{-1,1\}$.

As a shorthand, since we only consider random partitions from residual allocation, we will sometimes refer to Z (or its law) as the Bernoulli convolution of the measure μ. The

Bernoulli convolution is invariant under rearrangements of the sequence $(X_i)_{i\geq 1}$. The cases $p=0$ and $p=1$ are trivial and uninteresting, since $Z=0$ and $Z=1$, respectively.

If μ has an atom at 0 of value $c>0$, i.e. $\mu(\{0\})=c$, then the sequence $(Y_1,Y_2,\dots)$ — and therefore $(X_1,X_2,\dots)$ — contains a density c of elements that are equal to 0; this does not affect Z. In other words, the Bernoulli convolutions of μ and $c\delta_0+(1-c)\mu$ are the same for all $c\in[0,1)$. We avoid this trivial degeneracy by restricting our attention to measures that do not have an atom at 0.

Given $p\in(0,1)$, the question is whether the Bernoulli(p) convolution characterises the random partition obtained from residual allocation. We show that it is the case for $p\neq {}^1\!/_2$.

Theorem 1.2 (Uniqueness for $p\neq {}^1\!/_2$)**.** *Let $p\in(0,1)\setminus\{{}^1\!/_2\}$. If μ and ν are two probability measures on $[0,1]$ such that $\mu(\{0\})=\nu(\{0\})=0$, and the corresponding residual allocation models have identical Bernoulli(p) convolution, then $\mu=\nu$.*

We also show that Theorem 1.2 fails for $p={}^1\!/_2$. Our non-uniqueness results hold for GEM (or Poisson–Dirichlet) measures of arbitrary parameters.

Theorem 1.3 (Non-uniqueness for $p={}^1\!/_2$)**.** *Let $\theta>0$ and $\mu=\mathtt{Beta}(1,\theta)$. Then there exist infinitely many $\nu\neq\mu$ such that $\nu(\{0\})=0$, and such that μ and ν have identical Bernoulli(${}^1\!/_2$) convolutions.*

The non-uniqueness results are not explicit with the exception of $\mathtt{GEM}(2)$: We show that if an (absolutely continuous) measure ν satisfies

$$x\,\mathrm{d}\nu(x)=(1-x)\,\mathrm{d}\nu(1-x)\quad \text{for all } x\in[0,1], \tag{1.6}$$

then its residual allocation has the same Bernoulli convolution as $\mu=\mathtt{Beta}(1,2)$. Note that (1.6) holds true in the case $\mu=\mathtt{Beta}(1,2)$, for which $\mathrm{d}\mu(x)=2(1-x)\mathrm{d}x$. Another example is the Dirac measure at $x={}^1\!/_2$, $\nu=\delta_{1/2}$, which formally satisfies (1.6). We refer to Proposition 3.4 for details including conditions on the regularity of measures.

We prove Theorems 1.2 and 1.3 with the help of a stochastic identity for the random variable Z, see Lemma 2.1. This identity holds because of the self-similarity structure of residual allocations. The proofs of Theorem 1.2 and 1.3 can be found in Sections 2 and 3, respectively.

A natural question is whether Theorem 1.2 holds beyond residual allocations. Obviously, the Bernoulli convolution (1.4) may be defined for arbitrary random partitions $(X_i)_{i\geq 1}$. Alexander Holroyd has given an example showing that, in general, the Bernoulli convolution does *not* determine the random partition, even if the former is known for all $p\in(0,1)$; we explain Holroyd's example in Section 4. One may also allow more general random variables $(\varepsilon_i)_{i\geq 1}$; in this generality, Z is sometimes called a *random weighted average*. Pitman's recent review [14] contains a wealth of information about the theory of random weighted averages. In [14, Corollary 9] it is shown that the distributions of the random weighted averages Z, as $(\varepsilon_i)_{i\geq 1}$ range over all i.i.d. sequences of random variables with finite support, fully characterize the law of the random partition $(X_i)_{i\geq 1}$. This holds without any assumptions about the properties of the random partition. It is natural to ask whether the condition on the ε_i can be weakened.

2 Uniqueness when $p\neq {}^1\!/_2$ (proof of Theorem 1.2)

The following lemma will be used both to establish uniqueness for $p\neq {}^1\!/_2$ and non-uniqueness for $p={}^1\!/_2$. It is not new, see [9, Theorem 1] or [7, Theorem 7.1]; it is also discussed in [14, (119)].

Lemma 2.1 (Stochastic identity)**.** *Let $Y,Y_1,Y_2,\dots$ be i.i.d. random variables with values in $[0,1]$ and $(X_i)_{i\geq 1}$ defined by* (1.5)*; $\varepsilon,\varepsilon_1,\varepsilon_2,\dots$ be i.i.d.* $\mathtt{Bernoulli}(p)$ *random variables* *independent of the Y's; and Z and Z' be two identically distributed random variables with values in $[0,1]$, Z' being independent of Y and ε. The following stochastic identities are equivalent:*

(a) $Z \overset{\mathrm{d}}{=} \sum_{i\geq 1} \varepsilon_i X_i$;

(b) $Z \overset{\mathrm{d}}{=} \varepsilon Y + (1-Y)Z'$.

Proof. Assuming (a), we have

$$Z \overset{\mathrm{d}}{=} \varepsilon_1 Y_1 + (1-Y_1)\sum_{i\geq 2} \varepsilon_i \frac{X_i}{1-Y_1},$$

where the sequence $(X_i/(1-Y_1))_{i\geq 2}$ is independent of $X_1 = Y_1$ and has the same distribution as $(X_i)_{i\geq 1}$, which gives (b).

Assuming (b), we construct a sequence of random variables which all have the same distribution as Z and which converge weakly (in fact, almost surely) to $\sum_{i\geq 1}\varepsilon_i X_i$. Observe that there exist Z_1 and Z_2 two independent copies of Z, independent of ε_i and Y_i such that

$$\begin{aligned} Z_1 &\overset{\mathrm{d}}{=} \varepsilon_1 Y_1 + (1-Y_1)Z_1 \\ &\overset{\mathrm{d}}{=} \varepsilon_1 Y_1 + (1-Y_1)\big[\varepsilon_2 Y_2 + (1-Y_2)Z_2\big]. \end{aligned} \tag{2.1}$$

Iterating this further, we get $(Z_i)_{i\geq 1}$ such that for all $n \geq 1$,

$$\sum_{i=1}^{n} \varepsilon_i X_i + (1-Y_1)\cdots(1-Y_n)Z_n \overset{\mathrm{d}}{=} Z_1, \tag{2.2}$$

where X_i are as defined in (1.5). All terms in $\sum_{i=1}^{n}\varepsilon_i X_i$ are positive and the sums are bounded by 1, hence the series converges to $\sum_{i\geq 1}\varepsilon_i X_i$; the remainder $(1-Y_1)\cdots(1-Y_n)Z_n$ converges to 0 almost surely. As $n\to\infty$ we obtain (a). $\square$

We will show that all moments of $Y \sim \mu$ are determined by the Bernoulli convolution Z of the residual allocation model from μ. This holds for $p \in (0,1)\setminus\{1/2\}$. It does not hold for $p = 0$ (the Bernoulli convolution is always 0) and $p = 1$ (it is always 1). It also does not hold for $p = 1/2$, for reasons that are not obvious and that are discussed in Sect. 3.

Let us introduce numbers $a_{n,k}$ and c_n that depend on the law of Z, and numbers b_n that depend on the law of Y. For $n,k \in \mathbb{N}$ with $k \leq n$, let

$$\begin{aligned} a_{n,k} &= (-1)^k p \binom{n}{k} \mathbb{E}\big[(1-Z)^k\big], \\ c_n &= (1-p)\mathbb{E}[Z^n], \\ b_n &= \frac{1-\mathbb{E}\big[(1-Y)^n\big]}{\mathbb{E}[Y]}. \end{aligned} \tag{2.3}$$

Note that $b_0 = 0$, $b_1 = 1$, and $a_{1,1} + c_1 = 0$ since $\mathbb{E}[Z] = p$. We have the following relations.

Proposition 2.2 (Recurrence relation)**.** *For all $p \in [0,1]$ and all $n \geq 1$, we have*

$$c_n b_n + \sum_{k=1}^{n} a_{n,k} b_k = 0.$$

Proof. We expand $\mathbb{E}[Z^n]$ in two different ways. First,

$$\begin{aligned}\mathbb{E}[Z^n] &= (1-p)\mathbb{E}[Z^n] + p\mathbb{E}\big[\big(1-(1-Z)\big)^n\big] \\ &= (1-p)\mathbb{E}[Z^n] + p\sum_{k=0}^{n}(-1)^k\binom{n}{k}\mathbb{E}\big[(1-Z)^k\big].\end{aligned} \tag{2.4}$$

Second, using Lemma 2.1,

$$\begin{aligned}\mathbb{E}[Z^n] = \mathbb{E}\big[\big(\varepsilon Y + (1-Y)Z\big)^n\big] &= p\mathbb{E}\big[\big(Y+(1-Y)Z\big)^n\big] + (1-p)\mathbb{E}\big[\big((1-Y)Z\big)^n\big] \\ &= (1-p)\mathbb{E}\big[(1-Y)^n\big]\mathbb{E}[Z^n] + p\mathbb{E}\big[\big(1-(1-Y)(1-Z)\big)^n\big] \\ &= (1-p)\mathbb{E}\big[(1-Y)^n\big]\mathbb{E}[Z^n] + p\sum_{k=0}^{n}(-1)^k\binom{n}{k}\mathbb{E}\big[(1-Y)^k\big]\mathbb{E}\big[(1-Z)^k\big].\end{aligned} \tag{2.5}$$

Equating these identities, we get

$$0 = (1-p)\mathbb{E}[Z^n]\big\{1-\mathbb{E}\big[(1-Y)^n\big]\big\} + p\sum_{k=0}^{n}(-1)^k\binom{n}{k}\mathbb{E}\big[(1-Z)^k\big]\big\{1-\mathbb{E}\big[(1-Y)^k\big]\big\}. \tag{2.6}$$

We now divide by $\mathbb{E}[Y]$ and we obtain the claim of the proposition. □

The next lemma holds for $p \neq 1/2$ only.

Lemma 2.3 (Non-zero coefficient)**.** *For $p \in (0,1) \setminus \{1/2\}$, we have for all $n \geq 2$ that*

$$a_{n,n} + c_n \neq 0.$$

Proof. We have

$$a_{n,n} + c_n = (1-p)\mathbb{E}[Z^n] + (-1)^n p\mathbb{E}\big[(1-Z)^n\big]. \tag{2.7}$$

This is always positive for n even; we thus assume from now on that $n \geq 3$ is odd. From the definitions (1.5) and (1.4), we have

$$\mathbb{E}[Z^n] = \sum_{i_1,i_2,\ldots,i_n \geq 1}\mathbb{E}\big[\varepsilon_{i_1}\varepsilon_{i_2}\cdots\varepsilon_{i_n}\big]\mathbb{E}\big[X_{i_1}X_{i_2}\cdots X_{i_n}\big]. \tag{2.8}$$

Note that, if $\ell = \#\{i_1,\ldots,i_n\}$ denotes the number distinct indices among $i_1,\ldots,i_n \geq 1$, then

$$\mathbb{E}\big[\varepsilon_{i_1}\varepsilon_{i_2}\cdots\varepsilon_{i_n}\big] = p^\ell, \tag{2.9}$$

since $\varepsilon_i^k = \varepsilon_i$ for all $k, i \geq 1$. We thus get

$$\mathbb{E}[Z^n] = \sum_{\ell=1}^{n} p^\ell\mathbb{E}[S_{n,\ell}], \tag{2.10}$$

where $S_{n,\ell} = \sum X_{i_1}X_{i_2}\cdots X_{i_n}$ summed over all choices of indices $i_1,\ldots,i_n \geq 1$ such that $\#\{i_1,\ldots,i_n\} = \ell$. Note that $\mathbb{E}[S_{n,\ell}] > 0$ for all $\ell \geq 1$. Since, by definition, $1-Z = \sum_{i\geq 1}(1-\varepsilon_i)X_i$ we also have $\mathbb{E}[(1-Z)^n] = \sum_{\ell=1}^{n}(1-p)^\ell\mathbb{E}[S_{n,\ell}]$, and thus

$$a_{n,n} + c_n = p(1-p)\,\mathbb{E}[S_{n,\ell}]\sum_{\ell=1}^{n}\big(p^{\ell-1} - (1-p)^{\ell-1}\big). \tag{2.11}$$

While the term $\ell = 1$ is zero, all other terms are non-zero and have the same sign, which proves the claim since $n > 1$. □

We now turn to the proof of Theorem 1.2.

Proof of Theorem 1.2. It follows from Proposition 2.2 and Lemma 2.3 that, for $n \geq 2$,

$$b_n = -(a_{n,n} + c_n)^{-1} \sum_{k=1}^{n-1} a_{n,k} b_k. \tag{2.12}$$

Recall that $b_0 = 0$, $b_1 = 1$. The above equation shows that the b_n's are recursively determined by the $a_{n,k}$'s and c_n's, which only depend on the Bernoulli convolution Z. As $n \to \infty$, the sequence (b_n) converges to $1/\mathbb{E}[Y]$ — here we use our assumption that the measure μ does not have an atom at 0. It follows that $\mathbb{E}[Y]$ and $\mathbb{E}[(1-Y)^n]$ are determined by the Bernoulli convolution for all n. Then all moments of the original measure μ are known, hence the measure μ itself (see [4, Theorem 1.2]). □

3 Non-uniqueness when $p = 1/2$ (proof of Theorem 1.3)

In this section we set $p = 1/2$, unless indicated otherwise. We also assume that the Bernoulli convolution of parameter $1/2$ has a density $q(x)$ with respect to Lebesgue measure, and that $q(x) > 0$ for all $x \in (0,1)$. This will hold in particular in the case of $\mathtt{GEM}(\theta)$. Since $p = 1/2$ we then have that $q(x) = q(1-x)$ because $Z \overset{\mathrm{d}}{=} 1 - Z$.

Given a nonnegative measurable function ρ on $[0,1]$, we define the function $H\rho$ by

$$[H\rho](x) = \frac{1}{q(x)} \int_0^x q\Big(\frac{x-u}{1-u}\Big) \frac{\rho(u)}{1-u} \mathrm{d}u. \tag{3.1}$$

Let $\mathcal{R}_q$ be the cone of nonnegative measurable functions ρ such that the integral above is finite for all $0 \leq x \leq 1$. H is a linear operator on $\mathcal{R}_q$. As it turns out, it gives a relation between the density ρ of a probability measure on $[0,1]$, and the density q of the corresponding Bernoulli convolution. This may be seen by expanding the stochastic identity of Lemma 2.1 (b) and making a suitable change of variables. More precisely, we have:

Lemma 3.1 (Condition for non-uniqueness)**.** *Let q be a probability density function on $[0,1]$ such that $q(x) > 0$ on $(0,1)$ and $q(x) = q(1-x)$. Let $\rho \in \mathcal{R}_q$; we have*

$$[H\rho](x) + [H\rho](1-x) = 2, \quad \textit{for almost all } x \in [0,1], \tag{3.2}$$

if and only if

(a) ρ is a probability density function on $[0,1]$, and

(b) the Bernoulli($1/2$) convolution of the residual allocation model from ρ has density q.

Proof. Assume that (3.2) holds. For (a), we have, writing $h(x) = [H\,\rho](x)$,

$$\begin{aligned} 1 &= \int_0^1 q(x) \tfrac{h(x)+h(1-x)}{2} \,\mathrm{d}x = \int_0^1 q(x)h(x)\,\mathrm{d}x = \int_0^1 \mathrm{d}u\,\rho(u) \int_u^1 \mathrm{d}z \tfrac{1}{1-u} q\big(\tfrac{z-u}{1-u}\big) \\ &= \int_0^1 \mathrm{d}u\,\rho(u) \int_0^1 \mathrm{d}v\, q(v) = \int_0^1 \mathrm{d}u\,\rho(u), \end{aligned} \tag{3.3}$$

as claimed. (We used the change of variables $v = \frac{z-u}{1-u}$.)

For (b), we use (3.2) to get

$$q(x) = \tfrac{1}{2} \int_0^x q\Big(\frac{x-u}{1-u}\Big) \frac{\rho(u)}{1-u} \mathrm{d}u + \tfrac{1}{2} \int_0^{1-x} q\Big(\frac{x}{1-u}\Big) \frac{\rho(u)}{1-u} \mathrm{d}u. \tag{3.4}$$

It follows that for all continuous function f, we have

$$\begin{aligned}\int_0^1 & q(x)f(x)\,\mathrm{d}x \\ &= \tfrac{1}{2}\int_0^1 \frac{\rho(u)}{1-u}\mathrm{d}u\int_u^1 q\Big(\frac{x-u}{1-u}\Big)f(x)\mathrm{d}x + \tfrac{1}{2}\int_0^1 \frac{\rho(u)}{1-u}\mathrm{d}u\int_0^{1-u} q\Big(\frac{x}{1-u}\Big)f(x)\mathrm{d}x \\ &= \tfrac{1}{2}\int_0^1 \rho(u)\mathrm{d}u\int_0^1 f\big(u+(1-u)y\big)q(y)\mathrm{d}y + \tfrac{1}{2}\int_0^1 \rho(u)\mathrm{d}u\int_0^1 f\big((1-u)y\big)q(y)\mathrm{d}y.\end{aligned} \tag{3.5}$$

We used Fubini's theorem to get the second line, and the changes of variables $y = \frac{x-u}{1-u}$ and $y = \frac{x}{1-u}$ (for fixed u) to get the third line. The left side gives the expectation $\mathbb{E}[f(Z)]$ for the random variable with density q. The right side gives $\mathbb{E}[f(\varepsilon Y + (1-Y)Z)]$ for the independent random variables $\varepsilon \sim \mathtt{Bernoulli}(\frac{1}{2})$, Y with density ρ, and Z with density q. We recognise the stochastic identity of Lemma 2.1 (b). Hence q is the density of the Bernoulli convolution of ρ.

The other implication can be checked similarly: (3.5) holds by (b), hence also (3.4) for almost all x, which gives (3.2). □

The next step is to identify the Bernoulli convolution of GEM distributions. It turns out to be equal to Beta random variables. We consider general parameters p, although we only need the case $p = 1/2$ here.

Proposition 3.2 (Bernoulli convolution of GEM)**.** *Let $\theta > 0$ and $p \in [0,1]$. Then the Bernoulli convolution of $\mathtt{GEM}(\theta)$, i.e. of the residual allocation model from $\mathtt{Beta}(1,\theta)$ random variables, is the $\mathtt{Beta}(p\theta,(1-p)\theta)$ distribution.*

This result is not new, see e.g. [14, Prop. 27(iii)]. We sketch a proof using the connection between $\mathtt{GEM}(\theta)$ and $\mathtt{PD}(\theta)$, Kingman's characterization of $\mathtt{PD}(\theta)$ in terms of the Gamma-subordinator, as well as the following well-known lemma (see e.g. [10, Lemma 7.4]):

Lemma 3.3 (Beta–Gamma calculus)**.** *If Y_1 and Y_2 are independent, with respective distributions $\mathtt{Gamma}(\theta_1,1)$ and $\mathtt{Gamma}(\theta_2,1)$, then*

1. *Y_1+Y_2 has distribution $\mathtt{Gamma}(\theta_1+\theta_2,1)$,*
2. *$Y_1/(Y_1+Y_2)$ has distribution $\mathtt{Beta}(\theta_1,\theta_2)$,*
3. *Y_1+Y_2 and $Y_1/(Y_1+Y_2)$ are independent.*

Proof of Proposition 3.2. Let $\xi = (\xi_1,\xi_2,\dots)$ be the points of a Poisson process with intensity measure $\theta x^{-1}\mathrm{e}^{-x}\,\mathrm{d}x$ on $(0,\infty)$ in decreasing order. Let $S = \sum_{i\ge1}\xi_i$ and $X_i = \xi_i/S$ for all $i \ge 1$, then $S \sim \mathtt{Gamma}(\theta,1)$ and $X = (X_1,X_2,\dots)$ is $\mathtt{PD}(\theta)$-distributed (see [2, Definition 2.5]). Let $(\varepsilon_i)_{i\ge1}$ be a sequence of i.i.d. $\mathtt{Bernoulli}(p)$ random variables. Let $\xi^{(1)}$ be the collection $(\xi_i\colon \varepsilon_i = 1)$ and $\xi^{(0)}$ its complement $(\xi_i\colon \varepsilon_i = 0)$. Note that $\xi^{(1)}$ and $\xi^{(0)}$ are independent Poisson processes with respective intensity measures $p\theta x^{-1}\mathrm{e}^{-x}\,\mathrm{d}x$ and $(1-p)\theta x^{-1}\mathrm{e}^{-x}\,\mathrm{d}x$ on $(0,\infty)$. Set $Y_1 = \sum_{i\ge1}\xi_i^{(1)}$ and $Y_0 = \sum_{i\ge1}\xi_i^{(0)}$. Then Y_1 and Y_0 have distributions $\mathtt{Gamma}(p\theta,1)$ and $\mathtt{Gamma}((1-p)\theta,1)$ respectively (this can be checked using the Laplace transform and Campbell's formula as in [10, Lemma 7.3]). Since $Z = Y_1/(Y_0+Y_1)$, Lemma 3.3 implies that $Z \sim \mathtt{Beta}(p\theta,(1-p)\theta)$, which concludes the proof. □

We now consider a special case of Theorem 1.3, namely $\theta = 2$.

Proposition 3.4 (Non-uniqueness for $\mathtt{GEM}(2)$)**.** *Let ρ be a probability density function on $[0,1]$ such that $\int_0^1 \frac{\rho(u)}{1-u}\,\mathrm{d}u < \infty$. Then the corresponding residual allocation model has the same Bernoulli(1/2) convolution as $\mathtt{GEM}(2)$, if and only if*

$$x\rho(x) = (1-x)\rho(1-x) \quad \textit{for almost all } x \in [0,1]. \tag{3.6}$$

Note that there exist many solutions to (3.6): Starting from an arbitrary nonnegative integrable function f on $[0, \frac{1}{2}]$, one can set $f(x) = \frac{1-x}{x} f(1-x)$ for $x \in (\frac{1}{2}, 1]$ and take $\rho(x) = f(x)/\int f$. As mentioned before, the density of the $\mathtt{Beta}(1,2)$ random variable is $2(1-x)$ and it satisfies Eq. (3.6).

Proof. The Bernoulli($1/2$) convolution of $\mathtt{GEM}(2)$ is equal to $\mathtt{Beta}(1,1)$, i.e. the uniform probability measure on $[0,1]$, by Proposition 3.2. The operator H takes a simpler form and Eq. (3.2) becomes

$$\int_0^x \frac{\rho(u)}{1-u}\,\mathrm{d}u + \int_0^{1-x} \frac{\rho(u)}{1-u}\,\mathrm{d}u = 2, \tag{3.7}$$

for all $x \in [0,1]$. We get (3.6) by differentiating with respect to x. This proves the "only if" direction.

Conversely, if ρ is a probability density function on $[0,1]$ that satisfies Eq. (3.6), then

$$\begin{aligned}[H\rho](x) + [H\rho](1-x) &= \int_0^x \frac{\rho(u)}{1-u}\mathrm{d}u + \int_0^{1-x} \frac{\rho(u)}{1-u}\mathrm{d}u = \int_0^1 \frac{\rho(u)}{1-u}\mathrm{d}u \\ &= \int_0^1 \frac{\rho(u)}{1-u}(1-u+u)\mathrm{d}u = \int_0^1 \rho(u)\mathrm{d}u + \int_0^1 \rho(1-u)\mathrm{d}u = 2,\end{aligned} \tag{3.8}$$

and (3.2) holds true. □

The case of the $\mathtt{GEM}(\theta)$ distribution with $\theta \neq 2$ is more complicated and we do not give a full characterisation of all possibilities. We only prove the existence of many solutions.

We rely on the theory of fractional derivatives and integrals, see e.g. [16, Ch 1] for an extended exposition. For $\alpha > 0$, let $\mathcal{I}^\alpha$ denote the *fractional integral operator* (in the sense of Riemann–Liouville):

$$[\mathcal{I}^\alpha f](x) = \frac{1}{\Gamma(\alpha)} \int_0^x \frac{f(u)}{(x-u)^{1-\alpha}}\,\mathrm{d}u, \tag{3.9}$$

for all $x \in [0,1]$ and all functions f such that the above integral converges absolutely. Its inverse is the *fractional derivative operator* $\mathcal{D}^\alpha$. Writing $\alpha = [\alpha] + \{\alpha\}$ with $[\alpha] \in \mathbb{N}_0$ and $\{\alpha\} \in [0,1)$, it is given by

$$[\mathcal{D}^\alpha f](x) = \frac{1}{\Gamma(1-\{\alpha\})} \frac{\mathrm{d}^{[\alpha]+1}}{\mathrm{d}x^{[\alpha]+1}} \int_0^x \frac{f(t)}{(x-t)^{\{\alpha\}}}\,\mathrm{d}t. \tag{3.10}$$

We introduce the function φ on $[0,1]$ by

$$\varphi(u) = \frac{\rho(u)}{(1-u)^{\theta-1}}. \tag{3.11}$$

We now rewrite Eq. (3.2) using the fractional integral operator in the case where the probability density q is that of $\mathtt{Beta}(\theta/2, \theta/2)$. Taking $q(x) = \frac{\Gamma(\theta)}{\Gamma(\theta/2)^2} x^{\theta/2-1}(1-x)^{\theta/2-1}$ in Eq. (3.1), Lemma 3.1 can be reformulated as follows.

Lemma 3.5 (Non-uniqueness for $\mathtt{GEM}(\theta)$)**.** *Let $\theta > 0$. Assume that φ is a nonnegative function on $[0,1]$ that satisfies*

$$\frac{1}{x^{\theta/2-1}}[\mathcal{I}^{\theta/2}\varphi](x) + \frac{1}{(1-x)^{\theta/2-1}}[\mathcal{I}^{\theta/2}\varphi](1-x) = \frac{2}{\Gamma(\theta/2)} \quad \textit{for all } x \in [0,1]. \tag{3.12}$$

Then $\rho(x) = (1-x)^{\theta-1}\varphi(x)$ is a probability function on $[0,1]$ and the Bernoulli($1/2$) convolution of the residual allocation model from ρ has density $\mathtt{Beta}(\theta/2, \theta/2)$.

The claim about non-uniqueness, Theorem 1.3, is now a consequence of Lemma 3.5.

Proof of Theorem 1.3. We are looking for nonnegative solutions φ of (3.12); then $\rho(x) = (1-x)^{\theta-1}\varphi(x)$ is a solution. Let ε be a function on $[0,1]$ that is antisymmetric around $1/2$, i.e. $\varepsilon(x) = -\varepsilon(1-x)$, and consider the equation

$$[\mathcal{I}^{\theta/2}\varphi](x) = \frac{2}{\Gamma(\theta/2)}\big[x^{\theta/2} + x^{\theta/2-1}\varepsilon(x)\big] \tag{3.13}$$

with $x \in [0,1]$. Solutions of this equation are also solutions of (3.12). Applying the fractional derivative operator on both sides, and using $\mathcal{D}^\alpha\mathcal{I}^\alpha = \mathtt{id}$, we get

$$\begin{aligned}\varphi(x) &= \frac{2}{\Gamma(\theta/2)}\mathcal{D}^{\theta/2}\big[x^{\theta/2} + x^{\theta/2-1}\varepsilon(x)\big](x) \\ &= \frac{2}{\Gamma(\theta/2)\Gamma(1-\{\theta/2\})}\frac{\mathrm{d}^{[\theta/2]+1}}{\mathrm{d}x^{[\theta/2]+1}}\int_0^x \frac{t^{\theta/2} + t^{\theta/2-1}\varepsilon(t)}{(x-t)^{\{\theta/2\}}}\mathrm{d}t.\end{aligned} \tag{3.14}$$

Conversely, if we assume in addition that $\varepsilon(x) = \mathcal{O}(x)$ at $x = 0$, we can use [16, Eq. (2.60)] to verify that (3.13) is satisfied. Indeed, all derivatives in [16, Eq. (2.60)] vanish at $x = 0$.

The contribution of the term $t^{\theta/2}$ can be calculated explicitly; it gives the constant θ. We can also make the change of variables $t \mapsto ux$ and we get

$$\varphi(x) = \theta + \frac{2}{\Gamma(\theta/2)\Gamma(1-\{\theta/2\})}\frac{\mathrm{d}^{[\theta/2]+1}}{\mathrm{d}x^{[\theta/2]+1}}\left[x^{[\theta/2]}\int_0^1 \frac{u^{\theta/2-1}\varepsilon(ux)}{(1-u)^{\{\theta/2\}}}\mathrm{d}u\right]. \tag{3.15}$$

The case $\varepsilon \equiv 0$ leads to $\varphi(x) = \theta$, i.e. $\rho = \mathtt{Beta}(1,\theta)$. But we can also choose $\varepsilon \not\equiv 0$ to be small and smooth enough such that the last term is uniformly bounded by θ. Then $\varphi(x) \geq 0$ for all $x \in [0,1]$. □

4 Comments

4.1 Other examples of non-uniqueness for $p = 1/2$

For $p = 1/2$ there is another example of non-uniqueness of the Bernoulli convolution for $\mathtt{GEM}(2)$, using the Brownian bridge. Namely, let $X_1 \geq X_2 \geq \cdots$ be a ranked list of the excursion lengths away from 0 of a standard Brownian bridge on $[0,1]$, and let ε_i be the indicator that the bridge is positive on the corresponding excursion. Then the ε_i are i.i.d. $\mathtt{Bernoulli}(1/2)$, independent of the X_i, and the Bernoulli($1/2$) convolution $Z = \sum_{i\geq 1}\varepsilon_i X_i$ equals the time spent positive by the bridge. Lévy showed that the latter is uniformly distributed on $[0,1]$, which as we saw coincides with the Bernoulli($1/2$) convolution of $\mathtt{GEM}(2)$. See e.g. [14, Section 2.4] for more information.

We can also use the Brownian pseudo-bridge to get an example of non-uniqueness of the Bernoulli($1/2$) convolution for $\mathtt{GEM}(1)$. Indeed, the ranked list of excursions is given by the two-parameter Poisson-Dirichlet distribution $\mathtt{PD}(1/2, 0)$ and the time spent positive is $\mathtt{Beta}(1/2, 1/2)$; see [15].

4.2 Further questions

It would be interesting to investigate the extent to which Theorems 1.2 and 1.3 hold for other classes of random partitions $(X_i)_{i\geq 1}$ than those formed by residual allocation. One could for example consider more general residual allocation models where the sequence (Y_i) is not i.i.d. but e.g. given by a discrete-time stochastic process. Another natural class of random partitions are those built from subordinators (see [14, Section 5.2]). Briefly, in this case $(X_i)_{i\geq 1}$ is formed by normalising an exhaustive list of the jumps of a subordinator with no drift component. We pose the following two questions:

Question 4.1. Are there analogs of Theorems 1.2 and 1.3 for random partitions built from subordinators?

Question 4.2. For $p \neq 1/2$, are there natural examples of random partitions whose Bernoulli convolutions are identical to those of $\mathtt{GEM}(\theta)$, or other residual allocation models?

4.3 Holroyd's example

If one makes *no* assumptions about the structure of the partition $(X_i)_{i\geq 1}$ then the Bernoulli convolution does not determine the law of the random partition, even if the former is known for *all* $p \in (0,1)$. This is shown by the following example due to A. Holroyd.

The example deals with partitions of just three elements. We consider random variables X_1, X_2, X_3 such that $X_1 \geq X_2 \geq X_3 \geq 0$ and $X_1 + X_2 + X_3 = 1$, as well as independent $\mathtt{Bernoulli}(p)$ random variables $\varepsilon_1, \varepsilon_2, \varepsilon_3$. The first observation is that the law of the Bernoulli(p)-convolution Z is determined by the marginals for X_1, X_2, and $X_1 + X_2$ — no matter what the values of $\varepsilon_1, \varepsilon_2, \varepsilon_3$ are, the random variables X_1 and X_2 appear in the above form. This holds for all p. It is thus enough to show that we can find random variables $\tilde{X}_1$ and $\tilde{X}_2$, distinct from X_1, X_2, such that

$$\begin{aligned} \tilde{X}_1 &\overset{\mathrm{d}}{=} X_1 \\ \tilde{X}_2 &\overset{\mathrm{d}}{=} X_2 \\ \tilde{X}_1 + \tilde{X}_2 &\overset{\mathrm{d}}{=} X_1 + X_2. \end{aligned} \tag{4.1}$$

Let $f(x_1, x_2)$ denote the joint probability density function of (X_1, X_2). It is supported on the set $\Delta \in [0,1]^2$ such that

$$x_1 \geq x_2 \geq 1 - x_1 - x_2 \geq 0, \tag{4.2}$$

see Fig. 1. We can find a square in the set Δ, and define the function $g(x_1, x_2)$ that takes values $\{-1, 0, +1\}$ as shown in Fig. 1. If f is positive on Δ, then $\tilde{f} = f + \eta g$ is positive for η small enough. The function $\tilde{f}$ is the probability density function for $(\tilde{X}_1, \tilde{X}_2)$.

The marginals for $X_1, \tilde{X}_1$ are obtained by integrating $f, \tilde{f}$ along vertical lines. They are clearly identical. Same for the marginals for $X_2, \tilde{X}_2$, obtained by integrating along horizontal lines. And same for the marginals for $X_1 + X_2, \tilde{X}_1 + \tilde{X}_2$, obtained by integrating along oblique lines of slope -1.

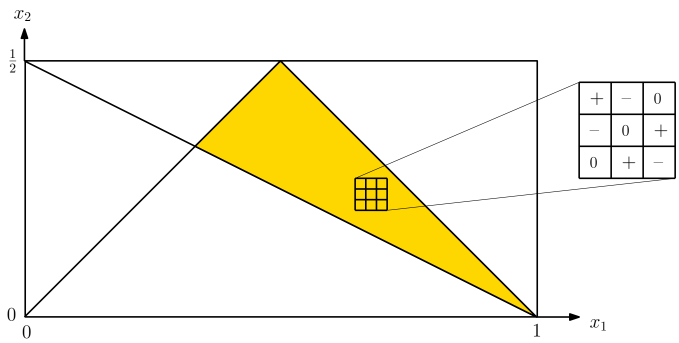

Figure 1: Domain Δ characterised by (4.2) and the square that defines the function g.

Holroyd's example can be generalised to random partitions with infinitely many elements as follows. Let $a \in (\frac{3}{4}, 1]$. Choose (X_1, X_2, X_3) with the constraint $X_1 + X_2 + X_3 = a$; then choose an arbitrary random partition on the remaining interval $[0, 1-a]$. The domain (4.2) is replaced by $x_1 \geq x_2 \geq a - x_1 - x_2 \geq 1 - a$ (it is nonempty for $a > 3/4$). The same argument then applies. It is also possible to take a to be random.

References

[1] M. Aizenman, B. Nachtergaele, *Geometric aspects of quantum spin states*, Comm. Math. Phys. 164, 17–63 (1994). MR-1288152

[2] J. Bertoin, *Random fragmentation and coagulation processes*, Cambridge Studies Adv. Math. 102, Cambridge University Press (2006). MR-2253162

[3] V. Betz, D. Ueltschi, *Spatial random permutations and Poisson-Dirichlet law of cycle lengths*, Electr. J. Probab. 16, 1173–1192 (2011). MR-2820074

[4] P. Billingsley, *Convergence of Probability Measures*, 2nd ed, Wiley (1999). MR-1700749

[5] J.E. Björnberg, J. Fröhlich, D. Ueltschi, *Quantum spins and random loops on the complete graph*, Comm. Math. Phys. (2020); https://doi.org/10.1007/s00220-019-03634-x

[6] J.E. Björnberg, M. Kotowski, B. Lees, P. Miłoś, *The interchange process with reversals on the complete graph*, Electr. J. Probab. 24, paper no. 108 (2019). MR-4017126

[7] P. Diaconis, D. Freedman, *Iterated random functions*, SIAM review 41.1, 45–76 (1999). MR-1669737

[8] W.J. Ewens, *The sampling theory of selectively neutral alleles*, Theor. Popul. Biol. 3, 87–112 (1972). MR-0325177

[9] P.D. Feigin, and R.L. Tweedie, *Linear functionals and Markov chains associated with Dirichlet processes*, Math. Proc. Cambridge Phil. Soc. 105(3) (1989). MR-0985694

[10] C. Goldschmidt, D. Ueltschi, P. Windridge, *Quantum Heisenberg models and their probabilistic representations*, Entropy and the Quantum II, Contemp. Math. 552, 177–224 (2011). MR-2868048

[11] J.F.C. Kingman, *Random discrete distributions*, J. Royal Statist. Soc. B 37, 1–22 (1975). MR-0368264

[12] J.F.C. Kingman, *The representation of partition structures*, J. London Math. Soc. 2.2, 374–380 (1978). MR-0509954

[13] Y. Peres, W. Schlag and B. Solomyak, *Sixty years of Bernoulli convolutions*, Fractal geometry and stochastics II, 39–65, Birkhäuser, Basel (2000). MR-1785620

[14] J. Pitman, *Random weighted averages, partition structures and generalized arcsine laws*, arXiv:1804.07896 (2018)

[15] J. Pitman, M. Yor, *The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator*, Ann. Probab. 25, 855–900 (1997). MR-1434129

[16] S.G. Samko, A.A. Kilbas and O.I. Marichev, *Fractional integrals and derivatives: theory and applications*, Gordon and Breach (1993). MR-1347689

[17] O. Schramm, *Compositions of random transpositions*, Israel J. Math. 147, 221–243 (2005). MR-2166362

[18] S. Sosnovskiy, *On financial applications of the two-parameter Poisson-Dirichlet distribution*, arXiv:1501.01954 (2015)

[19] J.E. Steif and J. Tykesson, *Generalized Divide and Color models*, to appear in ALEA, arXiv:1702.04296. MR-3986920

[20] B. Tóth, *Improved lower bound on the thermodynamic pressure of the spin* $1/2$ *Heisenberg ferromagnet*, Lett. Math. Phys. 28, 75–84 (1993). MR-1224836

[21] D. Ueltschi, *Random loop representations for quantum spin systems*, J. Math. Phys. 54, 083301, 1–40 (2013). MR-3135479

Acknowledgments. We thank Christina Goldschmidt for discussions and for pointing out the examples from the Brownian bridge and pseudo-bridge, and the reference [14]; Jon Warren for further discussions about the pseudo-bridge; Jeff Steif and Johan Tykesson for discussions about their paper [19]; and Alexander Holroyd for explaining his example. We also thank the three referees for positive and helpful comments. We gratefully acknowlege support from the UoC Forum *Classical and quantum dynamics of interacting particle systems*. JEB gratefully acknowledges support from *Vetenskapsrådet* grant 2015-0519 and *Ruth och Nils-Erik Stenbäcks stiftelse*. CM is grateful to the EPSRC for support through the fellowship EP/R022186/1.