

©Copyright 2017

Avi Levy

Novel uses of the Mallows model in coloring and matching

Avi Levy

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Christopher Hoffman, Chair

Alexander Holroyd

Sara Billey

Program Authorized to Offer Degree:
Mathematics

University of Washington

Abstract

Novel uses of the Mallows model in coloring and matching

Avi Levy

Chair of the Supervisory Committee:
Professor Christopher Hoffman
Mathematics

A natural model of a highly ordered random ranking is the Mallows model. Disorder is measured by the number of inversions; these are pairs of elements whose order is reversed. The Mallows model assigns to each ranking of a finite set a weight proportional to a parameter, q , raised to the number of inversions. Rankings of a finite set may be regarded as permutations, and thus the Mallows model consists of a probability measure on permutations.

Originally introduced by Colin L. Mallows in statistical ranking theory [19], this model has enjoyed a recent flurry of interest in contexts including mixing times [2, 7], statistical physics [25, 26], learning theory [6], and longest increasing subsequences [1, 3, 20].

In this work, we apply the model in a novel manner, using its intrinsic mathematical properties to resolve open questions in probability theory posed in [13]. These questions involve neither permutations nor rankings. Instead, we will use the model to construct random proper colorings of the integers possessing certain properties whose joint satisfiability was previously unknown. The first of these properties is that the colorings are finitely dependent: restrictions of the colorings to subsets separated by a fixed, finite, distance are stochastically independent of one another. The second is that they are expressible as finitary factors of iid with finite mean coding radius: this technical condition means, roughly speaking, that the colorings can be efficiently constructed by a parallel distributed algorithm.

In addition to coloring applications, we explore connections between the Mallows model

and stable matching. The celebrated Gale-Shapley algorithm [8] provided the first proof of existence of stable matchings for arbitrary preferences. This is a fundamental result in the economic theory of stable allocations, a theory which appeared in the citation of the 2012 Nobel Prize in Economics awarded to Lloyd Shapley and Alvin E. Roth. Variants of the Gale-Shapley algorithm play a crucial role in practical applications of stable matching, including in the National Resident Matching Program [23], which matches medical school students with residency programs. We introduce a model of stable matching in which the preferences are Mallows-distributed. A novel feature of this model is that preferences are highly correlated. Perhaps surprisingly, this leads to a vastly greater number of stable matchings than in the uniformly random case: exponential in the number of individuals, rather than polynomial.

Finally, we show that the Mallows permutation itself can be regarded as the solution of a stable matching problem. Here, one presupposes a global ranking of the individuals, with random and independent incompatibilities. It turns out that for a finite set of individuals, there is always a unique stable matching. We show that the conditional distribution of this matching, given that it is perfect, is the Mallows measure. We also consider infinite analogues of this result, uncovering new phenomena not present in the finite case.

TABLE OF CONTENTS

	Page
Chapter 1: Introduction	1
1.1 History of the Mallows model	1
1.2 Recent results involving the Mallows model	2
Chapter 2: Application to finitely dependent coloring	4
Mallows permutations and finite dependence	4
Introduction	6
Permutations, Codes, Colorings, and Graphs	15
Finite dependence	23
Reversibility	29
Convergence of Lehmer codes	34
Proof of main theorem	39
Painting algorithm and conditioning	45
Bit-finitary factors	48
Compact Markov chains	50
Higher dimensions and shifts of finite type	51
Open problems, References	52
Finitely dependent insertion processes	56
Introduction	57
Weighted insertion	62
Uniform weight graphs	69
Complete weighted digraphs with property (C)	76
Shifts of finite type	79
References	80
Chapter 3: Stable matching with Mallows preferences	82

3.1	Introduction	82
3.2	Coupling with an ergodic system	84
3.3	Cutpoint criteria	90
3.4	Two positive probabilities	96
3.5	Proof of main theorem	104
Chapter 4: The Mallows permutation as a stable matching		111
4.1	The finite case	111
4.2	Infinite intervals bounded from below	113
4.3	Stable matching of the integers	115
Bibliography		118

ACKNOWLEDGMENTS

I am deeply indebted to Alexander E. Holroyd, who acted as my mentor throughout this journey and during an internship at Microsoft Research, and who has greatly influenced my mathematical development. I am grateful for the support I have received from my advisor Christopher Hoffman, whose advice has broadened my perspectives. I thank James A. Morrow, whose tireless passion for inspiring mathematics in others shaped my formative early years in the graduate program.

I thank my recent collaborators, Omer Angel, Thomas Rothvoß and Elchanan Mossel, whose insight and wisdom have inspired me in new directions, as well as my first collaborators Sergey Bereg and I. Hal Sudborough, who led me onto the path of academic research.

I thank my fellow graduate student and collaborator Tom Hutchcroft, for his help and patience, and whose deep understanding of mathematics has challenged and inspired me. I also thank Clayton Barnes, Matthew Junge, Peter Lin, Harish Ramadas, and Tvrtko Tadic, all of whom were graduate students with me at UW, and who have played a role in my mathematical and social development. I thank my collaborator David Jekel, along with all former members of the mathematics REU at UW who I have had the pleasure of working with, whose tireless passion for mathematics is infective.

DEDICATION

To my parents, for fostering my passion for learning.

Chapter 1

INTRODUCTION

1.1 *History of the Mallows model*

In the 1957 paper NON-NULL RANKING MODELS by Colin L. Mallows [19], the author called attention to a ‘major outstanding problem in ranking theory’, which was the absence of a suitable statistical ranking model in the following scenario.

There is a fixed set of individuals being assessed by a population of judges, or by the same judge in repeated trials, on a particular attribute whose ranking is known *a priori*. The random element is uncertainty of preference, the correlation being the result of real differences between the individuals, and the population is one of rankings conditional on a given objective order.

Mallows proposed two approaches to produce such a model, the first being ‘paired-comparison theory’. He described the general model arising from paired comparisons thusly.

The judge is assumed to arrive at a ranking of n objects U_1, U_2, \dots, U_n by first making all the $\binom{n}{2}$ comparisons between pairs *independently*, but then only accepting the results if they are consistent with a ranking of the n objects.

In fact, special cases of this model had been previously considered. For example, the Bradley-Terry model introduced in 1952 [5] consists of the special case in which

$$\mathbb{P}(U_i < U_j) = \frac{\pi_j}{\pi_i + \pi_j}, \quad 1 \leq i < j \leq n,$$

for some non-negative real numbers $\pi_1, \pi_2, \dots, \pi_n$.

The second approach proposed by Mallows was to consider a model in which the probability of a ranking u_1, u_2, \dots, u_n is a function only of Kendall's τ statistic, given by

$$\tau = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \text{sgn}(u_j - u_i).$$

This statistic had appeared in the 1948 book RANK CORRELATION METHODS, by Maurice Kendall [16]. Mallows then proposed to take the probability of a ranking to be proportional to a parameter raised to the power of Kendall's τ statistic, and the rest is history.

There is no loss of generality in assuming that the ranks u_1, u_2, \dots, u_n take values in $\llbracket 1, n \rrbracket$. If, furthermore, we assume that all ranks are distinct, then the sequence of ranks is a permutation σ of $\llbracket 1, n \rrbracket$. Thus Kendall's τ statistic is a rescaling of the **inversion number** $\text{inv}(\sigma)$, defined by

$$\text{inv}(\sigma) = \#\{1 \leq i < j \leq n: \sigma(i) > \sigma(j)\}.$$

We regard the Mallows model as a probability measure Mal_q on S_n , the set of permutations of $\llbracket 1, n \rrbracket$, given by

$$\text{Mal}_q(\{\sigma\}) = \frac{q^{\text{inv}(\sigma)}}{\sum_{\tau \in S_n} q^{\text{inv}(\tau)}},$$

referred to as the **Mallows measure** on $\llbracket 1, n \rrbracket$. This measure, as well as its infinite analogues, forms a common thread throughout our present work.

1.2 Recent results involving the Mallows model

Card shuffling is one of the original models which fueled the development of the theory of *mixing times*: the study of the rate at which a Markov chain converges to its stationary measure. In [2], the authors study a biased card shuffling Markov chain whose stationary measure is the Mallows measure. By employing a connection with the asymmetric exclusion process on \mathbb{Z} , the authors establish that whenever the Mallows parameter q is not equal to 1 (i.e., whenever the Mallows measure is not the uniform measure), the mixing time of this Markov chain is $O(n^2)$. This confirmed a conjecture made by Diaconis and Ram in [7].

The description of the Mallows measure as the stationary measure of a Markov chain as in [2] is similar to the characterization of the Mallows measure in terms of q -exchangeability. This term is defined as follows in [11]:

A random word $w = w_1 w_2 \cdots$ with letters $w_i \in \mathbb{A}$ over some alphabet $\mathbb{A} \subseteq \mathbb{R}$ is *exchangeable* if swapping the places of two neighboring letters w_i and w_{i+1} does not change the probability. . . . For [a] positive parameter $q \neq 1$, we define w to be *q -exchangeable* if, by swapping the places of two neighboring letters w_i and w_{i+1} , the probability is multiplied by a factor of $q^{\text{sgn}(w_{i+1}-w_i)}$.

The authors show that the most general q -exchangeable word of length n is produced by choosing a random word satisfying $w_1 \leq \cdots \leq w_n$ a.s., then shuffling the letters via an independent Mallows permutation of S_n with parameter q . An infinite version of the Mallows permutation, defined on permutations of $\mathbb{N} = \{1, 2, \dots\}$, is then introduced and characterized by its q -exchangeability (which becomes a statement about the Radon-Nikodym derivative of the measure with respect to its pushword under composition with an adjacent transposition). In the follow-up paper [10], the authors extend this to produce a Mallows measure on permutations of the integers which also satisfies q -exchangeability (but this no longer characterizes it). We return to the topic of Mallows measures on infinite sets in Chapter 4, where we will show that the Mallows measure on \mathbb{N} from [11] arises from a stable matching problem on \mathbb{N} with random compatibility.

Peled and his coauthors have studied large-scale properties of the Mallows permutation in several papers. In [3], the asymptotic length of the longest increasing subsequence in a Mallows permutation was determined in all regimes where $n \rightarrow \infty$ and $q \rightarrow 1$ simultaneously (this was determined in [20] for the case when $n(1-q)$ approaches a constant as $n \rightarrow \infty$). In [9], the asymptotic cycle structure of a Mallows random permutation was determined, and it was shown that in the regime $(1-q)^{-2} \gg n$ there are macroscopic cycles and the law of their normalized lengths approaches the Poisson-Dirichlet law.

Chapter 2

APPLICATION TO FINITELY DEPENDENT COLORING

We begin this chapter by presenting an application of the Mallows model to the problem of constructing a stationary, finitely dependent, proper coloring of the integers. This appears in the article `MALLOWS PERMUTATIONS AND FINITE DEPENDENCE` to follow. While several examples of such colorings have been constructed previously in [13, 14], the constructions of such processes had remained mysterious and in all but one case, no direct probabilistic construction of the colorings on \mathbb{Z} was known. Here we use the Mallows model to produce a family of colorings unifying all previous constructions. Moreover when the Mallows parameter is not equal to 1, we obtain a simple construction of the colorings directly on \mathbb{Z} . These colorings form a three parameter family, and when the parameters are tuned to satisfy a certain equation, one obtains finitely dependent colorings.

In the second article of this chapter, `FINITELY DEPENDENT INSERTION PROCESSES` [18], we consider a class of finitely dependent processes generalizing colorings. The class consists of random graph homomorphisms from the nearest-neighbor graph on the integers to a finite graph. (This is the same thing as a random proper q -coloring of \mathbb{Z} when the graph is the complete graph, without self-loops, on $\{1, \dots, q\}$.) Associated to every such graph is a natural analogue of the construction in [13]. We show that, even if one is allowed to choose this graph arbitrarily, one does not obtain any essentially new processes in this manner.

MALLOWS PERMUTATIONS AND FINITE DEPENDENCE

ALEXANDER E. HOLROYD, TOM HUTCHCROFT, AND AVI LEVY

ABSTRACT. We use the Mallows permutation model to construct a new family of finitely dependent proper colorings of the integers. These colorings are expressible both as finitary factors of iid processes whose coding radii have exponential tails, and as functions of countable Markov chains. The existence of such processes resolves several open questions. Next we show that under mild conditions, finitary factors of iid processes with finite expected coding radii are functions of countable Markov chains. We deduce extensions of the colorings to higher dimensions and to shifts of finite type.

1 Introduction

A stochastic process indexed by a metric space is said to be finitely dependent if subsets of variables separated by some fixed distance are independent. Finitely dependent processes appear in classical limit theorems, statistical physics, and probabilistic combinatorics [4, 21, 27, 28, 33, 37, 43, 50]. The best-known class of finitely dependent processes are the block-factors: processes obtained from an iid sequence by applying a finite-range function. In 1965, Ibragimov and Linnik [33, 34] suggested that there may exist finitely dependent processes not expressible as block factors. While finitely dependent processes enjoyed significant attention in the intervening years [1, 2, 16–18, 23, 35, 36, 53], it was only in 1993 that Ibragimov and Linnik’s question was resolved in the affirmative by Burton, Goulet, and Meester [14]. Many subsequent works [11, 13, 19, 37, 43, 46, 47] explored the properties of such processes, but the question remained: are there ‘natural’ stationary finitely dependent processes that are not block-factors?

Recently, Holroyd and Liggett [30] answered this question in the affirmative, showing that *proper coloring* distinguishes between these classes of processes. More precisely, they constructed stationary finitely dependent colorings of \mathbb{Z} , and provided a simple argument showing that no block-factor is a coloring. A process $(X_i)_{i \in \mathbb{Z}}$ is a **q -coloring** if each X_i takes values in $\{1, \dots, q\}$, and almost surely $X_i \neq X_{i+1}$ for all i . A stationary q -coloring is **k -dependent** if the random sequences $(X_i)_{i < 0}$ and $(X_i)_{i \geq k}$ are independent of one another. A process is finitely dependent if it is k -dependent for some k , and it is a coloring if it is a q -coloring for some q . By an argument of Schramm [32], there is no stationary 1-dependent 3-coloring of \mathbb{Z} . Holroyd and Liggett [30] constructed a stationary 1-dependent 4-coloring

and a stationary 2-dependent 3-coloring (implying trivially that stationary k -dependent q -colorings exist for all $k \geq 1$ and $q \geq 3$ other than $(k, q) = (1, 3)$). These colorings were constructed in [30] by specifying cylinder probabilities and appealing to the Kolmogorov extension theorem, without a direct probabilistic construction on \mathbb{Z} .

Here is a way to formalize this last concept. We say that $X = (X_i)_{i \in \mathbb{Z}}$ is a **finitary factor** of an iid process, or simply that X is **ffiid**, if it is equal in law to $F(Y)$ where $Y = (Y_i)_{i \in \mathbb{Z}}$ is an iid sequence and F is a **translation-equivariant** function (i.e. one that commutes with translations of \mathbb{Z}) satisfying the following property: for almost every sequence y (with respect to the law of Y), there exists $r < \infty$ such that $F(y)_0 = F(y')_0$ whenever y' agrees with y on $\{-r, \dots, r\}$. Let $R(y)$ be the minimal such r . The random variable $R = R(Y)$ is the **coding radius** of the finitary factor. In other words, X_0 is determined by examining only those variables Y_i within a finite but random distance R from the origin. Finitary factors generalize block-factors: the latter are finitary factors with bounded coding radius.

In [29] it was shown that the 1-dependent 4-coloring of [30] is ffiid with infinite expected coding radius. However, the following question mentioned therein remained unanswered: does there exist a finitely dependent coloring that is ffiid with finite mean coding radius?

We resolve this question as well as several others from [29–31] by constructing a new family of finitely dependent colorings whose coding radii have exponential tails.

Theorem 1. *There exists a stationary, reversible, finitely dependent proper coloring of \mathbb{Z} that is symmetric under permutations of the colors and that can be expressed in each of the following ways:*

- (i) *as a finitary factor of an iid process, with exponential tail on the coding radius; and also*
- (ii) *as a function of a countable Markov chain with exponential tail on the return time to any given state.*

More precisely, there exists a stationary k -dependent q -coloring with all of the above properties for each of

$$(k, q) = (1, 5), (2, 4), (3, 3),$$

as well as for all larger q in each case.

By the statement that a process $X = (X_i)_{i \in \mathbb{Z}}$ is a ‘function of a countable Markov chain’ we mean that there exists a Markov chain $(Y_i)_{i \in \mathbb{Z}}$ on a countable state space S and a function h on S such that $X \stackrel{d}{=} (h(Y_i))_{i \in \mathbb{Z}}$. A process X is **reversible** if $(X_i)_{i \in \mathbb{Z}} \stackrel{d}{=} (X_{-i})_{i \in \mathbb{Z}}$, and

a q -coloring X is symmetric under permutations of the colors if $(X_i)_{i \in \mathbb{Z}} \stackrel{d}{=} (\sigma(X_i))_{i \in \mathbb{Z}}$ for any permutation σ of $\{1, \dots, q\}$.

The existence of colorings satisfying properties (i) and (ii) of Theorem 1 resolves open problems (v) and (ii) of [30], respectively. The two properties are related for general stationary processes; we explain how in Theorem 2.

We will show that our colorings with $k = 1$ and $q \geq 5$ coincide with the colorings constructed in [31]. The purpose of that paper was to show that there are symmetric finitely dependent q -colorings with $q \geq 5$. Thus Theorem 1 shows that these colorings are *ffiid* with finite mean coding radius, resolving [31, Problem 4].

The 1-dependent q -colorings from Theorem 1 have the further property that if one conditions on the absence of color q , the resulting coloring is equal in law to the 2-dependent $(q-1)$ -coloring. (No other colorings from the theorem may be obtained from one another by conditioning in this manner.) Each of the k -dependent q -colorings constructed in Theorem 1 is **strictly k -dependent**, i.e., k -dependent but not $(k-1)$ -dependent.

Sections 4 and 5 of [30] show that finitely dependent colorings can be written neither as block factors nor as functions of finite state Markov chains, respectively. The former result is a consequence of an earlier result in [49], where it appears in a different form, motivated by applications in distributed computing. Further consequences and extensions appear in [3] and [32]. For example, block factors must contain arbitrarily long constant sequences with positive probability.

We reiterate a natural conjecture suggested in [29]: there exists a k -dependent q -coloring that is a finitary factor of an iid process with finite mean coding radius if and only if $k \geq 1$, $q \geq 3$, and $(k, q) \notin \{(1, 3), (2, 3), (1, 4)\}$. Theorem 1 establishes half of this conjecture: it remains to be seen whether for $(k, q) \in \{(2, 3), (1, 4)\}$ there exists a k -dependent q -coloring that is a finitary factor of an iid process with finite mean coding radius.

Coloring has applications in computer science. For instance, colors may represent time schedules or communication frequencies for machines in a network, and adjacent machines are not permitted to conflict with each other. Finite dependence implies a security benefit — an adversary who gains knowledge of some colors learns nothing about the others, except within a fixed finite distance. A *ffiid* coloring with finite mean coding radius is desirable for the purpose of efficient computation. Such a coloring can be computed by the machines in distributed fashion, based on randomness generated locally, combined with communication with machines within a random distance of finite mean. All machines follow the same protocol, and no central authority is needed. See e.g. [44, 49] for more information.

Outline of proof

We next discuss the main ideas behind the proof of Theorem 1, which involves an intricate interplay of various ideas from combinatorics and physics.

At the heart of our construction (as well as those of [30]) is the following simple but mysterious picture. Imagine that integers arrive in a *random* order. When an integer arrives, it is assigned a uniformly random color among those not present in its current neighbors, by which we mean the nearest integers to its left and right that arrived previously. As a useful alternative description, the random order gives rise to a graph, which we call the *constraint graph*, in which two integers are adjacent if and only if they were neighbors at some time. (The constraint graph was also considered in [29], and may be interpreted as the planar dual of the binary search tree [22] of a permutation.) The final coloring is a uniformly random proper coloring of the constraint graph.

The proof of finite dependence begins with a version of this picture restricted to a finite interval, and involves remarkable cancellations that occur only when the set-up is precisely correct. (Indeed, it is surprising that they can occur at all). The required arrival order is not uniformly random. Rather, it arises by re-weighting a simple underlying probability measure by the number of proper colorings of the constraint graph. The fact that such a re-weighting can produce colorings with exceptional properties is reminiscent of the theory of two-dimensional quantum gravity, in which statistical mechanics models are studied on random planar maps that are weighted according their partition function for the model. There, as here, the model on the appropriately weighted random map has special properties that are not enjoyed by the same model on, say, a Euclidean lattice. See e.g. [24] and references therein.

For the 1-dependent 4-coloring and 2-dependent 3-coloring of [30], the underlying measure is uniform over permutations of an interval. For our new construction, the underlying measure is the *Mallows* measure, in which each permutation is weighted by a parameter t raised to the power of the number of its inversions. (An inversion is a pair of elements whose order is reversed.) The Mallows measure was originally introduced in statistical ranking theory [45], and has enjoyed a recent flurry of interest in contexts including mixing times [8, 20], statistical physics [56, 57], learning theory [12], and longest increasing subsequences [6, 9, 48]. The computations and combinatorial identities required to prove finite dependence in our case are t -analogues of those in [30]. (The more usual terminology is ‘ q -analogue’, but in this article q is reserved for the number of colors. See [55] for background on q -analogues.) Since the Mallows measure is not reflection-invariant, the reversibility claimed in Theorem 1 requires a further highly non-trivial combinatorial argument.

The Mallows parameter t must be chosen carefully. Specifically, for the q -coloring to be k -dependent, the parameters q , k , and t must satisfy the ‘tuning equation’

$$qt(t^k - 1) = (t + 1)(t^{k+1} - 1). \quad (1)$$

The tuning equation arises by setting a certain coefficient equal to zero in a recurrence for the cylinder probabilities of the colorings. Finite dependence of the colorings stems from this cancellation. This is reminiscent of a phenomenon in the theory of Schramm–Loewner evolution, in which $\text{SLE}(\kappa)$ curves possess additional distributional symmetries for special values of κ , stemming from cancellations in the coefficients of a stochastic differential equation [41].

For the three cases $(k, q) = (1, 5), (2, 4), (3, 3)$ highlighted in Theorem 1, the required values of t are respectively

$$\frac{3 - \sqrt{5}}{2}, \quad \frac{3 - \sqrt{5}}{2}, \quad \text{and} \quad \frac{1 + \sqrt{13} + \sqrt{2(\sqrt{13} - 1)}}{4}.$$

The equality between the t values for the pair of cases $(1, 5)$ and $(2, 4)$ generalizes to the pair $(1, q)$ and $(2, q - 1)$ with $q \geq 4$. This is behind the conditioning property mentioned earlier.

When restricted to finite intervals, the above construction yields a consistent family of random colorings, which extends to a coloring of \mathbb{Z} via Kolmogorov extension. However, proving that this random coloring satisfies properties (i) and (ii) of the theorem requires a more direct construction. To achieve this, we extend the random arrival picture to \mathbb{Z} . This presents several challenges. On a finite interval, the re-weighting introduces an extra factor every time an integer arrives at either end of the interval of its predecessors.

For the uniform model introduced in [30], it turns out that these endpoint arrivals are sufficiently rare that their effect washes out in the limit, and the associated random order on \mathbb{Z} is in fact uniform. However, this means that the constraint graph has many long edges. (A typical edge has infinite mean length, by the well-known record value waiting time property.) This is the reason for the power law tail in the finitary factor construction of the 1-dependent 4-coloring in [29]. (Since it is also necessary to properly color the constraint graph, it turns out that this framework does not yield a finitary factor construction of the 2-dependent 3-coloring at all. See [29] and the earlier discussion.)

The situation for our model is very different. For a fixed parameter t , the Mallows permutation of a sufficiently large finite interval can be naturally viewed as a perturbation of the identity, with a strong left-to-right bias in the corresponding order. Consequently, (right) endpoint arrivals now have a positive density, and their re-weighting effect is *not*

washed out in the limit. The resulting random order on \mathbb{Z} follows a new (and quite natural) two-parameter variant of the Mallows measure, which we call the *bubble-biased Mallows measure*. (Infinite-interval versions of the standard Mallows measure were constructed in [26].) As a result of the endpoint arrivals, the constraint graph is much better behaved than in the previous case. It decomposes into a sequence of finite ‘bubbles’, joined at their endpoints. The length of a bubble has exponential tails, allowing us to prove properties (i) and (ii) in Theorem 1. Some further technical details are involved in making the transition from finite intervals to \mathbb{Z} rigorous. In particular, it is useful to consider convergence of the Lehmer code of a permutation (see e.g. [7] for a definition).

The distinction between our new construction and that of [30] may be interpreted via the language of phase transition. For the tuning equation (1) to have a solution in t , the parameters k and q must satisfy the inequality $qk \geq 2(k+1)$. This is satisfied with equality along a critical curve $qk = 2(k+1)$ in the (k, q) plane – see Figure 1. On the curve we have $t = 1$, and there are precisely two integer solutions, $(1, 4)$ and $(2, 3)$, giving the colorings of [30]. (The Mallows measure reduces to the uniform measure when $t = 1$.) On one side of the curve, the construction does not work, while on the other side we obtain the colorings of this article. This fits the signature of a phase transition: an abrupt qualitative change in behavior, with power laws at criticality, and exponential decay in the off-critical regime. But we believe that the same phase transition phenomenon applies to finitely dependent colorings of \mathbb{Z} in complete generality, not just to the specific construction here (although currently no other constructions are known, besides trivial embellishments). Indeed, it is proved in [30] that no stationary 1-dependent 3-coloring exists (so no solution exists on that side of the curve). We believe that the stationary 1-dependent 4-coloring and 2-dependent 3-coloring (the ‘critical’ cases) are unique. (Some evidence for the former case is given in [30].) Moreover, we conjecture that no stationary 1-dependent 4-coloring or 2-dependent 3-coloring is fffd with finite mean coding radius.

We reiterate that in [29], the 1-dependent 4-coloring was shown to be fffd with infinite expected coding radius. We do not have an analogous explicit representation of the 2-dependent 3-coloring as a finitary factor of iid. A result of Smorodinsky [54] states that stationary finitely dependent processes of equal entropy are finitarily isomorphic, implying that the 2-dependent 3-coloring of [30] is fffd. Unfortunately, [54] contains only a brief sketch of the proof, and the details of the argument do not seem to be available.

An algorithm for finitely dependent coloring.

There is a remarkably direct construction of the colorings from the main theorem which we now present. While properties (i) and (ii) in the theorem (as well as stationarity, coloring

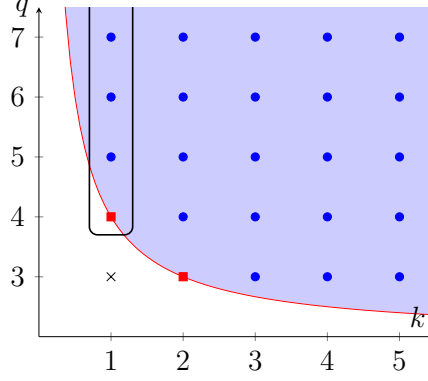


FIGURE 1. Phase diagram for k -dependent q -colorings. The phase boundary is the curve $qk = 2(k + 1)$. The two lattice points on this curve correspond to the 1-dependent 4-coloring and 2-dependent 3-coloring of [30]. In the region $qk < 2(k + 1)$ there do not exist k -dependent q -colorings (the \times at $(k, q) = (1, 3)$ indicates the non-trivial case ruled out by an argument of Schramm in [32]). The 1-dependent q -colorings from [31] correspond to the outlined region. When $qk > 2(k + 1)$ there exists a k -dependent q -coloring that is ffid with finite expected coding radius, by Theorem 1.

constraints, and color symmetry) follow in a straightforward manner from the description, the finite dependence and reversibility properties are more subtle, and will be explained later.

Algorithm 1 (The painting algorithm). *Input positive integers q and k satisfying $qk > 2(k + 1)$. Output the k -dependent q -coloring $(X_i)_{i \in \mathbb{Z}}$ constructed as follows.*

Stage 0. *Let t be the unique solution to (1) in $(0, 1)$. Set $s = t(q - 2)/(q - 1 - t)$. Start with X_i unassigned for every $i \in \mathbb{Z}$.*

Stage 1. *Let $B = (B_i)_{i \in \mathbb{Z}}$ be an iid Bernoulli process with each $B_i \in \{0, 1\}$ taking value 1 with probability s . To each i with $B_i = 1$, assign a random color $X_i \in \{1, \dots, q\}$, in such a way that, conditional on X , the subsequence $(X_i : B_i = 1)$ of assigned values is the trajectory of a simple symmetric random walk on the complete graph with vertex set $\{1, \dots, q\}$ at stationarity.*

Stage 2. *Consider pairs of nearest integers $a < b$ that were assigned colors in Stage 1, that is, $B_a = B_b = 1$ while $B_i = 0$ for $a < i < b$. Independently for each such pair and conditional on Stage 1, we fill in the missing colors via the following recursive procedure. Let K be a random element of $\{a + 1, \dots, b - 1\}$ with $\mathbb{P}(K = \ell) = c t^\ell$, where c is a constant of proportionality. Assign X_K a uniformly random color in $\{1, \dots, q\} \setminus \{X_a, X_b\}$. Conditional on the previous steps, recursively apply the same procedure to each of the intervals $\{a, \dots, K\}$ and $\{K, \dots, b\}$ until all integers have been assigned colors.*

It is important to note that the conditional law, given Stage 1, of the coloring $(X_i)_{i=a}^b$ for integers $a < b$ in Stage 2 is *not* simply the conditional law of the final process $(X_i)_{i \in \mathbb{Z}}$ restricted to $\{a, \dots, b\}$ given X_a and X_b . For example, it is possible that $(X_1, X_2, X_3, X_4) = (1, 2, 1, 2)$ (for instance if all 4 values are assigned at Stage 1). However, if Stage 1 assigns $X_1 = 1$ and $X_4 = 2$ but not X_2 or X_3 , it is then impossible for Stage 2 to fill the interval in this way, because both colors 1 and 2 are unavailable for the first insertion. Our construction is more subtle than such naïve conditioning would suggest. In particular, the specific choices of the parameters s and t of the Bernoulli and geometric processes are crucial, as we shall see.

Bit-finitary factors of iid and countable Markov chains.

In this section we investigate the connection between items (i) and (ii) of Theorem 1, that is, the connection between the expressibility of a process as a finitary factor of iid with certain desirable properties, and the expressibility of the same process as a function of a Markov chain with certain other desirable properties.

Rudolph [52] proved that a bi-infinite trajectory of a mixing, countable, positive recurrent Markov chain satisfying a finite entropy condition can be expressed as a finitary factor of iid if and only if the time for the chain to hit any particular state has an exponential tail. In this section, we provide a complement to Rudolph's result, giving a sufficient condition under which a finitary factor process $X = F(U)$ of an iid process U is also expressible as a function of a countable Markov chain. Intuitively, the condition is that F has finite mean coding radius and that F only has to query a finite (but random) number of the random bits of U in order to compute each of its outputs.

For the purposes of this section, it is convenient for us to think of our iid random variables as taking values in the space $\{0, 1\}^{\mathbb{N}}$, rather than the more usual $[0, 1]$. Endow $\{0, 1\}^{\mathbb{N}}$ with the product of the uniform measure on $\{0, 1\}$. Thus, each of our random variables is an infinite string of independent uniformly random bits. Given $x \in (\{0, 1\}^{\mathbb{N}})^{\mathbb{Z}} = \{0, 1\}^{\mathbb{Z} \times \mathbb{N}}$, a finite set $S \subset \mathbb{Z} \times \mathbb{N}$, a discrete space A and a function $f : \{0, 1\}^{\mathbb{Z} \times \mathbb{N}} \rightarrow A$, we say that $f(x)$ is **determined by the restriction of x to S** if there exists an element $a \in A$ such that $f(x') = a$ for almost every $x' \in \{0, 1\}^{\mathbb{Z} \times \mathbb{N}}$ such that the restrictions of x and x' to S coincide.

We say that a factor $F : \{0, 1\}^{\mathbb{Z} \times \mathbb{N}} \rightarrow A^{\mathbb{Z}}$ is **bit-finitary** if for almost every $x \in \{0, 1\}^{\mathbb{Z} \times \mathbb{N}}$ (with respect to the product of the uniform measure on $\{0, 1\}$), there exists a finite set $S \subset \mathbb{Z} \times \mathbb{N}$ such that $F(x)_0$ is determined by the restriction of x to S . For $i \in \mathbb{Z}$ we let $r_i(x)$ be the minimal integer $r \geq 0$ such that $F(x)_i$ is determined by the restriction of x to $\{i - r, \dots, i + r\} \times \{0, \dots, d\}$ for some $d < \infty$. We say that $(r_i(x))_{i \in \mathbb{Z}}$ are the

bit-finitary coding radii of F . We say that F has finite expected bit-finitary coding radius if $\mathbb{E}r_0(U) < \infty$.

Theorem 2. *Let $U = (U_{i,j})_{(i,j) \in \mathbb{Z} \times \mathbb{N}}$ be a collection of uniform $\{0, 1\}$ random variables, let A be a discrete space, and let $F : \{0, 1\}^{\mathbb{Z} \times \mathbb{N}} \rightarrow A$ be a bit-finitary factor with finite expected bit-finitary coding radius. Then there exists a countable set \mathcal{K} and a factor $G : \{0, 1\}^{\mathbb{Z} \times \mathbb{N}} \rightarrow \mathcal{K}^{\mathbb{Z}}$ such that the following claims hold.*

- (i) *The process $G_i(U)$ is a Markov chain.*
- (ii) *There exists a function $h : \mathcal{K} \rightarrow A$ such that*

$$h \circ G_i(U) = F_i(X)$$

for all $i \in \mathbb{Z}$ almost surely.

It is not hard to see that the construction of the finitary factors in Theorem 1 can be taken to be bit-finitary, so that it would be possible to deduce item (ii) of that theorem from Theorem 2. (For this example, however, there is a more obvious construction of the Markov chain, which we use to prove item (ii) of Theorem 1 directly.)

Compact Markov chains.

While we now have an entire family of finitely dependent colorings, we lack a simple proof of finite dependence (recall that our argument relies on delicate cancellations in the finite dimensional distributions). There is, however, a more transparent manner in which finitely dependent colorings may arise. Given a (discrete-time) Markov chain on a compact metric space that always moves by at least some fixed amount, one obtains a coloring by partitioning the state space into sufficiently small sets and considering the sequence of sets encountered by a trajectory of the chain. If one could arrange for the chain to mix after a finite number of transitions, this would yield a finitely dependent coloring.

We show that there is no reversible Markov chain with these properties. On the other hand, there is a simple example if one allows the state space to be non-compact.

Proposition 3. *If $(X_n)_{n \in \mathbb{Z}}$ is a stationary reversible Markov process on a compact metric space (S, d) and X_0 is independent of X_k for some $k > 0$, then there does not exist $\varepsilon > 0$ such that $d(X_0, X_1) \geq \varepsilon$ almost surely.*

Proposition 4. *There exists a stationary Markov process $(X_n)_{n \in \mathbb{Z}}$ on a non-compact countable metric space (S, d) such that X_0 is independent of X_2 and $d(X_0, X_2) \geq 1$ a.s.*

Shifts of finite type and higher dimensions.

From a finitely dependent coloring of \mathbb{Z} one may construct finitely dependent colorings in higher dimensions, as well as finitely dependent processes satisfying more general local constraints (namely non-lattice shifts of finite type), as shown in [30]. Applying this to the colorings from Theorem 1 yields the next two corollaries, which we state after giving the necessary definitions.

The hypercubic lattice is the graph with vertex set \mathbb{Z}^d and an edge between u and v whenever $\|u - v\|_1 = 1$; the graph is also denoted \mathbb{Z}^d . A process on \mathbb{Z}^d is stationary if it is invariant in law under all translations of \mathbb{Z}^d , and it is *ffiid* if it is equal in law to $F(Y)$ where Y is an iid process on \mathbb{Z}^d and F is a translation-equivariant function satisfying the following property: for almost every sequence y (with respect to the law of Y), there exists $r < \infty$ such that $F(y)_0 = F(y')_0$ whenever y' agrees with y on $\{-r, \dots, r\}^d$. Let $R(y)$ be the minimal such r . We call the random variable $R = R(Y)$ the **coding radius** of the process. A process indexed by a graph is **k -dependent** if its restrictions to two subsets of V are independent whenever the subsets are at graph-distance greater than k from each other.

The following is a consequence of our Theorem 1 combined with methods of [30].

Corollary 5. *Let $d \geq 2$. There exist integers $q = q(d)$ and $k = k(d)$ such that:*

- (i) *there exists a *ffiid* 1-dependent q -coloring of \mathbb{Z}^d with exponential tail on the coding radius;*
- (ii) *there exists a stationary *ffiid* k -dependent 4-coloring of \mathbb{Z}^d with exponential tail on the coding radius.*

Coloring is a special case of the following more general notion, in which the requirement that adjacent colors differ is replaced with arbitrary local constraints. A **shift of finite type** is a set of configurations S characterized by an integer k and a set $W \subseteq \{1, \dots, q\}^k$ as follows:

$$S = S(q, k, W) := \{x \in \{1, \dots, q\}^{\mathbb{Z}} : (x_{i+1}, \dots, x_{i+k}) \in W \ \forall i \in \mathbb{Z}\}.$$

We call the shift of finite-type **non-lattice** if for some $w \in W$ we have that

$$\gcd\{t \geq 1 : \exists x \in S \text{ s.t. } (x_1, \dots, x_k) = (x_{t+1}, \dots, x_{t+k}) = w\} = 1.$$

Corollary 6. *Let S be a non-lattice shift of finite type on \mathbb{Z} . There exists an integer k (depending on S) and a k -dependent *ffiid* process X with exponential tail on the coding radius such that the random sequence X belongs to S almost surely.*

Intermediate in generality between q -colorings and shifts of finite type is the class of stochastic processes taking values in the vertex set of a finite graph such that realizations of the process are a.s. paths. For the complete graph on q vertices K_q , such a process is precisely a q -coloring. A natural modification of the construction of [30] was systematically investigated in [42], in which the graph K_q was replaced with a weighted graph. It was found that, other than straightforward modifications of the 1-dependent 4-coloring and 2-dependent 3-coloring of [30], no other finitely dependent processes arise in this manner. It would be interesting to see if the obvious t -analogue of this result is true.

Outline of the paper

Section 2 covers background material and simple facts about the combinatorial objects we will use in the proof of the main theorem. Section 3 constructs the colorings in the main theorem by starting on finite intervals and using Kolmogorov extension. Section 4 shows that the colorings are reversible. Sections 5 and 6 complete the proof of the main theorem by providing a second construction of the colorings as a finitary factor with exponential tails on the coding radius.

The remaining results claimed in the introduction are proven in Sections 8 through 10. Open problems are in Section 11.

2 Permutations, Codes, Colorings, and Graphs

This section introduces notation and basic facts used in the proof of Theorem 1. As stated previously, the essence of this theorem is that finitely dependent colorings arise as t -analogues of the random colorings in [30]. The t -analogue of a positive integer n is $[n]_t := 1 + t + \dots + t^{n-1} = (1 - t^n)/(1 - t)$. Many numerical equalities that are combinatorial in nature generalize to polynomial identities between t -analogues. This phenomenon appears frequently in algebraic combinatorics [55]. The t -**factorial** $[n]_t!$ and the t -**binomial coefficient** $\binom{n}{k}_t$ are defined via the formulas

$$[n]_t! := \prod_{k=1}^n [k]_t \quad \text{and} \quad \binom{n}{k}_t := \frac{[n]_t!}{[k]_t! [n-k]_t!}.$$

There are $n!$ permutations in S_n . A t -analogue of this fact is that

$$[n]_t! = \sum_{\sigma \in S_n} t^{\text{inv}(\sigma)}, \tag{2}$$

where the inversion number $\text{inv}(\sigma)$ is defined to be

$$\text{inv}(\sigma) := \#\{1 \leq i < j \leq n: \sigma(i) > \sigma(j)\}.$$

Equation (2) is well known and easy to prove, if one uses the right bijection (see e.g. [55, Prop. 1.3.17]). It also follows from the proof of Lemma 8 later in this section.

The **Mallows measure** Mal_t with parameter t is the probability measure on S_n assigning to each permutation $\sigma \in S_n$ a probability of $t^{\text{inv}(\sigma)}/[n]_t!$. A **Mallows random permutation** is a random element of S_n whose law is a Mallows measure. Since $\text{inv}(\sigma) = \text{inv}(\sigma^{-1})$, a Mallows random permutation is equal in law to its inverse. Various statistics of Mallows random permutations are related to geometric random variables. We will need a rather extensive array of variants of the geometric distribution.

Let X be a random variable.

(i) X is an **i -truncated, t -geometric** random variable if

$$\mathbb{P}(X = j) = \frac{t^j}{1 + t + \dots + t^i}, \quad 0 \leq j \leq i.$$

(ii) X is a **u -zero-weighted, i -truncated, t -geometric** random variable if

$$\mathbb{P}(X = j) = \frac{u^{\mathbf{1}_{[j=0]}} t^j}{u + t + \dots + t^i}, \quad 0 \leq j \leq i.$$

(iii) X is a **u -max-weighted, i -truncated, t -geometric** random variable if

$$\mathbb{P}(X = j) = \frac{u^{\mathbf{1}_{[j=n]}} t^j}{1 + t + \dots + t^{n-1} + ut^n}, \quad 0 \leq j \leq n.$$

(iv) X is a **u -end-weighted, i -truncated, t -geometric** random variable if

$$\mathbb{P}(X = j) = \frac{u^{\mathbf{1}_{[j \in \{0, i\}]}} t^j}{u + t + \dots + t^{i-1} + ut^i}, \quad 0 \leq j \leq i.$$

(v) X is a **u -zero-weighted, t -geometric** random variable if

$$\mathbb{P}(X = j) = \frac{u^{\mathbf{1}_{[j=0]}} t^j}{u + \frac{t}{1-t}}, \quad 0 \leq j < \infty.$$

Intervals of integers are sets of the form $I \cap \mathbb{Z}$ where I is an interval of real numbers. We write $\llbracket a, b \rrbracket$ for $[a, b] \cap \mathbb{Z}$ and we use similar blackboard-bold notation for other types of intervals as well. The cardinality of a set is denoted by $\#S$.

As we will be constructing colorings directly on \mathbb{Z} , it is convenient to introduce notation for permutations of arbitrary integer intervals I , which may be finite or infinite. A permutation of I is a bijection from I to itself, and we write $\text{Sym}(I)$ for the set of all such bijections. We identify $\text{Sym}(I)$ with the subset of $\text{Sym}(\mathbb{Z})$ consisting of permutations fixing

all elements of $\mathbb{Z} \setminus I$. A **finite permutation** is a permutation fixing all but finitely many integers. For permutations $\sigma, \tau \in \text{Sym}(I)$, we write $\sigma \circ \tau$ for the permutation mapping i to $\sigma(\tau(i))$. For a sequence of permutations $\{\sigma_j\}_{j \in J}$ indexed by a finite interval $J = \llbracket a, b \rrbracket$ of \mathbb{Z} , we denote the composite permutation by

$$\bigcirc_{j \in J} \sigma_j := \sigma_a \circ \sigma_{a+1} \circ \cdots \circ \sigma_{b-1} \circ \sigma_b.$$

The Lehmer code is a standard way of encoding permutations of finite intervals by a sequence of integers [39]. We will use an extension of this to (possibly infinite) intervals I of \mathbb{Z} . For such intervals, we define the **Lehmer code** to be the map $\mathcal{L}: \text{Sym}(I) \rightarrow \llbracket 0, \infty \rrbracket^I$ given by

$$\mathcal{L}(\sigma)_i = \#\{j \in I: j > i \text{ and } \sigma(j) < \sigma(i)\}, \quad \sigma \in \text{Sym}(I), \quad i \in I.$$

The Lehmer code is a refinement of the **inversion number**,

$$\text{inv}(\sigma) := \#\{(i, j) \in I^2: j > i \text{ and } \sigma(j) < \sigma(i)\},$$

in the sense that $\sum_{i \in I} \mathcal{L}(\sigma)_i = \text{inv}(\sigma)$. A variant of the Lehmer code is the **insertion code**, which is the map $\tilde{\mathcal{L}}: \text{Sym}(I) \rightarrow \llbracket 0, \infty \rrbracket^I$ given by

$$\tilde{\mathcal{L}}(\sigma)_i := \mathcal{L}(\sigma)_{\sigma^{-1}(i)} = \#\{j \in I: j < i \text{ and } \sigma^{-1}(j) > \sigma^{-1}(i)\}, \quad \sigma \in \text{Sym}(I), \quad i \in I.$$

The entries of $\tilde{\mathcal{L}}(\sigma)$ are a permutation of those in $\mathcal{L}(\sigma)$, so $\sum_{i \in I} \tilde{\mathcal{L}}(\sigma)_i = \text{inv}(\sigma)$.

Clearly $0 \leq \mathcal{L}(\sigma)_i \leq \sup I - i$ and $0 \leq \tilde{\mathcal{L}}(\sigma)_i \leq i - \inf I$ for all $i \in I$. Let

$$\begin{aligned} \Omega_I &= \{\ell \in \llbracket 0, \infty \rrbracket^I: 0 \leq \ell_i \leq \sup I - i, \forall i \in I\} \quad \text{and} \\ \tilde{\Omega}_I &= \{\ell \in \llbracket 0, \infty \rrbracket^I: 0 \leq \ell_i \leq i - \inf I, \forall i \in I\}, \end{aligned}$$

so that $\mathcal{L}(\text{Sym}(I)) \subseteq \Omega_I$ and $\tilde{\mathcal{L}}(\text{Sym}(I)) \subseteq \tilde{\Omega}_I$. When I is finite, \mathcal{L} and $\tilde{\mathcal{L}}$ are bijections from $\text{Sym}(I)$ to Ω_I and $\tilde{\Omega}_I$, respectively, with explicit inverse functions which we now describe.

For a finite interval J of \mathbb{Z} , let $\pi_J^- \in \text{Sym}(\mathbb{Z})$ denote the permutation fixing $\mathbb{Z} \setminus J$ and cyclically decrementing J , i.e.,

$$\pi_J^-(j) = \begin{cases} j - 1, & \min J < j \leq \max J \\ \max J, & j = \min J \\ j, & j \in \mathbb{Z} \setminus J. \end{cases}$$

The permutation π_J^- has a single cycle and it is of size $\#J$.

Let $\mathcal{D}: \llbracket 0, \infty \rrbracket^I \rightarrow \text{Sym}(\mathbb{Z})$ denote the map

$$\mathcal{D}(\ell) = \bigcirc_{i \in I} \pi_{\llbracket i, i+\ell_i \rrbracket}^-, \quad \ell \in \llbracket 0, \infty \rrbracket^I \quad (3)$$

and let $\tilde{\mathcal{D}}: \llbracket 0, \infty \rrbracket^I \rightarrow \text{Sym}(\mathbb{Z})$ denote the map

$$\tilde{\mathcal{D}}(\ell) = \bigcirc_{i \in I} \pi_{\llbracket i-\ell_i, i \rrbracket}^-, \quad \ell \in \llbracket 0, \infty \rrbracket^I. \quad (4)$$

Lemma 7. *Suppose that I is a finite interval of \mathbb{Z} . Then \mathcal{L} is a bijection from $\text{Sym}(I)$ to Ω_I , with inverse given by the restriction of \mathcal{D} to Ω_I . Similarly $\tilde{\mathcal{L}}$ is a bijection from $\text{Sym}(I)$ to $\tilde{\Omega}_I$, with inverse given by the restriction of $\tilde{\mathcal{D}}$ to $\tilde{\Omega}_I$.*

The proof, which is straightforward, is omitted. Similar results are well known and appear for example in [9, 26, 55]. One may visually interpret the relationship between \mathcal{L} and \mathcal{D} using Figure 2.

Lemma 8. *For $1 \leq i \leq n$ let G_i be an $(n-i)$ -truncated t -geometric random variable, and suppose that G_1, \dots, G_n are independent. Then the law of $\mathcal{D}(G_1, \dots, G_n)$ is the Mallows measure on S_n with parameter t .*

Proof. It follows from Lemma 7 that for all $\sigma \in S_n$,

$$\mathbb{P}(\mathcal{D}(G_1, \dots, G_n) = \sigma) = \mathbb{P}((G_1, \dots, G_n) = \mathcal{L}(\sigma)).$$

By the independence of G_1, \dots, G_n , the right side equals

$$\prod_{i=1}^n \mathbb{P}(G_i = \mathcal{L}(\sigma)_i).$$

The result now follows since $\sum_{i=1}^n \mathcal{L}(\sigma)_i = \text{inv}(\sigma)$. \square

Having discussed permutations and their encodings, we now relate these objects to words. Fix a (possibly infinite) interval I of \mathbb{Z} and an integer $q \geq 1$. A **word** $x = (x_i)_{i \in I}$ indexed by I is a function from I to $\llbracket 1, q \rrbracket$, and its entries are referred to as **characters**. It is **proper** if $x_i \neq x_{i+1}$ whenever $i, i+1 \in I$. A **q -coloring** is defined to be a proper word. The **length** of a word, denoted by $|x|$, is the cardinality of its index set I . We denote the empty word by \emptyset . The **concatenation** of a word x indexed by $\llbracket a, b \rrbracket$ with a word y indexed by $\llbracket b, c \rrbracket$ is the word xy indexed by $\llbracket a, c \rrbracket$ whose restrictions to $\llbracket a, b \rrbracket$ and $\llbracket b, c \rrbracket$ are x and y , respectively. Similar notation is used for concatenations of words with individual characters. Given a word x indexed by I and a set $A \subseteq I$ of size m , the **subword** $(x_i: i \in A)$ is defined to be the word $x_{i_1} x_{i_2} \cdots x_{i_m}$, where i_k is the k^{th} smallest element of A . In other words, it is the subsequence of x indexed by A .

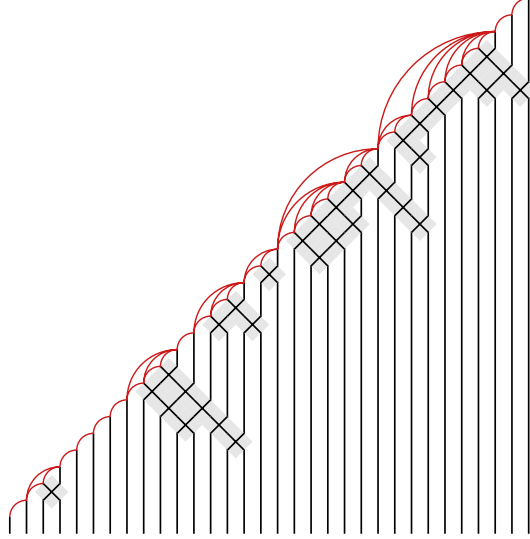


FIGURE 2. This figure depicts a permutation, its Lehmer code, and its constraint graph. Index the wires from left to right along their bottom end via a finite interval $I \subset \mathbb{Z}$. Let $\sigma \in \text{Sym}(I)$ be the permutation obtained by following the wires upwards. For $i \in I$, the quantity $\mathcal{L}(\sigma)_i$ (the Lehmer code of σ at i) is equal to the number of crosses emanating down and to the right from the point on the diagonal directly above the bottom of wire i . The constraint graph Γ_σ is drawn above the diagonal.

Following [30], we say that a permutation σ of I is a **proper building** of a word x if for each $t \in I$ the subword

$$x^\sigma(t) := (x_i : \sigma(i) \leq t)$$

of x is proper. We write $\sigma \vdash x$ if this occurs. Note that x is proper if and only if it has some proper building, in which case for instance the identity permutation is a proper building.

The following picture will be very useful. We regard $\sigma(i)$ as the arrival time of i . Then $x^\sigma(t)$ is the subword of x that has arrived by time t . At time step t , the integer $\sigma^{-1}(t)$ arrives, and the character $x_{\sigma^{-1}(t)}$ is inserted into $x^\sigma(t-1)$ (or the empty word, if $t = \min I$). The insertion code $\tilde{\mathcal{L}}(\sigma)_t$ has a natural interpretation in terms of arrivals, from which its name derives: it is the distance from the right at which $x_{\sigma^{-1}(t)}$ is inserted in $x^\sigma(t-1)$. More precisely, for all $t > \min I$ there are subwords u and v of x such that

$$x^\sigma(t-1) = uv, \quad x^\sigma(t) = ux_{\sigma^{-1}(t)}v, \quad \text{and} \quad |v| = \tilde{\mathcal{L}}(\sigma)_t.$$

For example if $\sigma = 25431$ and $t = 4$, then

$$x^\sigma(t-1) = x_1x_4x_5, \quad x_{\sigma^{-1}(t)} = x_3, \quad x^\sigma(t) = x_1x_3x_4x_5,$$

and $\tilde{\mathcal{L}}(\sigma)_t = |x_4x_5| = 2$.

The condition $\sigma \vdash x$ can also be expressed in the language of graphs. Let I be the index set of x . Then for a graph G with vertex set I , we say that x is a **proper coloring** of G if $x_i \neq x_j$ whenever i and j are adjacent in G . Given a permutation σ of I , we define its **constraint graph** Γ_σ to have vertex set I and an undirected edge between i and j , where $i < j$, if and only if $\sigma(i) < \sigma(k) > \sigma(j)$ for all $k \in (i, j)$. In other words, two integers are adjacent in Γ_σ if and only if they both arrive prior to any integer between them. It is immediate from the definitions that $\sigma \vdash x$ if and only if x is a proper coloring of Γ_σ . See Figure 2 for an example of a constraint graph.

Next we consider decompositions of graphs that arise naturally in the context of buildings. Let G be a graph whose vertex set I is a (possibly infinite) interval of \mathbb{Z} . An integer $i \in I$ is a **bubble endpoint** of G if there do not exist j and k with $j < i < k$ such that j and k are adjacent in G . A **bubble** of G is a subgraph induced by a finite interval of \mathbb{Z} whose endpoints are consecutive bubble endpoints, and $\text{bub}(G)$ denotes the set of all bubbles. Note that every endpoint of I is a bubble endpoint.

Lemma 9. *Let G be a graph whose vertex set I is a (possibly infinite) interval of \mathbb{Z} . Then $G = \bigcup \text{bub}(G)$ iff the infimum and supremum of the set of bubble endpoints agree with those of I .*

The proof is immediate. Note that the condition on the set of bubble endpoints holds automatically if I is finite.

Next we consider bubbles of constraint graphs of permutations. An integer i is a **record** of a permutation σ if it is either the maximum or minimum of the set

$$\{\sigma(j) : j \leq \sigma^{-1}(i)\}.$$

Records are a well-studied permutation statistic, both combinatorially [10] and probabilistically [5, 15, 25, 40]. A **founder** of a permutation σ is defined to be a record of σ^{-1} . Equivalently, i is a founder of σ iff there do not exist j and k with $j < i < k$ and $\sigma(j) < \sigma(i) > \sigma(k)$. Phrased in terms of the arrival times picture, i is a founder if and only if either:

- it arrives prior to all smaller elements of I , or
- it arrives prior to all larger elements of I .

We write $\mathcal{F}(\sigma)$ for the set of founders of σ .

Lemma 10. *Let σ be a permutation of an interval $I \subseteq \mathbb{Z}$ (which may be finite or infinite). Then the set of bubble endpoints of the constraint graph Γ_σ is $\mathcal{F}(\sigma)$. Furthermore if $I = \llbracket 0, n \rrbracket$, then $\#\mathcal{F}(\sigma) = \#\{i \in I: \tilde{\mathcal{L}}(\sigma)_i \in \{0, i\}\}$.*

In particular, this implies that $\#\text{bub}(\Gamma_\sigma)$ is one less than the number of founders of σ (which could be infinite for a permutation of an infinite interval). Notation involving Lehmer codes often simplifies when working on the interval $\llbracket 0, n \rrbracket$. Note that this interval has $n + 1$ elements.

Proof of Lemma 10. If i is not a bubble endpoint of Γ_σ , then there must exist $j < i < k$ such that j and k arrive prior to all elements of $\llbracket j, k \rrbracket$, and in particular i . Thus i is not a founder. Conversely if i is not a founder, then there exist $j < i < k$ with $\sigma(j) < \sigma(i) > \sigma(k)$. Now choose j maximal and k minimal satisfying these conditions to obtain an edge of Γ_σ passing over i , showing that i is not a bubble endpoint. This establishes the first claim.

For the second claim, observe that

$$\mathcal{F}(\sigma) = \{i \in I: \tilde{\mathcal{L}}(\sigma)_{\sigma(i)} \in \{0, \sigma(i)\}\} = \sigma(\{i \in I: \tilde{\mathcal{L}}(\sigma)_i \in \{0, i\}\}).$$

Thus $\#\mathcal{F}(\sigma) = \#\{i \in I: \tilde{\mathcal{L}}(\sigma)_i \in \{0, i\}\}$. □

Let $\text{Col}_q(G)$ be the number of proper q -colorings of a graph G .

Lemma 11. *For any $q \geq 3$, any finite interval I of \mathbb{Z} , and any permutation σ of I ,*

$$\text{Col}_q(\Gamma_\sigma) = q(q-2)^{\#I-1} \left(\frac{q-1}{q-2}\right)^{\#\text{bub}(\Gamma_\sigma)}.$$

Proof. Without loss of generality assume that $I = \llbracket 0, n \rrbracket$. For each $i \in \llbracket 0, n \rrbracket$, let Γ_σ^i denote the subgraph of Γ_σ induced by the set of the first i vertices to arrive, $\sigma^{-1}(\llbracket 0, i \rrbracket)$. The graph Γ_σ^1 can be colored in q ways. Suppose x is a proper coloring of Γ_σ^i . If $i + 1$ is a record of σ^{-1} , then $i + 1$ has degree 1 in Γ_σ^{i+1} , and there are exactly $q - 1$ colorings of Γ_σ^{i+1} extending x . Otherwise, $i + 1$ has degree 2 in Γ_σ^{i+1} and, since both of the neighbors of $i + 1$ in Γ_σ^i have different colors in x , there are exactly $q - 2$ proper colorings of Γ_σ^{i+1} extending x . Thus by Lemma 10,

$$\text{Col}_q(\Gamma_\sigma^{i+1}) = \begin{cases} (q-1) \text{Col}_q(\Gamma_\sigma^i), & \text{if } i+1 \text{ is a record of } \sigma^{-1} \\ (q-2) \text{Col}_q(\Gamma_\sigma^i), & \text{otherwise.} \end{cases}$$

Since the records of σ^{-1} are the bubble endpoints of $\text{bub}(\Gamma_\sigma)$, the lemma follows. □

A key element of the proof of our main theorem is a joint probability measure on colorings and permutations of a finite interval. The marginal law of the coloring will provide the finite-dimensional distributions for our coloring of \mathbb{Z} . The marginal law of the permutation will

belong to the following family of permutation measures. Let $I \subset \mathbb{Z}$ be a finite interval and let u and t be non-negative real parameters. The **bubble-biased Mallows measure** with parameters t and u is the probability measure $\text{BMal} = \text{BMal}_{t,u} = \text{BMal}_{t,u}^I$ on $\text{Sym}(I)$ given by

$$\text{BMal}(\{\sigma\}) = \frac{u^{\#\text{bub}(\Gamma_\sigma)} t^{\text{inv}(\sigma)}}{\sum_{\tau \in \text{Sym}(I)} u^{\#\text{bub}(\Gamma_\tau)} t^{\text{inv}(\tau)}}, \quad \sigma \in \text{Sym}(I). \quad (5)$$

Since $\#\text{bub}(\Gamma_\sigma)$ is one less than the number of records of σ^{-1} , we also have that

$$\text{BMal}(\{\sigma\}) = \frac{u^{\#\{\text{records of } \sigma^{-1}\}} t^{\text{inv}(\sigma)}}{\sum_{\tau \in \text{Sym}(I)} u^{\#\{\text{records of } \tau^{-1}\}} t^{\text{inv}(\tau)}}, \quad \sigma \in \text{Sym}(I).$$

The key property of BMal is that, for $q \geq 3$ and $u = \frac{q-1}{q-2}$, we have that

$$\text{BMal}_{t,u}(\{\sigma\}) = \frac{\text{Col}_q(\Gamma_\sigma) t^{\text{inv}(\sigma)}}{\sum_{\tau \in \text{Sym}(I)} \text{Col}_q(\Gamma_\tau) t^{\text{inv}(\tau)}}, \quad \sigma \in \text{Sym}(I), \quad (6)$$

which follows by combining Lemma 11 with (5). That is, BMal is the law of a Mallows random permutation biased by the number of proper q -colorings of its constraint graph. We extend the definition of the Mallows measure from permutations of $\llbracket 1, n \rrbracket$ to permutations of an arbitrary finite interval I by declaring it to be the special case $u = 1$ of $\text{BMal}_{t,u}$. We remark that, even though $\text{Mal}_t(\sigma) = \text{Mal}_t(\sigma^{-1})$ for all σ , the quantities $\text{BMal}_{t,u}(\sigma)$ and $\text{BMal}_{t,u}(\sigma^{-1})$ differ in general when $u \neq 1$. This fact adds significant complications to our proof of the main theorem.

Lemma 12. *For all $i \in \llbracket 0, n \rrbracket$, let G_i be a u -end-weighted, i -truncated, t -geometric random variable, and suppose that G_0, \dots, G_n are independent. Then the law of $\tilde{\mathcal{G}}(G_0, \dots, G_n)$ is $\text{BMal}_{t,u}^{\llbracket 0, n \rrbracket}$.*

Proof. It follows from Lemma 7 that for all $\sigma \in \text{Sym}(\llbracket 0, n \rrbracket)$,

$$\mathbb{P}(\tilde{\mathcal{G}}(G_0, \dots, G_n) = \sigma) = \mathbb{P}((G_0, \dots, G_n) = \tilde{\mathcal{L}}(\sigma)).$$

By the independence of G_0, \dots, G_n , the right side equals

$$\prod_{i=0}^n \mathbb{P}(G_i = \tilde{\mathcal{L}}(\sigma)_i) = \prod_{i=0}^n \frac{u^{1[\tilde{\mathcal{L}}(\sigma)_i \in \{0, i\}]} t^{\tilde{\mathcal{L}}(\sigma)_i}}{u + t + \dots + ut^{i-1} + ut^i}.$$

The exponent of u in this product is $\#\mathcal{F}(\sigma)$ by Lemma 10, and the exponent of t is $\text{inv}(\sigma)$.

Thus

$$\mathbb{P}(\tilde{\mathcal{G}}(G_0, \dots, G_n) = \sigma) = \frac{u^{\#\mathcal{F}(\sigma)} t^{\text{inv}(\sigma)}}{\prod_{i=0}^n (u + t + \dots + ut^{i-1} + ut^i)}.$$

The result now follows since $\#\mathcal{F}(\sigma) = \#\text{bub}(\Gamma_\sigma) + 1$. \square

3 Finite dependence

The purpose of this section is to construct random colorings of \mathbb{Z} that are finitely dependent by starting on finite intervals and appealing to Kolmogorov extension.

Proposition 13. *Fix $t \in [0, 1]$ and $q \geq 3$. Let $u = \frac{q-1}{q-2}$. Then there is a random q -coloring of \mathbb{Z} such that for every finite interval I , its restriction to I has the law of a uniform q -coloring of the constraint graph of a $\text{BMal}_{t,u}$ -distributed permutation of I .*

The random coloring is strictly k -dependent iff (q, k, t) satisfies the tuning equation

$$qt[k]_t = [2]_t[k+1]_t. \quad (7)$$

Note that the equation (7) is equivalent to (1) provided $t \neq 1$.

Consider the probability measure $\text{Joint} = \text{Joint}_{t,q,n}$ on $S_n \times \llbracket 1, q \rrbracket^n$ given by

$$\text{Joint}\left(\{(\sigma, x)\}\right) = \frac{\mathbf{1}[\sigma \vdash x]t^{\text{inv}(\sigma)}}{Z(t, q, n)},$$

where the normalizing constant $Z(t, q, n)$ is

$$Z(t, q, n) = \sum_{\sigma \in S_n} \sum_{x \in \llbracket 1, q \rrbracket^n} \mathbf{1}[\sigma \vdash x]t^{\text{inv}(\sigma)} = \sum_{\sigma \in S_n} \text{Col}_q(\sigma)t^{\text{inv}(\sigma)}.$$

The permutation marginal of Joint is BMal , by (6). If the random pair (σ, x) has law Joint , then the conditional law of x given σ is the uniform measure on proper q -colorings of Γ_σ . We denote the marginal probability mass function of x by $P^{\text{col}} = P_{t,q,n}^{\text{col}}$,

$$P^{\text{col}}(x) = \sum_{\sigma \in S_n} \frac{\mathbf{1}[\sigma \vdash x]t^{\text{inv}(\sigma)}}{Z(t, q, n)}. \quad (8)$$

Lemma 14. *Let $t \in [0, 1]$ and $q \geq 3$. There exists a measure $\text{MalCol} = \text{MalCol}_{q,t}$ on $\llbracket 1, q \rrbracket^{\mathbb{Z}}$ such that if $X = (X_i)_{i \in \mathbb{Z}}$ is random with law MalCol , then X is stationary and*

$$\mathbb{P}[(X_{i+1}, \dots, X_{i+n}) = x] = P_{t,q,n}^{\text{col}}(x)$$

for all $i \in \mathbb{Z}$, for all $n \geq 0$, and for all $x \in \llbracket 1, q \rrbracket^n$.

The random colorings of \mathbb{Z} which we construct in the proof of the main theorem have law $\text{MalCol}_{q,t}$ for certain values of t .

By inspection of (8), observe that if σ is a Mallows random permutation then

$$P^{\text{col}}(x) = \frac{[n]_t!}{Z(t, q, n)} \mathbb{P}(\sigma \vdash x), \quad x \in \llbracket 1, q \rrbracket^n. \quad (9)$$

This characterization of P^{col} will be used in Section 4 to prove reversibility.

The **building number** $B_t(x)$ is the unnormalized version of $P_{t,q,n}^{\text{col}}$ given by

$$B_t(x) := \sum_{\sigma \in \mathcal{S}_n} \mathbf{1}[\sigma \vdash x] t^{\text{inv}(\sigma)} = Z(t, q, n) P^{\text{col}}(x), \quad x \in \llbracket 1, q \rrbracket^n.$$

This specializes when $t = 1$ to the number of proper buildings, which was a key player in the earlier construction of [30]. Observe that

$$\sum_{x \in \llbracket 1, q \rrbracket^n} B_t(x) = Z(t, q, n).$$

and that $B_t(\emptyset) = 1$.

The reason we use $B_t(x)$ (rather than using P^{col} directly) is that it satisfies simpler recurrences, as we will now see. We abbreviate a word $x = (x_i)_{i \in \llbracket 1, n \rrbracket}$ by writing $x = x_1 \cdots x_n$, and we use the notation $\widehat{x}_i := x_1 \cdots x_{i-1} x_{i+1} \cdots x_n$.

Lemma 15. *For all $n \geq 1$, all words $x \in \llbracket 1, q \rrbracket^n$, and all real $t \geq 0$ we have*

$$B_t(x) = \mathbf{1}[x \text{ is proper}] \sum_{i=1}^n t^{n-i} B_t(\widehat{x}_i), \quad (10)$$

and

$$B_t(x) = \sum_{i=1}^n t^{n-i} B_t(\widehat{x}_i) - [2]_t \sum_{j=2}^n \mathbf{1}[x_{j-1} = x_j] t^{n-j} B_t(\widehat{x}_j). \quad (11)$$

Equation (10) is a t -analogue of [30, Prop. 9]. The variant recurrence (11) (which was not used in [30]) simplifies a large amount of casework. As an alternative to the proof below, one may deduce (11) from (10) via the Möbius Inversion Formula for posets [55, Section 3.7].

Proof of Lemma 15. To prove equation (10), observe that the permutation σ is a proper building of x with $\sigma^{-1}(n) = i$ if and only if x is proper and the permutation $\widehat{\sigma}_i := \sigma_1 \cdots \sigma_{i-1} \sigma_{i+1} \cdots \sigma_n \in \mathcal{S}_{n-1}$ is a proper building of \widehat{x}_i . Now (10) follows from the easy observation that $\text{inv}(\sigma) = \text{inv}(\widehat{\sigma}_i) + n - i$.

To establish (11), write $\mathbf{1}[x \text{ is proper}]$ as $\mathbf{1}[x_1 \neq x_2] \cdots \mathbf{1}[x_{n-1} \neq x_n]$. Then by (10),

$$B_t(x) = \prod_{j=2}^n (1 - \mathbf{1}[x_{j-1} = x_j]) \sum_{i=1}^n t^{n-i} B_t(\widehat{x}_i). \quad (12)$$

Next observe that for any pair of distinct indices $i \neq j$, the expression

$$\mathbf{1}[x_{j-1} = x_j] \mathbf{1}[x_{i-1} = x_i] B_t(\widehat{x}_k)$$

vanishes for all k . Indeed, any word x with $x_{j-1} = x_j$ and $x_{i-1} = x_i$ must still have adjacent repeated indices even after deleting an arbitrary symbol, and so the resulting word has no proper buildings. Expanding (12) and discarding such terms,

$$B_t(x) = \left(1 - \sum_{j=2}^n \mathbf{1}[x_{j-1} = x_j]\right) \sum_{i=1}^n t^{n-i} B_t(\hat{x}_i).$$

If the expression $\mathbf{1}[x_{j-1} = x_j] B_t(\hat{x}_k)$ is non-zero, then $k \in \{j-1, j\}$ and $\hat{x}_k = \hat{x}_j$. Thus

$$\begin{aligned} B_t(x) &= \sum_{i=1}^n t^{n-i} B_t(\hat{x}_i) - \sum_{j=2}^n \mathbf{1}[x_{j-1} = x_j] \sum_{i=1}^n t^{n-i} B_t(\hat{x}_i) \\ &= \sum_{i=1}^n t^{n-i} B_t(\hat{x}_i) - \sum_{j=2}^n \mathbf{1}[x_{j-1} = x_j] (t^{n-j} + t^{n-j+1}) B_t(\hat{x}_j). \end{aligned}$$

Factoring out $[2]_t = t + 1$ from the second term in the latter expression yields (11). \square

Using these recurrences, we show that the marginals of MalCol are consistent.

Proposition 16 (Consistency). *For all $x \in \llbracket 1, q \rrbracket^n$ and all $a \in \llbracket 1, q \rrbracket$ we have that*

$$\sum_{a \in \llbracket 1, q \rrbracket} P_{t,q,n+1}^{\text{col}}(ax) = \sum_{a \in \llbracket 1, q \rrbracket} P_{t,q,n+1}^{\text{col}}(xa) = P_{t,q,n}^{\text{col}}(x).$$

In the following proof and for the remainder of this section, we write \star to denote a dummy variable that is summed over $\llbracket 1, q \rrbracket$. For example, given a function $f: \llbracket 1, q \rrbracket^k \rightarrow \mathbb{R}$ and a word x , we write $f(x\star^k)$ as a shorthand for $\sum_{y \in \llbracket 1, q \rrbracket^k} f(xy)$.

Proof. To prove that $P_{t,q,n+1}^{\text{col}}(\star x) = P_{t,q,n}^{\text{col}}(x)$, we establish by induction on n that

$$B_t(x\star) = (q[n+1]_t - [2]_t[n]_t) B_t(x), \quad x \in \llbracket 1, q \rrbracket^n. \quad (13)$$

This is clear when $n = 0$. Let x be a word of length n , and suppose that (13) holds for all words of length at most $n-1$. Also suppose that x is proper, for otherwise (13) is trivial.

Applying equation (11) from Lemma 15, we see that for any $a \in \llbracket 1, q \rrbracket$ we have that

$$B_t(xa) = t \sum_{i=1}^n t^{n-i} B_t(\hat{x}_i a) + B_t(x) - \mathbf{1}[x_n = a] [2]_t B_t(x). \quad (14)$$

Summing over all $a \in \llbracket 1, q \rrbracket$ yields that

$$B_t(x\star) = t \sum_{i=1}^n t^{n-i} B_t(\hat{x}_i \star) + q B_t(x) - [2]_t B_t(x).$$

Hence by the inductive hypothesis

$$B_t(x\star) = t(q[n]_t - [2]_t[n-1]_t) \sum_{i=1}^n t^{n-i} B_t(\hat{x}_i) + qB_t(x) - [2]_t B_t(x). \quad (15)$$

By (10) we have that $\sum_{i=1}^n t^{n-i} B_t(\hat{x}_i) = B_t(x)$. Substituting this into (15) and using the trivial identity $t[n]_t + 1 = [n+1]_t$, we deduce (13).

Combining (13) with (3) yields that $P_{t,q,n+1}^{\text{col}}(x\star) = P_{t,q,n}^{\text{col}}(x)$. By an analogous argument, we have that $P_{t,q,n+1}^{\text{col}}(\star x) = P_{t,q,n}^{\text{col}}(x)$ as well. \square

Lemma 14 now follows easily.

Proof of Lemma 14. The family of cylinder measures in the statement of the lemma is consistent, by Proposition 16. Thus by the Kolmogorov extension theorem [38] there exists a random coloring, X , for which (14) holds. Stationarity of X is immediate. \square

Next we derive an expression for the normalizing constant $Z(t, q, n)$. When (q, k, t) satisfies the tuning equation (7), extra simplifications occur, as can be seen already in the following lemma.

Lemma 17. *For all integers $n \geq 1$ we have that*

$$Z(t, q, n) = \prod_{j=1}^n (q[j]_t - [2]_t[j-1]_t). \quad (16)$$

Moreover, when $qt[k]_t = [2]_t[k+1]_t$, equation (16) may be rewritten as

$$Z(t, q, n) = [n]_t! \left(\frac{q}{[k+1]_t} \right)^n \binom{k+n}{k}_t. \quad (17)$$

Proof. It follows from (13) that if x is a word of length n then $B_t(x\star) = (q[n+1]_t - [2]_t[n]_t)B_t(x)$. Summing over $x \in [1, q]^n$ yields that

$$Z(t, q, n+1) = (q[n+1]_t - [2]_t[n]_t)Z(t, q, n).$$

Now a simple induction establishes (16).

Next, suppose that (q, k, t) satisfies the tuning equation (7). We show that for every $j \geq 1$,

$$(q[j]_t - [2]_t[j-1]_t)[k+1]_t = q[k+j]_t. \quad (18)$$

Indeed, the tuning equation allows us to substitute $qt[k]_t$ in place of $[2]_t[k+1]_t$. Furthermore it is easy to see that $[k+j]_t = [j]_t + t^j[k]_t$. Thus (18) reduces to

$$q[j]_t[k+1]_t - qt[k]_t[j-1]_t = q[j]_t + qt^j[k]_t,$$

i.e., $q[j]_t([k+1]_t - 1) = q[k]_t(t^j + t[j-1]_t)$, which is apparent since both quantities simplify to $qt[j]_t[k]_t$.

Now (17) follows from (16) by using (18) to rewrite each factor of the product. \square

Proposition 18 (*k-dependence*). *Suppose that the integers $q \geq 3$ and $k \geq 1$ and the real number $t \geq 0$ satisfy the tuning equation $qt[k]_t = [2]_t[k+1]_t$ (7). Then for all $x \in \llbracket 1, q \rrbracket^m$ and $y \in \llbracket 1, q \rrbracket^n$ we have that*

$$\sum_{a \in \llbracket 1, q \rrbracket^k} P^{\text{col}}(xay) = P^{\text{col}}(x)P^{\text{col}}(y). \quad (19)$$

This proposition implies that a coloring with law MalCol is k -dependent whenever (q, k, t) satisfies the tuning equation (7). In the proof, we use a well-known t -binomial coefficient identity appearing in [55, eq. (17b)] which states that, for integers r and s ,

$$\binom{r}{s}_t = \binom{r-1}{s}_t + t^{r-s} \binom{r-1}{s-1}_t. \quad (20)$$

Proof of Proposition 18. We show that for $x \in \llbracket 1, q \rrbracket^m$ and $y \in \llbracket 1, q \rrbracket^n$,

$$B_t(x \star^k y) = [k]_t! \left(\frac{q}{[k+1]_t} \right)^k \binom{m+n+2k}{m+k}_t B_t(x)B_t(y). \quad (21)$$

By normalizing both sides of (21), it will follow that $P^{\text{col}}(x \star^k y) = c_{m,k,n} P^{\text{col}}(x)P^{\text{col}}(y)$ for some constant $c_{m,k,n}$. But both sides are probability mass functions, so $c_{m,k,n} = 1$ and thus the lemma follows directly from (21).

We prove (21) by induction on m and n . The case $m = n = 0$ follows from the special case $n = k$ of (17) in Lemma 17.

Suppose that (21) holds for all words x and y with lengths $m-1$ and n respectively, and for all words x and y with lengths m and $n-1$ respectively. Furthermore, suppose that x and y are proper, since the desired result holds trivially otherwise. Let $a = a_1 a_2 \cdots a_k$ denote a word; a will be summed over $\llbracket 1, q \rrbracket^k$ below. Applying equation (11) of Lemma 15 yields that

$$\begin{aligned} B_t(xay) &= \sum_{i=1}^m t^{m+k+n-i} B_t(\widehat{x}_i ay) + \sum_{i=1}^k t^{k+n-i} B_t(x \widehat{a}_i y) + \sum_{i=1}^n t^{n-i} B_t(xa \widehat{y}_i) \\ &\quad - [2]_t \mathbf{1}[x_m = a_1] t^{k+n-1} B_t(x \widehat{a}_1 y) \\ &\quad - [2]_t \sum_{i=1}^{k-1} \mathbf{1}[a_i = a_{i+1}] t^{k+n-i-1} B_t(x \widehat{a}_i y) \\ &\quad - [2]_t \mathbf{1}[a_k = y_1] t^{n-1} B_t(x \widehat{a}_k y). \end{aligned}$$

Summing over all $a \in \llbracket 1, q \rrbracket^k$ implies that

$$\begin{aligned}
B_t(x \star^k y) &= \sum_{i=1}^m t^{m+k+n-i} B_t(\widehat{x}_i \star^k y) + q \sum_{i=1}^k t^{k+n-i} B_t(x \star^{k-1} y) + \sum_{i=1}^n t^{n-i} B_t(x \star^k \widehat{y}_i) \\
&\quad - [2]_t t^{k+n-1} B_t(x \star^{k-1} y) \\
&\quad - [2]_t \sum_{i=1}^{k-1} t^{k+n-i-1} B_t(x \star^{k-1} y) \\
&\quad - [2]_t t^{n-1} B_t(x \star^{k-1} y) \\
&= t^{k+n} \sum_{i=1}^m t^{m-i} B_t(\widehat{x}_i \star^k y) + \sum_{i=1}^n t^{n-i} B_t(x \star^k \widehat{y}_i) \\
&\quad + (qt^m [k]_t - t^{m-1} [2]_t [k+1]_t) B_t(x \star^{k-1} y). \tag{22}
\end{aligned}$$

Note that we have not yet used the assumption that (q, k, t) satisfies the tuning equation (7). Crucially, the coefficient of $B_t(x \star^{k-1} y)$ in (22) vanishes when the tuning equation is satisfied, so that

$$B_t(x \star^k y) = t^{k+n} \sum_{i=1}^m t^{m-i} B_t(\widehat{x}_i \star^k y) + \sum_{i=1}^n t^{n-i} B_t(x \star^k \widehat{y}_i). \tag{23}$$

By the inductive hypothesis and equation (10) from Lemma 15, (23) expands to

$$[k]_t! \left(\frac{q}{[k+1]_t} \right)^k \left[t^{k+n} \binom{m+n+2k-1}{m+k-1}_t + \binom{m+n+2k-1}{m+k}_t \right] B_t(x) B_t(y).$$

Using (20), the above expression simplifies to

$$B_t(x \star^k y) = [k]_t! \left(\frac{q}{[k+1]_t} \right)^k \binom{m+n+2k}{m+k}_t B_t(x) B_t(y). \quad \square$$

The final result we will need for Proposition 13 is a converse to the previous lemma.

Lemma 19. *Let $q \geq 3$ and $t \geq 0$ be given. Suppose that k is a number such that, for all $m \geq 0$ and $n \geq 0$ and for all words $x \in \llbracket 1, q \rrbracket^m$ and $y \in \llbracket 1, q \rrbracket^n$,*

$$\sum_{a \in \llbracket 1, q \rrbracket^k} P^{\text{col}}(xay) = P^{\text{col}}(x) P^{\text{col}}(y).$$

Then there exists an integer $k' \leq k$ such that (q, k', t) satisfies the tuning equation (7).

Proof. From (22) in the proof of Proposition 18, if x is a word of length m and y is a word of length n then

$$\begin{aligned}
B_t(x \star^k y) &= t^{k+n} \sum_{i=1}^m t^{m-i} B_t(\widehat{x}_i \star^k y) + \sum_{i=1}^n t^{n-i} B_t(x \star^k \widehat{y}_i) \\
&\quad + (qt^m [k]_t - t^{m-1} [2]_t [k+1]_t) B_t(x \star^{k-1} y),
\end{aligned}$$

which holds for all q, k , and t . In particular, taking x and y to be words of length one, and subtracting two instances of the last equation yields

$$B_t(1 \star^k 2) - B_t(1 \star^k 1) = (qt^k[k]_t - t^{k-1}[2]_t[k+1]_t) (B_t(1 \star^{k-1} 2) - B_t(1 \star^{k-1} 1)).$$

Upon iterating this identity, we obtain that

$$B_t(1 \star^k 2) - B_t(1 \star^k 1) = (B_t(12) - B_t(11)) \prod_{k'=1}^k (qt^{k'}[k']_t - t^{k'-1}[2]_t[k'+1]_t).$$

Under our hypotheses on k , the left side of the previous equation vanishes. Since $B_t(12) = t+1$ is non-zero but $B_t(11) = 0$, one of the factors in the product on the right vanishes. \square

We conclude this section by proving the proposition stated at the very beginning.

Proof of Proposition 13. We show that, for all $t \in [0, 1]$ and all $q \geq 3$, the measure $\text{MalCol}_{q,t}$ is the law of a random coloring $X = (X_i)_{i \in \mathbb{Z}}$ satisfying all conditions in the proposition. Namely,

- for each finite interval I , the restricted coloring $(X_i)_{i \in I}$ is equal in law to a uniform q -coloring of the constraint graph of a $\text{BMal}_{t,u}$ -distributed permutation of I , and
- the coloring X is strictly k -dependent if and only if (q, k, t) satisfies the tuning equation $qt[k]_t = [2]_t[k+1]_t$ (7).

The claims in the first bullet follow from Lemma 14 and the comments involving Joint preceding that lemma. The second bullet follows by combining Lemma 14 with Proposition 18 and Lemma 19. \square

4 Reversibility

The primary purpose of this section is to prove the following.

Proposition 20. *For $q \geq 3$ and $t \in [0, 1]$, if $(X_i)_{i \in \mathbb{Z}}$ has law $\text{MalCol}_{q,t}$, then so does $(X_{-i})_{i \in \mathbb{Z}}$.*

This proposition plays an auxiliary role in the proof of Theorem 1, and its proof is technical. Readers eager for the proof of the main theorem may safely proceed to the next section.

The secondary purpose of this section is to justify the following claim from the introduction. In fact, it is a corollary of the previous result.

Corollary 21. *For each $q \geq 4$, there exists a unique $t \in [0, 1]$ such that $(q, 1, t)$ satisfies (7). For this t the symmetric 1-dependent q -colorings from [31] have law $\text{MalCol}_{q,t}$.*

Throughout this section we fix a finite interval I of \mathbb{Z} and a word x indexed by I . Recall from Section 2 that a permutation σ of I is a proper building of x if and only if each of the subwords

$$x^\sigma(t) = (x_i : \sigma(i) \leq t), \quad t \in I$$

is proper. The word $x^\sigma(t)$ is obtained from $x^\sigma(t-1)$ by inserting $x_{\sigma^{-1}(t)}$ at position $\tilde{\mathcal{L}}(\sigma)_t$ from the right.

For $i, j \in I$ we write $(i \ j)$ for the permutation transposing i and j . We also define $\Delta_k(\sigma) := \tilde{\mathcal{L}}(\sigma)_{k+1} - \tilde{\mathcal{L}}(\sigma)_k$.

Lemma 22. *Let I be a finite interval of \mathbb{Z} and let k be an integer such that $k, k+1 \in I$. Let δ be an integer and let $\ell = (\ell_i : i \in I \setminus \{k, k+1\})$ be a sequence. Let $A_{\delta, \ell}$ be the set of permutations σ of I such that restriction of $\tilde{\mathcal{L}}(\sigma)$ to $I \setminus \{k, k+1\}$ is ℓ and $\Delta_k(\sigma) = \delta$. Then for all words $x \in \llbracket 1, q \rrbracket^I$,*

$$\#\{\sigma \in A_{\delta, \ell} : \sigma \vdash x\} = \#\{\sigma \in A_{\delta, \ell} : (k \ k+1) \circ \sigma \vdash x\}. \quad (24)$$

We remark that if it were the case that for all graphs G

$$\#\{\sigma \in A_{\delta, \ell} : \Gamma_\sigma = G\} = \#\{\sigma \in A_{\delta, \ell} : \Gamma_{(k \ k+1) \circ \sigma} = G\}, \quad (25)$$

then (24) would follow immediately, since $\sigma \vdash x$ is equivalent to the assertion that x is a proper coloring of the graph Γ_σ . As we will see in the proof, (25) does hold in many cases but not all. For example, it does not hold in the case $I = \llbracket 1, 4 \rrbracket$, $k = 3$, and $\delta = \ell_1 = \ell_2 = 0$ with the graph

$$G = (I, E), \quad \text{where} \quad E = \{(1, 2), (2, 3), (1, 3), (3, 4), (1, 4)\}.$$

We will use the following result in the proof of the lemma.

Lemma 23. *Let I be a finite interval, let k be an integer such that $\min I \leq k < \max I$, and let σ be a permutation of I . Then*

$$\begin{aligned} \tilde{\mathcal{L}}((k \ k+1) \circ \sigma) &= (\dots, \ell_{k-1}, \ell_{k+1} - \mathbf{1}[\ell_{k+1} > \ell_k], \ell_k + \mathbf{1}[\ell_{k+1} \leq \ell_k], \ell_{k+2}, \dots), \\ \text{where } \tilde{\mathcal{L}}(\sigma) &= (\dots, \ell_{k-1}, \ell_k, \ell_{k+1}, \ell_{k+2}, \dots). \end{aligned}$$

In particular, $\Delta_k((k \ k+1) \circ \sigma) = 1 - \Delta_k(\sigma)$.

The proof is straightforward and is omitted.

Proof of Lemma 22. Fix x , k , δ , and ℓ . For any permutation σ of I and any set S of such permutations, we write

$$\sigma' := (k \ k+1) \circ \sigma, \quad S' = \{\sigma' : \sigma \in S\}.$$

Note that $\sigma'' = \sigma$ and $\#S = \#S'$. The only difference between σ and σ' is that the integers arriving at times k and $k+1$ are interchanged. In particular, $x^\sigma(t) = x^{\sigma'}(t)$ for all $t \neq k$. Let E be the set of permutations of I that are proper buildings of x . Using this notation, (24) may be expressed as

$$\#A_{\delta,\ell} \cap E = \#A_{\delta,\ell} \cap E'. \quad (26)$$

Say that a permutation σ is **almost proper** (at k , with respect to x) if $x^\sigma(t)$ is a proper word for all $t \neq k$, and let P be the set of almost proper permutations. Observe that $E \subseteq P$ and that $P' = P$, from which it follows that $E' \subseteq P$ as well.

Since $\#A_{\delta,\ell} \cap (P \setminus E) = \#A_{\delta,\ell} \cap P - \#A_{\delta,\ell} \cap E$ and similarly with E' in place of E , (26) is equivalent to

$$\#A_{\delta,\ell} \cap (P \setminus E) = \#A_{\delta,\ell} \cap (P \setminus E'). \quad (27)$$

It follows from Lemma 23 that $(A_{\delta,\ell})' = A_{1-\delta,\ell}$. Thus the set $(A_{\delta,\ell} \cap (P \setminus E'))'$ is equal to $A_{1-\delta,\ell} \cap (P \setminus E)$, and therefore $\#A_{\delta,\ell} \cap (P \setminus E') = \#A_{1-\delta,\ell} \cap (P \setminus E)$. Substituting this into equation (27), we see that it is equivalent to

$$\#A_{\delta,\ell} \cap (P \setminus E) = \#A_{1-\delta,\ell} \cap (P \setminus E). \quad (28)$$

We establish (28) by exhibiting an involution of the set $F := P \setminus E$ that leaves ℓ fixed and interchanges δ with $1-\delta$. Observe that F is the set of permutations σ that are almost proper but $x^\sigma(k)$ is non-proper. Thus for $\sigma \in F$ the word $x^\sigma(k)$ is non-proper but becomes proper after deleting a single character. This means that there exists a unique pair of integers $r = r^\sigma$ and $s = s^\sigma$ such that $r < s$ and $x_r = x_s$ and r, s occur as consecutive indices in the sequence $(m : \sigma(m) \leq k)$ that indexes $x^\sigma(k)$. Observe that either $\sigma(r) = k$ or $\sigma(s) = k$. Let $a = \sigma^{-1}(k+1)$ denote the integer arriving at time $k+1$. Since $x^\sigma(k+1)$ is proper, x_a must be inserted between x_r and x_s . (This implies, in particular, that $\Delta_k(\sigma) \in \{0, 1\}$.)

Define the function f on F via

$$f(\sigma) = \sigma \circ (r^\sigma \ s^\sigma), \quad \sigma \in F.$$

That is, f interchanges the arrival times of r and s in σ . Since $x_{r^\sigma} = x_{s^\sigma}$ for $\sigma \in F$, it follows that $x^\sigma(t) = x^{f(\sigma)}(t)$ for all $t \in I$. Thus $f(\sigma) \in F$ as well. Furthermore $r^{f(\sigma)} = r^\sigma$ and similarly for s , from which it follows that f is involutive. All that remains is to verify that f preserves ℓ and interchanges δ with $1-\delta$, that is,

$$\tilde{\mathcal{L}}(f(\sigma))_i = \tilde{\mathcal{L}}(\sigma)_i, \quad i \in I \setminus \{k, k+1\}, \sigma \in F, \quad (29)$$

and

$$\Delta_k(f(\sigma)) = 1 - \Delta_k(\sigma), \quad \sigma \in F. \quad (30)$$

Suppose, for the sake of concreteness, that $\sigma(j) = k$; the other case $\sigma(i) = k$ is similar. The subwords of x built by σ and $f(\sigma)$ at times $k-1$, k , and $k+1$ are then

$$\begin{aligned} x^\sigma(k-1) &= u x_i v, & x^\sigma(k) &= u x_i \overset{\circ}{x_j} \boxed{v}, & x^\sigma(k+1) &= u x_i \overset{\circ}{x_a} \boxed{x_j v}, \\ x^{f(\sigma)}(k-1) &= u x_j v, & x^{f(\sigma)}(k) &= u \overset{\circ}{x_i} \boxed{x_j v}, & x^{f(\sigma)}(k+1) &= u x_i \overset{\circ}{x_a} \boxed{x_j v}, \end{aligned}$$

where we have circled the character that was just inserted, boxed the subword to its right, and set $u = (x_m : \sigma(m) < k, k < i)$ and $v = (x_m : \sigma(m) < k, m > j)$. Thus when $\sigma(j) = k$, we have that

$$\Delta_k(\sigma) = |x_j v| - |v| = 1 \text{ and } \Delta_k(f(\sigma)) = |x_j v| - |x_j v| = 0.$$

Similarly when $\sigma(i) = k$, we have that

$$\Delta_k(\sigma) = |x_j v| - |x_j v| = 0 \text{ and } \Delta_k(f(\sigma)) = |x_j v| - |v| = 1.$$

This establishes (30). Moreover (29) clearly holds when $i > k+1$, and for $i < k$ it follows since x_r and x_s occupy the same relative positions in the respective subwords $x^\sigma(i)$ and $x^{f(\sigma)}(i)$. \square

Fix $n \geq 0$. For a word $x = (x_i)_{i=0}^n$, we write $\bar{x} = (x_{n-i})_{i=0}^n$ for its reversal. For a permutation $\sigma \in \text{Sym}(\llbracket 0, n \rrbracket)$, its reversal is the permutation $\bar{\sigma}$ with $\bar{\sigma}(i) = \sigma(n-i)$.

Lemma 24. *Let σ be a Mallows-distributed permutation of $\llbracket 0, n \rrbracket$. Then*

$$\mathbb{P}(\sigma \vdash x) = \mathbb{P}(\bar{\sigma} \vdash x), \quad \forall x \in \llbracket 1, q \rrbracket^{\llbracket 0, n \rrbracket}.$$

Proof. Write \mathcal{G}_i and $\bar{\mathcal{G}}_i$ for the respective laws of X_i and $i - X_i$, where $0 \leq i \leq n$ and X_i is an i -truncated, t -geometric random variable. By Lemma 12, the random sequences $\tilde{\mathcal{L}}(\sigma)$ and $\tilde{\mathcal{L}}(\bar{\sigma})$ have laws $\mathcal{G}_0 \otimes \mathcal{G}_1 \otimes \cdots \otimes \mathcal{G}_n$ and $\bar{\mathcal{G}}_0 \otimes \cdots \otimes \bar{\mathcal{G}}_n$, respectively. Fix a word x and let B_x denote the set of tuples $\ell \in \{0\} \times \cdots \times \llbracket 0, n \rrbracket$ such that $\tilde{\mathcal{D}}(\ell)$ is a proper building of x . The desired result is equivalent to

$$\mathcal{G}_0 \otimes \cdots \otimes \mathcal{G}_n(B_x) = \bar{\mathcal{G}}_0 \otimes \cdots \otimes \bar{\mathcal{G}}_n(B_x). \quad (31)$$

Note that $\mathcal{G}_0 = \overline{\mathcal{G}_0}$. We will establish (31) by showing that

$$\begin{aligned}
\mathcal{G}_0 \otimes \cdots \otimes \mathcal{G}_n(B_x) &= \overline{\mathcal{G}_0} \otimes \mathcal{G}_1 \otimes \cdots \otimes \mathcal{G}_{n-1} \otimes \mathcal{G}_n(B_x) \\
&= \mathcal{G}_0 \otimes \overline{\mathcal{G}_1} \otimes \cdots \otimes \mathcal{G}_{n-1} \otimes \mathcal{G}_n(B_x) \\
\cdots &= \mathcal{G}_0 \otimes \mathcal{G}_1 \otimes \cdots \otimes \mathcal{G}_{n-1} \otimes \overline{\mathcal{G}_n}(B_x) \\
\cdots &= \mathcal{G}_0 \otimes \mathcal{G}_1 \otimes \cdots \otimes \overline{\mathcal{G}_{n-1}} \otimes \overline{\mathcal{G}_n}(B_x) \\
&\quad \dots\dots\dots \\
\cdots &= \overline{\mathcal{G}_0} \otimes \overline{\mathcal{G}_1} \otimes \cdots \otimes \overline{\mathcal{G}_{n-1}} \otimes \overline{\mathcal{G}_n}(B_x).
\end{aligned}$$

More precisely, we will show that for all $0 \leq k < m \leq n$ fixed,

$$\mu(B_x) = \mu'(B_x), \quad (32)$$

where

$$\begin{aligned}
\mu &= \mathcal{G}_0 \otimes \cdots \otimes \mathcal{G}_{k-1} \otimes \overline{\mathcal{G}_k} \otimes \mathcal{G}_{k+1} \otimes \mathcal{G}_{k+2} \otimes \cdots \otimes \mathcal{G}_{m-1} \otimes \overline{\mathcal{G}_m} \otimes \cdots \otimes \overline{\mathcal{G}_n} \text{ and} \\
\mu' &= \mathcal{G}_0 \otimes \cdots \otimes \mathcal{G}_{k-1} \otimes \mathcal{G}_k \otimes \overline{\mathcal{G}_{k+1}} \otimes \mathcal{G}_{k+2} \otimes \cdots \otimes \mathcal{G}_{m-1} \otimes \overline{\mathcal{G}_m} \otimes \cdots \otimes \overline{\mathcal{G}_n}.
\end{aligned}$$

Any two independent truncated geometric random variables are conditionally uniform on the set of possible values given their difference. From this it follows that a random tuple L with law μ is conditionally uniform on some set given $L_{k+1} - L_k$ and $(L_i : i \in \llbracket 0, n \rrbracket \setminus \{k, k+1\})$. Thus by Lemma 22, the following conditional probabilities are equal:

$$\mu(B_x \mid L_{k+1} - L_k, (L_i : i \neq k, k+1)) = \mu(B'_x \mid L_{k+1} - L_k, (L_i : i \neq k, k+1)),$$

where

$$B'_x := \left\{ \ell \in \{0\} \times \cdots \times \llbracket 0, n \rrbracket : (k \ k+1) \circ \tilde{\mathcal{D}}(\ell) \vdash x \right\}.$$

Integrating out the conditioning yields that

$$\mu(B_x) = \mu(B'_x). \quad (33)$$

If (L_k, L_{k+1}) has law $\overline{\mathcal{G}_k} \otimes \mathcal{G}_{k+1}$, then

$$(L_{k+1} - \mathbf{1}[L_{k+1} > L_k], L_k + \mathbf{1}[L_{k+1} \leq L_k])$$

has law $\mathcal{G}_k \otimes \overline{\mathcal{G}_{k+1}}$. When combined with Lemma 23, this yields that $\mu(B'_x) = \mu'(B_x)$. We now deduce (32) from this and (33). The lemma follows by repeated application of (32) as indicated above. \square

Corollary 25. *If $X = (X_i)_{i \in \mathbb{Z}}$ is a random q -coloring of \mathbb{Z} with law $\text{MalCol}_{q,t}$, then*

$$(X_0, X_1, \dots, X_n) \stackrel{d}{=} (X_n, X_{n-1}, \dots, X_0), \quad n \geq 0.$$

Proof. This follows directly from the previous lemma, since for any word x and for σ a random Mallows permutation, equation (9) implies that the quantities $\mathbb{P}(\sigma \vdash x)$ and $\mathbb{P}((X_0, \dots, X_n) = x)$ are constant multiples of one another. \square

Proof of Proposition 20. This is immediate from Corollary 25. \square

Proof of Corollary 21. By inspection, $(q, t, 1)$ satisfies (7) if and only if $q = (t+1)^2/t$, which is equivalent to

$$t^{1/2} + t^{-1/2} = \sqrt{q}. \quad (34)$$

Provided $\sqrt{q} \geq 2$, there exists $t = t(q) \in [0, 1]$ satisfying this equation. Then for any proper word x ,

$$B_t(x) = \sum_{i=0}^n t^{n-i} B_t(\hat{x}_i) \quad \text{and} \quad B_t(x) = \sum_{i=0}^n t^i B_t(\hat{x}_i),$$

where the first equation is due to Lemma 15 and the second follows by combining Lemma 15 with Corollary 25. Averaging these two equalities yields

$$B_t(x) = \sum_{i=0}^n \left(\frac{t^{n-i} + t^i}{2} \right) B_t(\hat{x}_i).$$

By (34), the coefficients of this recurrence agree with those of [31, Eq. 2.2]. \square

5 Convergence of Lehmer codes

Previously we showed that $\text{MalCol}_{q,t}$ is k -dependent whenever (q, k, t) satisfy the tuning equation (7) and that $\text{MalCol}_{q,t}$ is reversible. The key step remaining in the proof of the main theorem is to exhibit a finitary factor of an iid process having this law. We accomplish this over the course of the next two sections. The idea will be to give a probabilistic construction of the colorings on \mathbb{Z} . On a finite interval, this was done already in Proposition 13. Our goal now is to extend this construction to \mathbb{Z} by taking appropriate limits.

In this section we show that the law of the Lehmer code of a BMal -distributed permutation of $\llbracket -n, n \rrbracket$ is iid in the limit as $n \rightarrow \infty$ with a certain distribution. In the next section we will use this iid sequence to produce a random coloring of \mathbb{Z} satisfying all of the properties claimed in Theorem 1.

Both the Lehmer code, \mathcal{L} , and the insertion code, $\tilde{\mathcal{L}}$, play an important role in these sections. The zeros of the Lehmer code occur at key locations in the coloring (essentially,

they are the renewal times). On the other hand, the pushforward of BMal under the insertion code is the product of truncated geometric distributions. The pushforward under the Lehmer code is not a product measure for general finite intervals, but our main result in this section is that it tends towards a product measure as the interval approaches \mathbb{Z} .

Recall that the Lehmer code is the map $\mathcal{L}: \text{Sym}(I) \rightarrow \llbracket 0, \infty \rrbracket^I$ given by

$$\mathcal{L}(\sigma) = \left(\#\{j \in I: j > i \text{ and } \sigma(j) < \sigma(i)\} \right)_{i \in I} \quad \text{for } \sigma \in \text{Sym}(I).$$

Here we identify $\llbracket 0, \infty \rrbracket^I$ with the subset of $\llbracket 0, \infty \rrbracket^{\mathbb{Z}}$ consisting of sequences vanishing on $\mathbb{Z} \setminus I$, so that we can compare permutations on different intervals.

Recall from Section 2 that $\text{BMal}_{t,u}$, the bubble-biased Mallows measure on permutations, assigns to $\sigma \in \text{Sym}(I)$ a probability proportional to $u^{\#\text{bub}(\Gamma_\sigma)} t^{\text{inv}(\sigma)}$.

Proposition 26. *Let σ_n be a random permutation of $\llbracket -n, n \rrbracket$ with law $\text{BMal}_{t,u}$. As $n \rightarrow \infty$ the sequence $\mathcal{L}(\sigma_n)$ converges in law, with respect to the product topology on $\llbracket 0, \infty \rrbracket^{\mathbb{Z}}$, to an iid sequence of u -zero-weighted, t -geometric random variables.*

The remainder of this section is devoted to the proof of this proposition, which requires several simple but technical lemmas. Recall that $\mathcal{F}(\sigma)$ denotes the set of founders of the permutation σ (see Section 2).

Lemma 27. *Let σ be a permutation of $\llbracket 0, n \rrbracket$. Then $\sigma^{-1}(0)$ is a founder of σ and*

$$\sigma^{-1}(0) = \min\{0 \leq i \leq n: \mathcal{L}(\sigma)_i = 0\} \text{ and} \quad (35)$$

$$\mathcal{F}(\sigma) \cap \llbracket \sigma^{-1}(0), n \rrbracket = \{i \in I: \mathcal{L}(\sigma)_i = 0\} \text{ and} \quad (36)$$

$$\mathcal{F}(\sigma) \cap \llbracket 0, \sigma^{-1}(0) \rrbracket = \{i \in I: \mathcal{L}(\sigma)_i = \sigma(i)\}. \quad (37)$$

Proof. By the definition of a founder, $i \in \mathcal{F}(\sigma)$ if and only if either $\sigma(k) > \sigma(i)$ for all $k > i$ or $\sigma(j) > \sigma(i)$ for all $j < i$. When $i = \sigma^{-1}(0)$ both conditions hold. If $i > \sigma^{-1}(0)$, then the latter condition cannot hold, and so

$$\mathcal{F}(\sigma) \cap \llbracket \sigma^{-1}(0), n \rrbracket = \{i \in I: \sigma(k) > \sigma(i) \text{ for all } k > i\} = \{i \in I: \mathcal{L}(\sigma)_i = 0\}.$$

Similarly if $i < \sigma^{-1}(0)$, then the former condition cannot hold, and so

$$\mathcal{F}(\sigma) \cap \llbracket 0, \sigma^{-1}(0) \rrbracket = \{i \in I: \sigma(j) > \sigma(i) \text{ for all } j < i\} = \{i: \mathcal{L}(\sigma)_i = \sigma(i)\},$$

where the last equality follows since $\sigma(i) - \mathcal{L}(\sigma)_i = \#\{j < i: \sigma(j) < \sigma(i)\}$. \square

Lemma 28. *Let σ be a random permutation of $\llbracket m, n \rrbracket$ with law $\text{BMal}_{t,u}$ and let $i \in \llbracket m, n \rrbracket$. Given $\sigma^{-1}(m) < i$, the random variables $(\mathcal{L}(\sigma)_j)_{j=i}^n$ are conditionally independent of each other, with the conditional law of $\mathcal{L}(\sigma)_j$ being the u -zero-weighted, $(n-j)$ -truncated, t -geometric distribution.*

Proof. By a simple relabelling we assume that $m = 0$ without loss of generality. For each $\ell = (\ell_i, \dots, \ell_n) \in \llbracket 0, n-i \rrbracket \times \dots \times \{0\}$, let

$$\mathcal{A}(\ell) = \{\tau \in \text{Sym}(\llbracket 0, n \rrbracket) : \tau^{-1}(0) < i \text{ and } \mathcal{L}(\tau)_j = \ell_j \text{ for all } i \leq j \leq n\}.$$

Recall that $\#\mathcal{F}(\tau) = \#\text{bub}(\Gamma_\tau) + 1$. Thus

$$\mathbb{P}((\mathcal{L}(\sigma)_i, \dots, \mathcal{L}(\sigma)_n) = \ell \mid \sigma^{-1}(0) < i) = \frac{\mathbb{P}(\sigma \in \mathcal{A}(\ell))}{\mathbb{P}(\sigma^{-1}(0) < i)} = \frac{1}{Z} \sum_{\tau \in \mathcal{A}(\ell)} t^{\text{inv}(\tau)} u^{\#\mathcal{F}(\tau)},$$

where $Z = \mathbb{P}(\sigma^{-1}(0) < i)$.

Observe that

$$\begin{aligned} \sum_{\tau \in \mathcal{A}(\ell)} t^{\text{inv}(\tau)} u^{\#\mathcal{F}(\tau)} &= \sum_{\tau \in \mathcal{A}(\ell)} \prod_{j=0}^n t^{\mathcal{L}(\tau)_j} u^{\mathbf{1}[j \in \mathcal{F}(\tau)]} \\ &= \prod_{j=i}^n t^{\ell_j} u^{\mathbf{1}[\ell_j=0]} \left[\sum_{\tau \in \mathcal{A}(\ell)} \prod_{j=0}^{i-1} t^{\mathcal{L}(\tau)_j} u^{\mathbf{1}[j \in \mathcal{F}(\tau)]} \right], \end{aligned} \quad (38)$$

where in the second equality we have applied equation (36) of Lemma 27. Let

$$R_{u,n,i,t}(\ell) := \sum_{\tau \in \mathcal{A}(\ell)} \prod_{j=0}^{i-1} t^{\mathcal{L}(\tau)_j} u^{\mathbf{1}[j \in \mathcal{F}(\tau)]}$$

denote the bracketed expression in (38). The product appearing in $R_{u,n,i,t}$ factorizes into two factors, corresponding to equations (36) and (37) of Lemma 27 respectively:

$$R_{u,n,i,t}(\ell) = \sum_{\tau \in \mathcal{A}(\ell)} \left[\prod_{j=0}^{\tau^{-1}(0)-1} t^{\mathcal{L}(\tau)_j} u^{\mathbf{1}[\mathcal{L}(\tau)_j = \tau(j)]} \right] \left[\prod_{j=\tau^{-1}(0)}^{i-1} t^{\mathcal{L}(\tau)_j} u^{\mathbf{1}[\mathcal{L}(\tau)_j = 0]} \right]. \quad (39)$$

We wish to show that the function $R_{u,n,i,t}$ is constant, i.e., that for any pair of tuples ℓ and ℓ' belonging to $\llbracket 0, n-i \rrbracket \times \dots \times \{0\}$, it holds that $R_{u,n,i,t}(\ell) = R_{u,n,i,t}(\ell')$.

Recall that \mathcal{L} is a bijection from $\text{Sym}(\llbracket 0, n \rrbracket)$ to $\llbracket 0, n \rrbracket \times \dots \times \{0\}$ with inverse \mathcal{D} . For any $\tau \in \mathcal{A}(\ell)$, eq. (36) implies that $\tau^{-1}(0) < i$ if and only if there exists $j \in \llbracket 0, i \rrbracket$ such that $\mathcal{L}(\tau)_j = 0$. Thus $\mathcal{L}(\mathcal{A}(\ell))$ is the set of tuples in $\llbracket 0, n \rrbracket \times \dots \times \{0\}$ whose restriction to $\llbracket i, n \rrbracket$ is ℓ and whose restriction to $\llbracket 0, i \rrbracket$ has at least one entry that vanishes. The set

$\mathcal{L}(\mathcal{A}(\ell'))$ bears a similar description. Let P be the bijection from $\mathcal{L}(\mathcal{A}(\ell))$ to $\mathcal{L}(\mathcal{A}(\ell'))$ leaving the restriction to $\llbracket 0, i \rrbracket$ fixed, and let $Q = \mathcal{D} \circ P \circ \mathcal{L}$ be the corresponding bijection from $\mathcal{A}(\ell)$ to $\mathcal{A}(\ell')$.

By the explicit formula for \mathcal{D} in (3), it follows that

$$\tau(j) = \left(\bigcirc_{k=0}^j \pi_{\llbracket k, k + \mathcal{L}(\tau)_k \rrbracket}^- \right)(j), \quad j \in \llbracket 0, i \rrbracket.$$

Since P fixes the restriction of $\mathcal{L}(\tau)$ to $\llbracket 0, i \rrbracket$, we have that Q fixes the restriction of τ to $\llbracket 0, i \rrbracket$. Applying (36) once more shows that Q also fixes $\tau^{-1}(0)$. Thus, each summand of (39) is unchanged by the action of Q , from which it follows that $R_{u,n,i,t}(\ell) = R_{u,n,i,t}(\ell')$. Since ℓ and ℓ' were arbitrary, $R_{u,n,i,t}$ does not depend on ℓ . The lemma now follows from (38). \square

A real-valued random variable X is said to **stochastically dominate** another random variable Y if $\mathbb{P}(X > r) \geq \mathbb{P}(Y > r)$ for all r .

Lemma 29. *Fix $0 < t < s < 1$ and $u \geq 1$. Let S be an n -truncated, s -geometric random variable and let T be a u -end-weighted, n -truncated, t -geometric random variable. Then S stochastically dominates T for all*

$$n \geq n_0 := \log_{s/t} \left(u \cdot \frac{1-t}{1-s} \right).$$

For terminology regarding variants of geometric random variables, refer to equations (i) to (v) in Section 2.

Proof. Let M be a u -max-weighted n -truncated t -geometric variable. By inspection of the mass functions of M and T , it is apparent that M stochastically dominates T for all $u \geq 1$. Thus it remains to show that S stochastically dominates M .

For this, we argue that $\mathbb{P}(M < k) \geq \mathbb{P}(S < k)$ for all $k \in \llbracket 1, n \rrbracket$, which by the formulas in (i) and (iii) of Section 2 is equivalent to

$$\frac{\frac{1-t^k}{1-t}}{\frac{1-t^n}{1-t} + ut^n} \geq \frac{1-s^k}{1-s^{n+1}},$$

which is, in turn, equivalent to

$$\frac{1-t^k}{1-s^k} \geq \frac{1-t^n + ut^n(1-t)}{1-s^{n+1}}.$$

Since $t < s$, the left side of the latter inequality is a decreasing function of k , and thus it suffices to prove the inequality for $k = n$. In this case the inequality rearranges to

$$\left(\frac{s}{t} \right)^n \geq u \cdot \frac{1-t}{1-s} \cdot \frac{1-s^n}{1-t^n}.$$

By our choice of n_0 , the inequality holds for all $n \geq n_0$, as desired. \square

We use the previous lemma to prove the following tightness result for the bubble-biased Mallows measure.

Lemma 30. *Fix $u \in [1, \infty)$ and $t \in [0, 1)$. For each $n \geq 0$ let σ_n be a random permutation $\llbracket 0, n \rrbracket$ with law $\text{BMal}_{t,u}$. Then $(\sigma_n^{-1}(0))_{n \geq 0}$ is tight.*

Proof. We prove the result by finding a coupling of the permutations in which

$$\mathbb{P}(\sup_n \sigma_n^{-1}(0) = \infty) = 0,$$

from which tightness follows.

Let $(X_n^\sigma)_{n \geq 0}$ be a independent random variables with X_n^σ being u -end-weighted, n -truncated, and t -geometric. By Lemma 12, for all n the law of the random permutation $\sigma_n = \tilde{\mathcal{D}}(X_0^\sigma, \dots, X_n^\sigma)$ is $\text{BMal}_{t,u}^{\llbracket 0, n \rrbracket}$. By inspection of the formula (4) for $\tilde{\mathcal{D}}$, it follows that the sequence $\{\sigma_n^{-1}(0)\}_n$ is a.s. non-decreasing, and therefore

$$\mathbb{P}(\sup_n \sigma_n^{-1}(0) \geq k) = \lim_{n \rightarrow \infty} \mathbb{P}(\sigma_n^{-1}(0) \geq k), \quad \forall k \geq 0. \quad (40)$$

When $u = 1$, the law of σ_n is Mal_t , and the same holds for σ_n^{-1} by inversion symmetry of the Mallows measure. Since $\sigma_n^{-1}(0) = \mathcal{L}(\sigma_n^{-1})_0$, it follows by Lemma 8 that $\sigma_n^{-1}(0)$ is an n -truncated, t -geometric random variable. Combining this with (40) implies that $\sup_n \sigma_n^{-1}(0)$ is a.s. finite when $u = 1$.

We extend this result from the case $u = 1$ to $u > 1$ using a domination argument. Fix s such that $t < s < 1$ and take n_0 to be any integer larger than the constant $n_0(s, t, u)$ from Lemma 29, thereby guaranteeing that an s -geometric n -truncated random variable stochastically dominates a u -end-weighted, n -truncated, t -geometric random variable for all $n \geq n_0$. Let $(X_n^\tau)_{n \geq 0}$ be independent random variables with X_n^τ being n -truncated and s -geometric. We couple $(X_n^\sigma)_{n > n_0}$ and $(X_n^\tau)_{n > n_0}$ such that $X_n^\sigma \leq X_n^\tau$ a.s. for all $n > n_0$ using Strassen's Theorem [58]. By applying Lemma 12 with $(t, u) = (s, 1)$, it follows that for all $n \geq 0$, the random permutation $\tau_n = \tilde{\mathcal{D}}(X_0^\tau, \dots, X_n^\tau)$ has law Mal_s . Hence $\sup_n \tau_n^{-1}(0)$ is a.s. finite by the previous analysis of the case $u = 1$.

For all $i \in \llbracket 0, \infty \rrbracket$ and all $a_i \in \llbracket 0, i \rrbracket$, the probabilities $\mathbb{P}(X_i^\sigma = a_i)$ and $\mathbb{P}(X_i^\tau = a_i)$ are positive. We claim that, for any sequence of integers a_0, \dots, a_{n_0} satisfying these conditions, we have that

$$\begin{aligned} & \mathbb{P}\left(\sup_n \sigma_n^{-1}(0) = \infty \mid X_i^\sigma = a_i, \forall 0 \leq i \leq n_0\right) \\ & \leq \mathbb{P}\left(\sup_n \tau_n^{-1}(0) = \infty \mid X_i^\tau = a_i, \forall 0 \leq i \leq n_0\right). \end{aligned} \quad (41)$$

Indeed, this is equivalent to

$$\begin{aligned} \mathbb{P}\left(\sup_n \tilde{\mathcal{D}}(a_0, \dots, a_{n_0}, X_{n_0+1}^\sigma, \dots, X_n^\sigma)^{-1}(0) = \infty\right) \\ \leq \mathbb{P}\left(\sup_n \tilde{\mathcal{D}}(a_0, \dots, a_{n_0}, X_{n_0+1}^\tau, \dots, X_n^\tau)^{-1}(0) = \infty\right), \end{aligned}$$

which follows since the function $\tilde{\mathcal{D}}(\ell_0, \dots, \ell_n)^{-1}(0)$ is non-decreasing in each of its arguments, as seen by inspection of the formula (4) for $\tilde{\mathcal{D}}$.

But we have already shown that $\mathbb{P}(\sup_n \tau_n^{-1}(0) = \infty) = 0$, whereupon the right side of (41) is zero. Thus the left side is zero as well. This implies that $\mathbb{P}(\sup_n \sigma_n^{-1}(0) = \infty) = 0$, from which the lemma now follows. \square

Proof of Proposition 26. Let $L = (L_i)_{i \in \mathbb{Z}}$ be a sequence of iid u -zero-weighted, t -geometric random variables. For each n , let σ_n be a random permutation of $\llbracket -n, n \rrbracket$ with law $\text{BMal}_{t,u}$. The statement of the proposition is equivalent to the equality

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\left(\mathcal{L}(\sigma_n)_i\right)_{i \in I} = \ell\right) = \mathbb{P}\left(\left(L_i\right)_{i \in I} = \ell\right), \quad (42)$$

for all finite intervals $I \subset \mathbb{Z}$ and for all $\ell \in \llbracket 0, \infty \rrbracket^I$.

By considering the interval $\llbracket 0, 2n \rrbracket$ and shifting, Lemma 30 implies that

$$\lim_{n \rightarrow \infty} \mathbb{P}(\sigma_n^{-1}(-n) < \min I) = 1$$

Combining this with Lemma 28 establishes (42), proving the proposition. \square

6 Proof of main theorem

Fix $t \in (0, 1)$ and $q \geq 3$. In Section 3 we constructed a measure $\text{MalCol}_{q,t}$ which is the law of a k -dependent q -coloring of \mathbb{Z} whenever (q, k, t) satisfies the tuning equation (7). Here we give a construction, directly on the integers, of a random coloring with this law. This random coloring will arise as a uniform proper q -coloring of a certain random infinite graph. Thus, we begin by explaining what we mean by a ‘uniform proper q -coloring’ of an infinite graph.

First suppose that the graph in question is the nearest-neighbor graph on the integers. We define a uniform proper q -coloring of this graph to be the bi-infinite trajectory of a stationary simple random walk on the complete graph (without self-loops) with vertex set $\llbracket 1, q \rrbracket$.

For a graph G with vertex set \mathbb{Z} , recall that the bubble endpoints of G were defined in Section 2 to be those integers i such that there do not exist integers j and k adjacent in

G with $j < i < k$. Say that G is **good** if it is q -colorable, its set of bubble endpoints is unbounded from above and below, and consecutive bubble endpoints are adjacent in G . For any good graph G , we define a **uniform proper q -coloring** of G to be a random coloring equal in law to the output of the following algorithm.

Algorithm 2 (Uniform coloring algorithm). *Input a good graph G with vertex set \mathbb{Z} .*

- (i) *Let $(b_e)_{e \in \mathbb{Z}}$ be an increasing enumeration of the bubble endpoints of G .*
- (ii) *Let $(Y_e)_{e \in \mathbb{Z}}$ be a uniform proper q -coloring of \mathbb{Z} , and set $X_{b_e} = Y_e$ for all $e \in \mathbb{Z}$.*
- (iii) *Conditional on step (ii), for each bubble of G , choose independently a proper q -coloring of the bubble uniformly from among those that assign the colors from step (ii) to the bubble endpoints. Output the resulting coloring $(X_i)_{i \in \mathbb{Z}}$.*

Lemma 31. *Let G be a good graph. Let $(X_i)_{i \in \mathbb{Z}}$ be a uniform proper q -coloring of G . Then for every finite interval $I \subset \mathbb{Z}$ whose endpoints are bubble endpoints of G , the coloring $(X_i: i \in I)$ is distributed uniformly on the set of proper q -colorings of the subgraph of G induced by I .*

Proof. Write G_I for the subgraph of G induced by I . Let (b_0, b_1, \dots, b_n) be an increasing enumeration of the bubble endpoints of G contained in I , and let bub_i be the subgraph of G induced by $\llbracket b_{i-1}, b_i \rrbracket$ for $1 \leq i \leq n$. Then $G_I = \text{bub}_1 \cup \dots \cup \text{bub}_n$, and there is a bijection between proper q -colorings of G_I and proper q -colorings of $\text{bub}_1, \dots, \text{bub}_n$ that agree at their endpoints.

Thus if $U = (U_i: i \in I)$ is a uniform proper q -coloring of G_I , then the conditional law of U given $U_{b_0}, U_{b_1}, \dots, U_{b_n}$ coincides with that of $(X_i: i \in I)$ given $X_{b_0}, X_{b_1}, \dots, X_{b_n}$, by step (iii) of the algorithm. Since the subgraph of G induced by b_0, b_1, \dots, b_n is a path, the laws of $(X_{b_0}, X_{b_1}, \dots, X_{b_n})$ and $(U_{b_0}, U_{b_1}, \dots, U_{b_n})$ are equal by step (ii) of the algorithm. Thus the unconditional laws of U and $(X_i: i \in I)$ coincide. \square

Recall the definition of the constraint graph Γ_σ of a permutation σ of \mathbb{Z} . In the case when σ is a finite permutation of \mathbb{Z} , Γ_σ can be expressed in terms of the Lehmer code $\mathcal{L}(\sigma)$, a sequence in which all but finitely many entries vanish. We now extrapolate to a graph defined in terms of a more general sequence. Recall the map \mathcal{D} , inverse to \mathcal{L} , defined in Section 2.

Definition (of $\Gamma[\ell]$). *Let $\ell \in \llbracket 0, \infty \rrbracket^{\mathbb{Z}}$ be a sequence for which the zero set $(i \in \mathbb{Z}: \ell_i = 0)$ is unbounded from above and below. Let $(i_k)_{k \in \mathbb{Z}}$ be an increasing enumeration of $\mathcal{Z}(\ell)$, and*

for each $k \in \mathbb{Z}$ let $A_k = (\ell_i : i \in \llbracket i_k, i_{k+1} \rrbracket)$. Then the graph $\Gamma[\ell]$ is defined to be

$$\Gamma[\ell] := \bigcup_{k \in \mathbb{Z}} \Gamma_{\mathcal{D}(A_k)}.$$

This generalizes the definition of the constraint graph in that, if σ is a finite permutation of \mathbb{Z} , then $\Gamma[\mathcal{L}(\sigma)] = \Gamma_\sigma$. In fact, this follows from the next lemma.

Lemma 32. *For any integers $a \leq i < j \leq b$ and for any sequence $\ell \in \llbracket 0, \infty \rrbracket^{\mathbb{Z}}$ with zero set unbounded from above and below, the integers i and j are adjacent in $\Gamma[\ell]$ if and only if they are adjacent in the constraint graph of $\mathcal{D}(\ell_a, \ell_{a+1}, \dots, \ell_b)$.*

Proof. First suppose that there exists $k \in (i, j)$ such that $\ell_k = 0$. Then i and j are non-adjacent in $\Gamma[\ell]$ by definition. Likewise, they are non-adjacent in the constraint graph of the permutation $\sigma = \mathcal{D}(a, a+1, \dots, b)$ since k is a founder of σ by Lemma 27.

Otherwise, there exist $a' \leq i < j \leq b'$ such that a' and b' are consecutive zeros of ℓ . Thus by definition of $\Gamma[\ell]$, the integers i and j are adjacent in $\Gamma[\ell]$ iff they are adjacent in the constraint graph of $\sigma' = \mathcal{D}(a', a'+1, \dots, b')$. Thus, the lemma will follow once we show that i and j are adjacent in Γ_σ iff they are adjacent in $\Gamma_{\sigma'}$. Recall the explicit formula (3) expressing the function \mathcal{D} as a composition of cycles, from the discussion preceding Lemma 7 in Section 2. From this formula, it follows that

$$\sigma = \bigcirc_{k=a}^{i-1} \pi_{\llbracket k, k+\ell_k \rrbracket}^- \circ \bigcirc_{k=i}^j \pi_{\llbracket k, k+\ell_k \rrbracket}^- \circ \bigcirc_{k=j+1}^b \pi_{\llbracket k, k+\ell_k \rrbracket}^-$$

and

$$\sigma' = \bigcirc_{k=a'}^{i-1} \pi_{\llbracket k, k+\ell_k \rrbracket}^- \circ \bigcirc_{k=i}^j \pi_{\llbracket k, k+\ell_k \rrbracket}^- \circ \bigcirc_{k=j+1}^{b'} \pi_{\llbracket k, k+\ell_k \rrbracket}^-.$$

The cycle $\pi_{\llbracket k, k+\ell_k \rrbracket}^-$ leaves the relative ordering of $\llbracket i, j \rrbracket$ unchanged whenever $k \notin \llbracket i, j \rrbracket$. Thus it is only the shared middle factor which determines the relative ordering of σ and σ' on $\llbracket i, j \rrbracket$. This, in turn, determines whether $\sigma(i) < \sigma(k) > \sigma(j)$ for all $i < k < j$, which is equivalent to i and j being adjacent in σ , and respectively for σ' . \square

Corollary 33. *For $\ell \in \llbracket 0, \infty \rrbracket^{\mathbb{Z}}$ with zero set unbounded from above and below,*

- (i) *the set of bubble endpoints of $\Gamma[\ell]$ is equal to the zero set of ℓ , and*
- (ii) *$\Gamma[\ell]$ is good.*

Proof. By definition of $\Gamma[\ell]$, all zeros of ℓ are bubble endpoints. Conversely, suppose that $\ell_i > 0$. Let $j < i$ be maximal such that $\ell_j = 0$ and let $k > i$ be minimal such that $\ell_k = 0$. Then j is adjacent to k in the constraint graph of $\mathcal{D}(\ell_j, \ell_{j+1}, \dots, \ell_k)$, and therefore

by Lemma 32 j and k are also adjacent in $\Gamma[\ell]$. Thus i is not a bubble endpoint, proving part (i).

By part (i) the bubble endpoints of $\Gamma[\ell]$ are unbounded from above and below. By Lemma 9, $\Gamma[\ell]$ decomposes into a collection of finite bubbles joined at their endpoints. Now Lemma 11 implies that $\Gamma[\ell]$ is q -colorable. That consecutive bubble endpoints of $\Gamma[\ell]$ are adjacent follows from the corresponding property for a single bubble of the constraint graph of a permutation. Thus $\Gamma[\ell]$ is good. \square

We can now prove the following key result.

Proposition 34. *Set $u = \frac{q-1}{q-2}$ and let $L = (L_i)_{i \in \mathbb{Z}}$ be an iid sequence of u -zero-weighted, t -geometric random variables. Conditional on L , choose a uniform proper q -coloring of $\Gamma[L]$. Then the (unconditional) law of the resulting coloring of \mathbb{Z} is $\text{MalCol}_{q,t}$.*

We fix $u = \frac{q-1}{q-2}$ for the remainder of the section.

Proof of Proposition 34. By Corollary 33(ii), $\Gamma[L]$ is a.s. good. Let $X = (X_i)_{i \in \mathbb{Z}}$ be a uniform proper q -coloring of $\Gamma[L]$ and let $Y = (Y_i)_{i \in \mathbb{Z}}$ be a random coloring with law $\text{MalCol}_{q,t}$. Let σ_n be a random permutation of $\llbracket -n, n \rrbracket$ with law $\text{BMal}_{t,u}$ and let $Y^n = (Y_i^n : i \in \llbracket -n, n \rrbracket)$ be a uniform proper coloring of Γ_{σ_n} . By Proposition 13, the sequence Y^n is equal in law to $(Y_i : i \in \llbracket -n, n \rrbracket)$.

In Proposition 26 we showed that $\mathcal{L}(\sigma_n)$ converges in distribution to L . Thus by the Skorohod representation theorem [38], there exists a coupling of $(\sigma_n)_{n \geq 0}$ and L such that $\mathcal{L}(\sigma_n)$ a.s. converges to L . Fix such a coupling.

Fix a finite interval J of \mathbb{Z} and let I be the (random) smallest interval containing J whose endpoints are zeros of L . There is a random integer N which is almost surely finite such that on the event $N < n$ we have that $\mathcal{L}(\sigma_n)_i = L_i$ for all $i \in I$.

It follows from our earlier observations that $(X_i)_{i \in I}$ and $(Y_i^n)_{i \in I}$ have the same conditional distribution given L , σ_n , and the event that $N < n$. Indeed, under this conditioning both $(X_i)_{i \in I}$ and $(Y_i^n)_{i \in I}$ are uniformly distributed on the set of proper q -colorings of the subgraph of $\Gamma[L]$ induced by I by Lemmas 31 and 32. Since $J \subseteq I$ and $N < \infty$ a.s., we deduce that

$$(Y_i)_{i \in J} \stackrel{d}{=} (Y_i^n)_{i \in J} \xrightarrow[n \rightarrow \infty]{d} (X_i)_{i \in J}.$$

The claim follows since J was arbitrary. \square

The last proposition yields the following construction of a random coloring with law MalCol . Let L be the above iid sequence. Assign to the zero set of L a uniform proper q -coloring of \mathbb{Z} . Conditional on these colors and L , assign to the intervals between each pair of consecutive zeros i, j of L an independent, uniformly random proper q -coloring of the

constraint graph of $\mathcal{D}(L_i, \dots, L_j)$ conditioned to agree at i and j with the colors previously assigned. Then by Corollary 33, the resulting coloring of \mathbb{Z} is conditionally a uniform proper q -coloring of the constraint graph of $\Gamma[L]$ given L , from which it follows by Proposition 34 that the coloring has law MalCol . It remains to show, using this construction, that the colorings can be expressed both as finitary factors of iid processes and as functions of countable Markov chains. This is relatively routine, and we provide the details below.

Proposition 35. *There exists a countable state space S , a function $h : S \rightarrow \llbracket 1, q \rrbracket$, and a Markov process $(Y_i)_{i \in \mathbb{Z}}$ on S such that the process $(X_i)_{i \in \mathbb{Z}} = (h(Y_i))_{i \in \mathbb{Z}}$ has law $\text{MalCol}_{q,t}$. Moreover the return time of each state of S has exponential tail.*

Proof. Let L be an iid sequence of u -zero-weighted, t -geometric random variables. Conditional on L , let $X = (X_i)_{i \in \mathbb{Z}}$ be a uniform proper q -coloring of $\Gamma[L]$. By the previous proposition, X has law $\text{MalCol}_{q,t}$. It remains to express X as a function of a Markov process with the stated properties.

For each $k \in \mathbb{Z}$, let X^k and L^k denotes the shifted sequences

$$X^k = (X_i^k)_{i \in \mathbb{Z}} = (X_{i-k})_{i \in \mathbb{Z}} \quad \text{and} \quad L^k = (L_i^k)_{i \in \mathbb{Z}} = (L_{i-k})_{i \in \mathbb{Z}}.$$

For each k , let $f_k^+ = \min\{i > 0 : L_i^k = 0\}$ and $f_k^- = \max\{i \leq 0 : L_i^k = 0\}$. Let G_k be the subgraph of $\Gamma[L^k]$ induced by $\llbracket f_k^-, f_k^+ \rrbracket$. For each $k \in \mathbb{Z}$, let Y_k be the tuple

$$Y_k = \left(-f_k^-, f_k^+, G_k, (X_j^k)_{j \in \llbracket f_k^-, f_k^+ \rrbracket} \right).$$

The tuple Y_k takes values in the set S' of tuples (f^1, f^2, G, x) , where f^1 and f^2 are non-negative integers, G is a graph with vertex set $\llbracket -f^1, f^2 \rrbracket$, and x is a q -coloring of G . Note that S' is countable. Let S be the support of Y_0 on S' . We define $h : S \rightarrow \llbracket 1, q \rrbracket$ by setting $h(f^1, f^2, G, x) = x_0$, so that $(X_i)_{i \in \mathbb{Z}} = (h(Y_i))_{i \in \mathbb{Z}}$ as desired.

Clearly Y is stationary. To prove that Y is Markov, it suffices to show that $(Y_i)_{i > 0}$ and $(Y_i)_{i < 0}$ are conditionally independent given Y_0 . Since f_0^+ is the location of the first bubble endpoint of $\Gamma[L]$ to the right of the origin, it follows from the definition of a uniform proper q -coloring that $(Y_i)_{i > f_0^+}$ is conditionally independent of $(Y_i)_{i < f_0^+}$ given $Y_{f_0^+}$. Now $Y_{f_0^+}$ determines the sequence $Y_0, Y_1, \dots, Y_{f_0^+}$, as does Y_0 , so therefore $(Y_i)_{i > 0}$ and $(Y_i)_{i < 0}$ are conditionally independent given Y_0 . Thus Y is a Markov process.

That the return times have exponential tails follows in a straightforward manner. \square

Proposition 36. *There exists a ffd process with law $\text{MalCol}_{q,t}$ whose coding radius has exponential tail.*

Before proving Proposition 36, we show how to produce a f.i.i.d. uniform proper q -coloring of \mathbb{Z} with exponential tail on the coding radius, using a simple application of the technique of coupling from the past [51].

Let $(Z_i)_{i \in \mathbb{Z}}$ be an iid sequence, where each Z_i is chosen uniformly from the set $\{(a, b) \in \llbracket 1, q \rrbracket^2 : a \neq b\}$ of ordered pairs of distinct elements of $\llbracket 1, q \rrbracket$.

We claim that there is almost surely a unique sequence $X = (X_i)_{i \in \mathbb{Z}}$ satisfying the constraints

$$X_i = \begin{cases} Z_i^1 & \text{if } Z_i^1 \neq X_{i-1} \\ Z_i^2 & \text{if } Z_i^1 = X_{i-1}, \end{cases} \quad (43)$$

and furthermore that X can be computed as a finitary factor of Z with an exponential tail on the coding radius. Given its existence, it is easily seen that X is a uniform proper q -coloring of \mathbb{Z} .

First, notice that if $Z_i^1 \notin \{Z_{i-1}^1, Z_{i-1}^2\}$, then we must have $X_i = Z_i^1$. Thus, it is possible to compute X_i for arbitrary i by first finding the maximal $T_i \leq i$ such that $Z_{T_i}^1 \notin \{Z_{T_i-1}^1, Z_{T_i-1}^2\}$, setting $X_{T_i} = Z_{T_i}^1$, and then computing X_j for all $T_i \leq j \leq i$ by applying the recurrence (43). This shows that there is almost surely a unique solution X of (43), and that X can be computed as a finitary factor of Z with coding radii $(i - T_i)_{i \in \mathbb{Z}}$. Finally, we observe that $i - T_i$ is a geometric random variable, since

$$\mathbb{P}(i - T_i \geq n) = \prod_{k=1}^n \mathbb{P}\left(Z_{i-j+1}^1 \in \{Z_{i-j}^1, Z_{i-j}^2\}\right) = \left(\frac{2}{q}\right)^n.$$

Proof of Proposition 36. Consider the iid sequence $(Z_i, U_i, L_i)_{i \in \mathbb{Z}}$ where Z_i is chosen uniformly from the set $\{(a, b) \in \llbracket 1, q \rrbracket^2 : a \neq b\}$, U_i is chosen uniformly from $[0, 1]$, and L_i is a u -zero-weighted, t -geometric random variable. We construct the desired process in two steps. In the first step, we assign to $(i \in \mathbb{Z} : L_i = 0)$ a uniform proper q -coloring by applying the above procedure to $(Z_i : L_i = 0)$.

In the second step we assign colors to $(i \in \mathbb{Z} : L_i > 0)$. For such i let

$$f_i^- = \max\{j < i : L_j = 0\} \quad \text{and} \quad f_i^+ = \min\{j > i : L_j = 0\};$$

note that $i \in (f_i^-, f_i^+)$ and that f_i^\pm were assigned colors in the previous step. Conditional on the previous step, let $X^{\llbracket f_i^-, f_i^+ \rrbracket}$ be a uniform proper q -coloring of the constraint graph of $\mathcal{D}((\ell_k)_{k \in \llbracket f_i^-, f_i^+ \rrbracket})$ consistent with the colors assigned to f_i^\pm . Assume that $X^{\llbracket f_i^-, f_i^+ \rrbracket}$ is defined on the probability space $[0, 1]$ and assign to i the color $X_i^{\llbracket f_i^-, f_i^+ \rrbracket}(U_{f_i^-})$, i.e., $U_{f_i^-}$ is used as a seed to generate the random coloring of $\llbracket f_i^-, f_i^+ \rrbracket$.

It is easy to see that the coloring of \mathbb{Z} thus obtained is a finitary factor of the sequence $(Z_i, U_i, L_i)_{i \in \mathbb{Z}}$ and that its coding radius has exponential tail. By Lemma 31 it follows that,

conditional on L , the coloring thus produced is a uniform proper q -coloring of $\Gamma[L]$. Thus by Proposition 34, the coloring has law $\text{MalCol}_{q,t}$. \square

Recall the tuning equation (1), which is

$$qt(t^k - 1) = (t + 1)(t^{k+1} - 1).$$

Proof of Theorem 1. Combine Propositions 13, 20, 35, and 36 to conclude that if for integers $q \geq 3$ and $k \geq 1$ there exists $t \in (0, 1)$ satisfying the tuning equation (1), then the theorem holds in the case (k, q) .

That such a t exists for $(k, q) = (1, 5)$, $(2, 4)$, $(3, 3)$, and all larger k and q follows since these integers satisfy $qk > 2(k + 1)$, which implies that the polynomial

$$qt(1 - t^k) - (t + 1)(1 - t^{k+1})$$

is negative at $t = 0$, vanishes at $t = 1$, and has negative derivative there. Thus it has a root in $(0, 1)$, providing the desired solution of the tuning equation (1). \square

7 Painting algorithm and conditioning

The primary purpose of this section is to verify correctness of the painting algorithm (Algorithm 1). Recall that its input consists of positive integers q and k satisfying $qk > 2(k + 1)$ and its output is a random q -coloring of \mathbb{Z} , which we claim is k -dependent.

Proposition 37.

(i) For all positive integers q and k satisfying $qk > 2(k + 1)$, there exists a unique $t = t(q, k) \in (0, 1)$ such that (q, k, t) satisfies the tuning equation (1).

(ii) The output of Algorithm 1 has law $\text{MalCol}_{q,t(q,k)}$.

The other purpose of this section is to establish that if one conditions the 1-dependent q -coloring from the main theorem to only use colors in $\llbracket 1, q - 1 \rrbracket$, this results in the 2-dependent $(q - 1)$ -coloring from the theorem, and no other pairs of colorings from the theorem are related in this manner. This is equivalent to the following.

Proposition 38. *The only pairs (k, q) and (k', q') such that there exists $t \in (0, 1)$ for which (q, k, t) and (q', k', t) both satisfy the tuning equation (1) are*

$$(k, q) = (1, q) \quad \text{and} \quad (k', q') = (2, q - 1), \quad q \geq 5.$$

First we establish the following properties of the tuning equation.

Lemma 39. *Let q and k be integers with $k > 0$.*

(i) *There exists $t \in (0, 1)$ satisfying (1) if and only if $qk > 2(k + 1)$.*

(ii) *There is at most one $t \in [0, 1)$ satisfying (1).*

We remark that this lemma and others in this section hold more generally when q and k are real-valued, with the same proofs.

Proof. The ‘if’ direction of part (i) was established in the proof of Theorem 1. For the ‘only if’ direction, let

$$f_{q,k}(t) := qt(1 - t^k) - (t + 1)(1 - t^{k+1}).$$

We show that when $q \leq q' := 2(k + 1)/k$, the function $f_{q,k}(t)$ has no zeros in $(0, 1)$. Indeed, in this case $f_{q,k}(t) \leq f_{q',k}(t)$ and since

$$f''_{q',k}(t) = (k + 1)t^{k-1}((k + 2)t - (q - 1)k), \quad (44)$$

we have that $f''_{q',k}(t) = -(k + 1)(k + 2)(1 - t)t^{k-1}$ and therefore $f_{q',k}$ is strictly concave on $(0, 1)$. Combined with $f_{q',k}(1) = f'_{q',k}(1) = 0$, it follows that $f_{q',k}(t) < 0$ for all $t \in (0, 1)$. Thus $f_{q,k}(t) < 0$ when $q \leq 2(k + 1)/k$ and $t \in (0, 1)$, implying the ‘only if’ direction of (i).

It remains to establish part (ii) in the case when q and k satisfy $qk > 2(k + 1)$. Then $(q - 1)t > k + 2 \geq (k + 2)t$ for all $t \geq 1$, so $f''_{q,k}(t) < 0$ for $t \in [0, 1]$ by (44). This implies that $f_{q,k}$ has at most two zeros in $[0, 1]$, counting the zero at 1, so therefore $f_{q,k}$ has at most one zero in $[0, 1)$. \square

Proof of Proposition 37. The first part of the proposition follows from the Lemma 39.

For the second part, let Γ_{alg} be the graph on \mathbb{Z} in which integers $i < j$ are adjacent iff i and j are the first two elements of $[[i, j]]$ assigned colors by the algorithm. It is straightforward to verify that the bubble endpoints of Γ_{alg} are the integers assigned colors during Stage 1 of the algorithm, and that conditional on Γ_{alg} , the coloring X is a uniform proper q -coloring of Γ_{alg} .

Thus by Proposition 34, it suffices to show that Γ_{alg} is equal in law to $\Gamma[L]$ where $L = (L_i)_{i \in \mathbb{Z}}$ is an iid sequence of u -bubble-weighted, t -geometric random variables. The bubble endpoints of the two graphs have the same law by Corollary 33(i), since

$$\mathbb{P}(L_i \geq 1) = \frac{t/(1 - t)}{u + t/(1 - t)} = \frac{t(q - 2)}{q - 1 - t} = s,$$

where we have substituted $u = \frac{q-1}{q-2}$ in the second equality.

The result will follow from the claim that, for all integers $i < j$, the subgraphs of $\Gamma[L]$ and Γ_{alg} induced by $[[i, j]]$ have the same conditional law given that i and j are bubble endpoints

of $\Gamma[L]$ and Γ_{alg} , respectively. We establish this claim by showing that both conditional laws are equal to the law of the constraint graph of a Mallows-distributed permutation of $\llbracket i, j \rrbracket$. For $\Gamma[L]$ this is straightforward to verify using Lemma 32 together with the observation that the conditional law of a u -zero-weighted, t -geometric random variable conditioned to be positive does not depend on u . To prove the same for Γ_{alg} , we consider the permutation σ of $\llbracket i, j \rrbracket$ for which $\sigma(i) - i$ equals the number of elements of $\llbracket i, j \rrbracket$ that have been assigned colors prior to i . We regard $\sigma(i)$ as the arrival time of i . On the event that i and j are bubble endpoints of Γ_{alg} , the subgraph of Γ_{alg} induced by $\llbracket i, j \rrbracket$ is seen to coincide with the constraint graph of σ .

The integer with arrival time k appears in position $\tilde{\mathcal{L}}(\sigma)_k$ from the right in the subsequence of $\llbracket i, j \rrbracket$ consisting of integers that have previously arrived. Thus, the random choices made by the algorithm ensure that:

- $\tilde{\mathcal{L}}(\sigma)_k$ is $(k - i)$ -truncated t -geometric, and
- $(\tilde{\mathcal{L}}(\sigma)_k)_{k \in \llbracket i, j \rrbracket}$ is a sequence of independent random variables.

Thus by Lemma 12, the permutation σ has law Mal_t . It follows that for all integers $i < j$ the conditional laws of the subgraphs of $\Gamma[L]$ and Γ_{alg} induced by $\llbracket i, j \rrbracket$, given that i and j are bubble endpoints of the respective graphs, coincide.

From this we deduce that the conditional laws of $\Gamma[L]$ and Γ_{alg} given their respective sets of bubble endpoints are the same. Since we have previously established that these sets have the same distribution, it follows that $\Gamma[L]$ and Γ_{alg} are equal in law. \square

In the introduction, we claimed that if one conditions on the absence of color q in the 1-dependent q -colorings from the main theorem, the resulting $(q - 1)$ -coloring is 2-dependent, and that these are the only pairs of colorings in the theorem related to one another by conditioning in this manner. This follows from the properties of the tuning equation (1) established in the following lemma.

Lemma 40. *Consider positive integers q and k and a real number $t \in (0, 1)$.*

- (i) *The triple $(q, 1, t)$ satisfies (1) if and only if $(q - 1, 2, t)$ does.*
- (ii) *Given (q, t) , there is at most one k satisfying (1).*
- (iii) *Suppose that $qk > 2(k+1)$ and let $t(q, k)$ denote the unique solution of (1) satisfying $0 < t(q, k) < 1$. Then $t(q, k) > \frac{1}{q-1}$ and $\lim_{k \rightarrow \infty} t(q, k) = \frac{1}{q-1}$.*
- (iv) *Given t with $0 < t < 1$, there is at most one pair of integers (q, k) with $q > 1$ and $k \geq 2$ satisfying (1).*

Proof. Part (i) is an easy calculation. Part (ii) follows since (1) is equivalent to

$$t^{k+1} = \frac{(q-1)t-1}{q-1-t},$$

and this also implies part (iii).

Finally we establish part (iv). Solving for q in (1) yields

$$q = \left(1 + \frac{1}{t}\right) \left(t + \frac{1-t}{1-t^k}\right). \quad (45)$$

The right side of (45) is strictly decreasing in k . Furthermore as $k \rightarrow \infty$ it tends to $\frac{1}{t} + 1$, whereas when $k = 2$ it evaluates to $\frac{1}{t} + 1 + t$. Therefore if (q, k, t) is any solution of (45) with $q > 1$ and $k \geq 2$, we must have that

$$\frac{1}{t} + 1 < q \leq \frac{1}{t} + 1 + t, \quad k > 2. \quad (46)$$

Since $t < 1$ there can be at most one integer q satisfying (46), in which case by part (ii) there is at most one k satisfying (1). \square

8 Bit-finitary factors

In this section we prove Theorem 2.

Proof of Theorem 2. Let F and U be as in the statement of the theorem. For each $k \geq 1$, let

$$p(k) = \mathbb{P}(r_0(U) \leq k)$$

be the probability that $F(U)_0$ is determined by the restriction of U to $[-k, k] \times [0, d]$ for some $d \geq 0$. For each $k \geq 1$, let $d(k)$ be the minimal value of d such that the probability that $F(U)_0$ is determined by the restriction of U to $[-k, k] \times [0, d]$ is at least $p(k) - 2^{-k}$. In particular, $d(k) = 0$ if $p(k) \leq 2^{-k}$. Let $R_i(U)$ be minimal such that $F(U)_i$ is determined by the restriction of U to

$$S_i(U) := [i - R_i(U), i + R_i(U)] \times [0, d(R_i(U))].$$

The definition of $r_i(U)$ and $R_i(U)$ ensure that

$$\mathbb{P}(R_0(U) \geq k) \leq \mathbb{P}(r_0(U) \geq k) + 2^{-k}$$

for every $k \geq 0$ and hence that

$$\mathbb{E}[R_0(U)] = \sum_{k \geq 1} \mathbb{P}(R_0(U) \geq k) \leq \mathbb{E}[r_0(U)] + 1 < \infty.$$

For each $i \in \mathbb{Z}$, we define

$$T_i^+(U) = \sup \{j \geq i : R_k(U) \geq j - k \text{ for some } k \leq i\}$$

and

$$T_i^-(U) = \inf \{j \leq i : R_k(U) \geq k - j \text{ for some } k \geq i\}.$$

Similarly, for each $i, j \in \mathbb{Z}$ with $j \leq T_i^+$, we define

$$D_{i,j}(U) = \inf \{D \geq 0 : S_k \cap (\mathbb{N} \times \{j\}) \subseteq [0, D] \times \{j\} \text{ for all } k \leq i\}$$

Since $\mathbb{E}[R_0(U)]$ is finite, it follows from Borel-Cantelli that for each $i \in \mathbb{Z}$ there exist at most finitely many $k \in \mathbb{Z}$ for which $R_k(U) \geq |k - i|$ a.s., and this easily implies that $T_i^+(U)$, $T_i^-(U)$, and $D_{i,j}(U)$ are finite for every $i \in \mathbb{Z}$ and every $j \leq T_i^+$ a.s.

Let \mathcal{K} be the set of pairs (K, y) , where K is a finite subset $\mathbb{Z} \times \mathbb{N}$ and y is a function from K to $\{0, 1\}$. Clearly \mathcal{K} is countable. For each $i \in \mathbb{Z}$, we define a finite set $K_i(U) \subset \mathbb{Z} \times \mathbb{N}$ by

$$K_i(U) = \{(j, k) \in \mathbb{Z} \times \mathbb{N} : j \in [T_0^-(U), T_0^+(U)], k \in [0, D_{i,j}(U)]\},$$

and let

$$\bar{K}_i(U) = \{(j, k) \in \mathbb{Z} \times \mathbb{N} : j \in [T_0^-(U) - i, T_0^+(U) - i], k \in [0, D_{i,j}(U)]\}.$$

We define a factor $G : \{0, 1\}^{\mathbb{Z} \times \mathbb{N}} \rightarrow \mathcal{K}^{\mathbb{Z}}$ by setting

$$G_i(U) = (\bar{K}_i(U), U|_{K_i(U)}).$$

Fix an element $a_0 \in A$ arbitrarily, and define $h : \mathcal{K} \rightarrow A$ by letting $h(K, y) = a$ if $a \in A$ is such that $F(x)_0 = a$ for a.e. x such that $x|_K = y$, and letting $h(K, y) = a_0$ if no such $a \in A$ exists. The construction of G ensures that the value of $F(U)_0$ is determined by the restriction of U to $K_0(U)$, and it follows that $H \circ G = F$, where $H : \mathcal{K}^{\mathbb{Z}} \rightarrow A^{\mathbb{Z}}$ is the factor

$$H_i((K_j, y_j)_{j \in \mathbb{Z}}) = h(K_i, y_i).$$

It remains to prove only that the process $(G_i(U))_{i \in \mathbb{Z}}$ is a Markov chain. To see this, observe that the subsets of $\mathbb{Z} \times \mathbb{N}$ that are queried in order to compute $(G_i(U))_{i \geq 1}$ and $(G_i(U))_{i \leq -1}$ have intersection contained in the subset of $\mathbb{Z} \times \mathbb{N}$ that is queried to compute $G_0(U)$. It follows that $(G_i(U))_{i \geq 1}$ and $(G_i(U))_{i \leq -1}$ are conditionally independent given $G_0(U)$, so that $(G_i(U))_{i \in \mathbb{Z}}$ is indeed a Markov chain. \square

9 Compact Markov chains

By a **process** on a space S , we mean a random element of $S^{\mathbb{Z}}$ that is measurable with respect to the product Borel σ -algebra.

Question 1. *Does there exist a stationary Markov process $(X_i)_{i \in \mathbb{Z}}$ on a compact metric space (S, d) , an integer $k > 0$, and a real number $\varepsilon > 0$ such that*

(i) X_0 and X_k are independent, and

(ii) $d(X_0, X_1) \geq \varepsilon$ almost surely?

It is trivial to construct chains that satisfy either one of the two conditions. Here are two somewhat interesting examples. Firstly, let $S = [0, 1]^2$. Conditional on $X_0 = (u, v)$, let $X_1 = (v, U)$, where U is uniformly distributed on $[0, 1]$. This satisfies (i) with $k = 2$, but not (ii). Secondly, let S be the unit sphere $\{x \in \mathbb{R}^3 : \|x\|_2 = 1\}$. Conditional on X_0 , let X_1 be uniformly distributed on the circle $\{y \in S : \langle y, X_0 \rangle = 0\}$ (where $\langle \cdot, \cdot \rangle$ is the standard inner product on \mathbb{R}^3). This satisfies (ii) (with the Euclidean metric and $\varepsilon = 1$) but not (i).

Proof of Proposition 3. Let π denote the law of X_0 . Since S is Polish, there is a Markov kernel that is a regular conditional probability for $\mathbb{P}(X_1 \in A \mid X_0)$. This gives rise to a Markov transition operator P on $L^2(S, \pi)$. Since X is reversible, P is self-adjoint.

Since X_0 is independent of X_k , it follows that $P^i = P^k$ for all $i \geq k$. We claim that if $k \geq 2$, then $P^k = P^{k-1}$. Indeed, for all $f \in L^2(S, \pi)$,

$$\begin{aligned} \|P^{k-1}f - P^k f\|^2 &= \|P^{k-1}f\|^2 - 2\langle P^{k-1}f, P^k f \rangle + \|P^k f\|^2 \\ &= \langle f, P^{2k-2}f \rangle - 2\langle f, P^{2k-1}f \rangle + \langle f, P^{2k}f \rangle. \end{aligned} \quad (47)$$

Since $k \geq 2$, we have that $2k - 2 \geq k$ and therefore $P^{2k-2} = P^{2k-1} = P^{2k}$. Thus (47) vanishes for all f , so $P^{k-1} = P^k$. Hence by induction it follows that $P = P^2$.

Consider a cover of S by balls of radius $\varepsilon/4$. Fix a partition of unity $\{f_i\} \subset L^2(S, \pi)$ subordinate to this cover. Then Pf_i and f_i have disjoint support for all i . Thus $\langle Pf_i, f_i \rangle = 0$, implying that

$$\|Pf_i\|^2 = \langle Pf_i, Pf_i \rangle = \langle Pf_i, f_i \rangle = 0.$$

Consequently $\sum_i Pf_i = 0$, so P maps the constant 1 function to the zero function. But this contradicts the fact that P is a Markov transition operator. \square

Every process satisfying the conditions of Question 1 yields a finitely dependent coloring. Indeed, cover S by balls of radius $\varepsilon/2$, extract a finite subcover, and take any partition

$\{P_i\}_{i=1}^q$ of S subordinate to the finite subcover. Then the sequence of elements of the partition visited by the Markov process is a finitely dependent q -coloring.

Schramm established that no hidden-Markov finitely dependent q -coloring exists, for any q . (This was first published in [30, Prop. 3].) Combining this with the observation in the previous paragraph implies that there is no *finite* Markov chain satisfying the conditions of Question 1. On the other hand, any stationary stochastic process can trivially be expressed as a function of an *uncountable*-state Markov chain.

Proof of Proposition 4. We first describe the process $(X_n)_{n \in \mathbb{Z}}$, and then explain why it is Markov. Let $(\omega_{i,j} : i \in \mathbb{Z}, j \in \mathbb{N})$ be iid Bernoulli($\frac{1}{2}$) random bits indexed by the discrete half plane. The process $(X_n)_{n \in \mathbb{Z}}$ will be a deterministic function of these bits. Let

$$h(n) := \min\{j \in \mathbb{N} : \omega_{n-1,j} \neq \omega_{n,j}\}, \quad n \in \mathbb{Z},$$

be the height of the first discrepancy between the bits in columns $n - 1$ and n . Note that $h(n)$ is a.s. finite. Let X_n be the 2-by- $h(n)$ matrix

$$\left(\omega_{i,j} : (i,j) \in \{n-1, n\} \times \{1, \dots, h(n)\}\right),$$

consisting of bits in the two columns up to the discrepancy. The state space S is the set of binary matrices of width 2 with the two columns differing exactly in the last row; note that this set is countable. We equip S with the discrete metric $d(x, y) := \mathbf{1}[x \neq y]$.

We now show that the process $(X_n)_{n \in \mathbb{Z}}$ is Markov. Consider the σ -algebras

$$\mathcal{F}_n = \sigma((\omega_{n-1,j}, \omega_{n,j}) : j \in \mathbb{N}), \quad n \in \mathbb{Z}.$$

By considering cylinder events, it is easy to check that $\sigma(\mathcal{F}_{n+1}, \mathcal{F}_{n+2}, \dots)$ and $\sigma(\mathcal{F}_{n-1}, \mathcal{F}_{n-2}, \dots)$ are conditionally independent given \mathcal{F}_n . Since X_m is measurable with respect to \mathcal{F}_m for all $m \in \mathbb{Z}$, the Markov property follows.

Clearly the sequence $(X_n)_{n \in \mathbb{Z}}$ is stationary. Furthermore X_0 is independent of X_2 , since they are functions of disjoint sets of independent bits. All that remains is to verify that $X_n \neq X_{n+1}$. Suppose to the contrary that $X_n = X_{n+1}$. Then $h(n) = h(n+1)$; call the common value j . Since $X_n = X_{n+1}$ we have that $(\omega_{n-1,j}, \omega_{n,j}) = (\omega_{n,j}, \omega_{n+1,j})$. But since $h(n) = j$ we have that $\omega_{n-1,j} \neq \omega_{n,j}$. Thus we have derived a contradiction. \square

10 Higher dimensions and shifts of finite type

Proofs of Corollaries 5 and 6. A simple modification of the proofs of [30, Corollaries 5 and 6] establishes Corollaries 5 and 6, respectively. Namely, replace the 1-dependent 4-coloring

used in the proof of Corollary 20 of [30] with the fffd 1-dependent 5-coloring with exponential tail on the coding radius from Theorem 1. This results in an fffd process with exponential tails, since the maximum of a finite (deterministic) number of independent random variables with exponential tails still has exponential tails. \square

11 Open problems

- (i) For the pairs $(k, q) = (1, 4)$ and $(2, 3)$, can the k -dependent q -coloring be expressed as a finitary factor of iid with finite expected coding radius? (We suspect not.)
- (ii) For the pairs $(k, q) = (1, 5), (2, 4), (3, 3)$, and for all larger k and q , is there a unique stationary color-symmetric k -dependent q -coloring of \mathbb{Z} ?
- (iii) Are there finitary factors of iid that are not expressible as bit-finitary factors of iid?
- (iv) Is there a discrete-time Markov chain on a compact metric space that mixes perfectly in finitely many steps and always moves by at least some fixed distance?
- (v) For every word $x \in \mathbb{Z}^n$ and every $0 \leq i \leq \binom{n}{2}$, is there a ‘natural’ bijection between proper buildings of x having i and $\binom{n}{2} - i$ inversions?

References

- [1] J. Aaronson, D. Gilat, and M. Keane. On the structure of 1-dependent Markov chains. *J. Theoret. Probab.*, 5(3):545–561, 1992.
- [2] J. Aaronson, D. Gilat, M. Keane, and V. de Valk. An algebraic construction of a class of one-dependent processes. *Ann. Probab.*, 17(1):128–143, 1989.
- [3] N. Alon and O. N. Feldheim. A note on general sliding window processes. *Electron. Commun. Probab.*, 19:no. 66, 7, 2014.
- [4] N. Alon and J. H. Spencer. *The probabilistic method*. John Wiley & Sons, 2004.
- [5] N. Auger, M. Bouvel, C. Nicaud, and C. Pivoteau. Analysis of algorithms for permutations biased by their number of records. *arXiv preprint arXiv:1605.02905*, 2016.
- [6] R. Basu and N. Bhatnagar. Limit theorems for longest monotone subsequences in random mallows permutations. *arXiv preprint arXiv:1601.02003*, 2016.

- [7] E. F. Beckenbach, G. Pólya, D. H. Lehmer, and M. Phister. *Applied combinatorial mathematics*. Wiley New York, 1964.
- [8] I. Benjamini, N. Berger, C. Hoffman, and E. Mossel. Mixing times of the biased card shuffling and the asymmetric exclusion process. *Transactions of the American Mathematical Society*, 357(8):3013–3029, 2005.
- [9] N. Bhatnagar and R. Peled. Lengths of monotone subsequences in a Mallows permutation. *Probab. Theory Related Fields*, 161(3-4):719–780, 2015.
- [10] M. Bóna. *Combinatorics of permutations*. CRC Press, 2012.
- [11] A. Borodin, P. Diaconis, and J. Fulman. On adding a list of numbers (and other one-dependent determinantal processes). *Bull. Amer. Math. Soc. (N.S.)*, 47(4):639–670, 2010.
- [12] M. Braverman and E. Mossel. Sorting from noisy information. *arXiv preprint arXiv:0910.1191*, 2009.
- [13] E. I. Broman. One-dependent trigonometric determinantal processes are two-block-factors. *Annals of probability*, pages 601–609, 2005.
- [14] R. M. Burton, M. Goulet, and R. Meester. On 1-dependent processes and k -block factors. *Ann. Probab.*, 21(4):2157–2168, 10 1993.
- [15] H.-H. Chern, H.-K. Hwang, and Y.-N. Yeh. Distribution of the number of consecutive records. *Random Structures and Algorithms*, 17(3-4):169–196, 2000.
- [16] V. de Valk. The maximal and minimal 2-correlation of a class of 1-dependent 0-1 valued processes. *Israel J. Math.*, 62(2):181–205, 1988.
- [17] V. de Valk. A problem on 0 – 1 matrices. *Compositio mathematica*, 71(2):139–179, 1989.
- [18] V. de Valk. Hilbert space representations of m -dependent processes. *The Annals of Probability*, pages 1550–1570, 1993.
- [19] V. de Valk. *One-dependent processes: two-block factors and non-two-block factors*, volume 85 of *CWI Tract*. Stichting Mathematisch Centrum, Centrum voor Wiskunde en Informatica, Amsterdam, 1994.
- [20] P. Diaconis, A. Ram, et al. *Analysis of systematic scan Metropolis algorithms using Iwahori-Hecke algebra techniques*. Department of Statistics, Stanford University, 2000.

- [21] P. Erdos and L. Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. *Infinite and finite sets*, 10(2):609–627, 1975.
- [22] J. A. Fill. On the distribution of binary search trees under the random permutation model. *Random Structures & Algorithms*, 8(1):1–25, 1996.
- [23] A. Gandolfi, M. Keane, and V. De Valk. Extremal two-correlations of two-valued stationary one-dependent processes. *Probability Theory and Related Fields*, 80(3):475–480, 1989.
- [24] C. Garban. Quantum gravity and the kpz formula. *arXiv preprint arXiv:1206.0212*, 2012.
- [25] N. Glick. Breaking records and breaking boards. *American Mathematical Monthly*, pages 2–26, 1978.
- [26] A. Gnedin and G. Olshanski. The two-sided infinite extension of the Mallows model for random permutations. *Advances in Applied Mathematics*, 48(5):615–639, 2012.
- [27] L. Heinrich. Asymptotic Expansions in the Central Limit Theorem for a Special Class of m -Dependent Random Fields II–Lattice Case. *Mathematische Nachrichten*, 145(1):309–327, 1990.
- [28] W. Hoeffding, H. Robbins, et al. The central limit theorem for dependent random variables. *Duke math. J*, 15(3):773–780, 1948.
- [29] A. E. Holroyd. One-dependent coloring by finitary factors, arXiv:1411.1463.
- [30] A. E. Holroyd and T. M. Liggett. Finitely dependent coloring, arXiv:1403.2448.
- [31] A. E. Holroyd and T. M. Liggett. Symmetric 1-dependent colorings of the integers. *Electron. Commun. Probab.*, 20:no. 31, 8, 2015.
- [32] A. E. Holroyd, O. Schramm, and D. B. Wilson. Finitary coloring, arXiv:1412.2725.
- [33] I. Ibragimov and Y. V. Linnik. Independent and stationarily connected variables. *Izdat. Nauka, Moscow*, 1965.
- [34] I. A. Ibragimov and Y. V. Linnik. *Independent and stationary sequences of random variables*. Wolters-Noordhoff Publishing, Groningen, 1971. With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, translated from original in Russian (1965).
- [35] S. Janson. Runs in m -dependent sequences. *Ann. Probab.*, 12(3):805–818, 1984.

- [36] S. Janson et al. Renewal theory for m -dependent variables. *The Annals of Probability*, 11(3):558–568, 1983.
- [37] S. Janson et al. On degenerate sums of m -dependent variables. *Journal of Applied Probability*, 52(4):1146–1155, 2015.
- [38] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.
- [39] D. E. Knuth. *The art of computer programming: sorting and searching*, volume 3. Pearson Education, 1998.
- [40] I. Kortchemski. Asymptotic behavior of permutation records. *Journal of Combinatorial Theory, Series A*, 116(6):1154–1166, 2009.
- [41] G. F. Lawler. *Conformally invariant processes in the plane*. Number 114. American Mathematical Soc., 2008.
- [42] A. Levy. Finitely dependent insertion processes. *arXiv preprint arXiv:1510.08995*, 2015.
- [43] T. M. Liggett, R. H. Schonmann, A. M. Stacey, et al. Domination by product measures. *The Annals of Probability*, 25(1):71–95, 1997.
- [44] N. Linial. Distributive graph algorithms global solutions from local data. In *Proceedings of the 28th Annual Symposium on Foundations of Computer Science*, SFCS '87, pages 331–335, Washington, DC, USA, 1987. IEEE Computer Society.
- [45] C. L. Mallows. Non-null ranking models. I. *Biometrika*, 44:114–130, 1957.
- [46] F. Matúš. On two-block-factor sequences and one-dependence. *Proceedings of the American Mathematical Society*, 124(4):1237–1242, 1996.
- [47] F. Matus. Combining m -dependence with markovness. In *Annales de l'IHP Probabilités et statistiques*, volume 34, pages 407–423, 1998.
- [48] C. Mueller and S. Starr. The length of the longest increasing subsequence of a random Mallows permutation. *Journal of Theoretical Probability*, 26(2):514–540, 2013.
- [49] M. Naor. A lower bound on probabilistic algorithms for distributive ring coloring. *SIAM Journal on Discrete Mathematics*, 4(3):409–412, 1991.

- [50] G. O'Brien. Scaling transformations for $\{0, 1\}$ -valued sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 53(1):35–49, 1980.
- [51] J. Propp and D. Wilson. Coupling from the past: a user's guide. *Microsurveys in Discrete Probability*, 41:181–192, 1998.
- [52] D. J. Rudolph. A mixing markov chain with exponentially decaying return times is finitarily bernoulli. *Ergodic Theory and Dynamical Systems*, 2(01):85–97, 1982.
- [53] L. Rüschendorf and V. de Valk. On regression representations of stochastic processes. *Stochastic Processes and their Applications*, 46(2):183–198, 1993.
- [54] M. Smorodinsky. Finitary isomorphism of m -dependent processes. *Symbolic Dynamics and Its Applications (New Haven, Conn., 1991)*, pages 373–376, 1992.
- [55] R. P. Stanley. Enumerative Combinatorics. vol. 1, vol. 49 of Cambridge Studies in Advanced Mathematics, 1997.
- [56] S. Starr. Thermodynamic limit for the Mallows model on S_n . *Journal of mathematical physics*, 50(9), 2009.
- [57] S. Starr and M. Walters. Phase Uniqueness for the Mallows Measure on Permutations. *ArXiv e-prints*, Feb. 2015, 1502.03727.
- [58] V. Strassen. The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, pages 423–439, 1965.

FINITELY DEPENDENT INSERTION PROCESSES

AVI LEVY

ABSTRACT. A q -coloring of \mathbb{Z} is a random process assigning one of q colors to each integer in such a way that consecutive integers receive distinct colors. A process is k -dependent if any two sets of integers separated by a distance greater than k receive independent colorings. Holroyd and Liggett constructed the first stationary k -dependent q -colorings by introducing an insertion algorithm on the complete graph K_q . We extend their construction from complete graphs to weighted directed graphs. We show that complete multipartite analogues of K_3 and K_4 are the only graphs whose insertion process is finitely dependent and whose insertion algorithm is consistent. In particular, there are no other such graphs among all unweighted graphs and among all loopless complete weighted directed graphs. Similar results hold if the consistency condition is weakened to eventual consistency. Finally we show that the directed de Bruijn graphs of shifts of finite type do not yield k -dependent insertion processes, assuming eventual consistency.

1 Introduction

A **proper q -coloring of \mathbb{Z}** is a sequence of colors $(x_i)_{i \in \mathbb{Z}}$ with $x_i \in [q] := \{1, \dots, q\}$ such that $x_i \neq x_{i+1}$ for all i . A random q -coloring $(X_i)_{i \in \mathbb{Z}}$ is **stationary** if $(X_i)_{i \in \mathbb{Z}}$ and $(X_{i+1})_{i \in \mathbb{Z}}$ are equal in law. A stationary q -coloring is **k -dependent** if $(X_i)_{i < 0}$ and $(X_i)_{i \geq k}$ are independent, and **finitely dependent** if it is k -dependent for some $k \geq 0$.

The simplest examples of stationary finitely dependent processes are the **block factors**. These are stochastic processes of the form $\{f(Y_i, \dots, Y_{i+k})\}_{i \in \mathbb{Z}}$ where f is deterministic and $\{Y_i\}_{i \in \mathbb{Z}}$ form an i.i.d. sequence. In the 1960s, Ibragimov and Linnik first suggested that there may exist non-block factor stationary finitely dependent processes [15, 16]. Since then, examples of such processes have been constructed by several authors in the course of studying properties of finitely dependent processes [1, 2, 7, 9, 17]. Until recently, it has been believed that most ‘natural’ finitely dependent processes are block factors [6].

Yet block factors have subtle limitations. For example, these processes are never supported on proper colorings [3]. It turns out that finitely dependent processes do not have this limitation, although this fact is highly non-obvious and remains to be fully understood. This was discovered by Holroyd and Liggett in a recent breakthrough [13],

2010 *Mathematics Subject Classification*. 60G10; 05C15; 60C05.

Key words and phrases. Proper coloring, directed graph, weighted graph, m -dependence, stationary process, shift of finite type.

in which they disproved a conjecture of Schramm [14] by showing that stationary finitely dependent colorings of the integers exist. These are perhaps the first natural non-block factor finitely dependent processes.

Specifically, Holroyd and Liggett constructed symmetric 3- and 4-colorings with these properties. It is remarkable that, while their construction produces a q -coloring for each integer $q \geq 2$, only when $q \in \{3, 4\}$ is the coloring finitely dependent. Using a more complicated construction, these authors later obtained symmetric q -colorings for all $q \geq 4$ [12].

As described in [13], the q -colorings therein have the following characterization. For each integer $q \geq 2$, let $Z = (Z_1, \dots, Z_n)$ be a sequence of independent random variables each taking the values $1, 2, \dots, q$ with equal probability. Let σ be an independent uniformly random permutation of $1, \dots, n$, which we interpret as meaning that the symbol $Z_{\sigma(i)}$ arrives at time i . Let E be the event that, for every time $t = 1, \dots, n$, the subsequence of Z formed by those symbols that arrived up to time t (ordered as in the original sequence Z) forms a proper coloring (i.e. no two consecutive elements in the subsequence are equal). Then the conditional law of Z given E equals the law of (X_1, \dots, X_n) , where X is the q -coloring constructed in [13].

It was observed by Holroyd (personal communication) that the proper coloring condition in the previous paragraph may be replaced by a graph adjacency condition. The case of a q -coloring corresponds to the complete graph with vertex set $\{1, \dots, q\}$, denoted K_q . A general graph will encode which pairs of vertices may appear consecutively. See e.g. [5, Example 2.5] for more on this perspective. Since few stationary finitely dependent colorings are currently known, it is natural to pursue this generalization in the search for new finitely dependent processes.

Fix a finite graph G containing at least one edge. Let $Z = (Z_1, \dots, Z_n)$ be a sequence of independent uniformly random vertices of G . Let σ be an independent uniformly random permutation of $1, \dots, n$, which we interpret as meaning that the vertex $Z_{\sigma(i)}$ arrives at time i . Let E be the event that, for every time $t = 1, \dots, n$, the subsequence of Z formed by those vertices that arrived up to time t (ordered as in the original sequence Z) forms a path in G . Then let (Y_1, \dots, Y_n) denote a random tuple whose law equals the conditional law of Z given E . This generalizes the construction of the q -coloring in [13].

To produce a stochastic process $(Y_i)_{i \in \mathbb{Z}}$ from these random tuples, we must apply a limiting procedure. Let P_n denote the probability mass function of (Y_1, \dots, Y_n) . If for all n the mass functions of (Y_1, \dots, Y_n) and (Y_2, \dots, Y_{n+1}) equal P_n , then by the Kolmogorov Extension Theorem [10] there exists a unique stochastic process $(\tilde{Y}_i)_{i \in \mathbb{Z}}$ such that $(\tilde{Y}_1, \dots, \tilde{Y}_n)$ has mass function P_n for all $n \geq 1$. In this case we say that $\{P_n\}_{n=1}^\infty$

is **consistent** and that G satisfies property (C). More generally, we say that the mass functions are **eventually consistent** and that G satisfies property (EC) if the preceding condition holds for all sufficiently large n . In this case, we construct a process $(\tilde{Y}_i)_{i \in \mathbb{Z}}$ in the same manner as before by taking a limit over sufficiently large n . We call $(\tilde{Y}_i)_{i \in \mathbb{Z}}$ the **insertion process** associated to G .

Holroyd and Liggett showed that for all $q \geq 2$, the complete graph K_q has property (C) [13, Proposition 10]. Furthermore, they discovered that the insertion process associated to K_4 is 1-dependent and the insertion process associated to K_3 is 2-dependent and that these are the *unique* values of q for which the insertion process is finitely dependent [13, Proposition 13].

One may embellish these examples in our general setting. Replacing each of the 4 (resp. 3) colors with r copies of itself yields the complete multipartite graphs $K_{r,r,r,r}$ (resp. $K_{r,r,r}$). Both of these graphs are easily seen to have property (C) and to have a 1- (resp. 2-)dependent insertion process. We establish the remarkable fact that these are essentially the *only* (EC) graphs with a finitely dependent insertion process.

Theorem 1. *Let G be a finite graph satisfying property (C). Then the insertion process associated to G is k -dependent if and only if either: $G = K_{r,r,r}$ and $k \geq 2$, or $G = K_{r,r,r,r}$ and $k \geq 1$.*

If instead G satisfies property (EC), then the same characterization holds except G may in addition be a disjoint union of one of the above graphs with a collection of isolated vertices.

In the hope of uncovering other finitely dependent insertion processes, we consider weighted and directed graph variants. A **weighted digraph** on a vertex set V is simply a non-negative function on V^2 . For $i, j \in V$, we regard $w(i, j)$ as the weight of the directed edge from i to j , with 0 signifying that no edge is present. The weighted digraph is **undirected** if $w(i, j) = w(j, i)$ for all $i, j \in V$, it is **loopless** if $w(i, i) = 0$ for all $i \in V$, and it is **unweighed** if $w(i, j) \in \{0, 1\}$ for all $i, j \in V$. Given a weighted digraph with vertex set V , the following iterative algorithm computes a random element of $V \times V^2 \times \dots$, where the n^{th} coordinate is denoted by (X_1^n, \dots, X_n^n) . Let $X_0^n = X_{n+1}^n = \Delta$ denote an element not in V , and extend the definition of $w(i, j)$ by setting $w(\Delta, i) = w(i, \Delta) = 1$ for all $i \in V$. The obvious generalization of Holroyd and Liggett's construction in the context of weighted digraphs is given by the following algorithm.

Algorithm 1.

- (i) Let X_1^1 be a uniformly random element of V . Set $n = 1$.

(ii) If $\sum_{i=0}^n \sum_{v \in V} w(X_i^n, v) \cdot w(v, X_{i+1}^n) > 0$, let I and \mathcal{V} be random with

$$\mathbb{P}(I = i, \mathcal{V} = v \mid X_1^n, \dots, X_n^n) = \frac{w(X_i^n, v) \cdot w(v, X_{i+1}^n)}{\sum_{i,v} w(X_i^n, v) \cdot w(v, X_{i+1}^n)}, \quad i \in \{0, \dots, n\}, \quad \mathcal{V} \in V,$$

and proceed to (iii). Otherwise, halt.

(iii) Set $(X_1^{n+1}, \dots, X_{n+1}^{n+1}) = (X_1^n, \dots, X_I^n, \mathcal{V}, X_{I+1}^n, \dots, X_n^n)$, increment n , and go to (ii).

The weighted digraph is said to satisfy property (EC) if this algorithm does not halt and if the law of (X_1^n, \dots, X_n^n) is eventually consistent as $n \rightarrow \infty$. The projective limit of this sequence of laws is the law of the associated insertion process.

A weighted digraph is **uniform of weight w** if $w(i, j) \in \{0, w\}$ for all vertices i and j . When $w = 1$, this is simply an unweighted digraph. Such digraphs are obtained from undirected graphs by applying a weight of w to each edge. This is a strict generalization of the unweighted case, since Algorithm 1 weights endpoints differently than other locations (w instead of w^2). The endpoint-weighted insertion procedure in [11, §4] corresponds to a uniform weight digraph with $w \neq 1$.

Theorem 2. *Let G be an undirected, loopless weighted digraph of uniform weight w satisfying property (C). Then the insertion process associated to G is k -dependent if and only if either: $G = K_{r,r,r}$ and $k \geq 2$, or $G = K_{r,r,r,r}$ and $k \geq 1$.*

If instead G satisfies property (EC), then the same characterization holds except G may in addition be a disjoint union of one of the above graphs with a collection of isolated vertices.

Note that Theorem 1 is the special case $w = 1$ of Theorem 2.

We establish a similar result for directed, complete weighted digraphs. That is, we no longer assume that all edges have the same weight, but we do assume that all edges have positive weight. This is neither a special case nor a generalization of Theorem 2, although clearly this family of weighted digraphs is much richer than before.

Theorem 3. *Let G be a loopless weighted digraph satisfying property (C) such that $w(i, j) > 0$ for all distinct vertices i and j . Then the insertion process associated to G is k -dependent if and only if G is unweighted and either: $G = K_3$ and $k \geq 2$; or $G = K_4$ and $k \geq 1$.*

Another way to generalize graph coloring is to consider shifts of finite type [14, 19]. A **loopless shift of finite type** is a set of the form

$$S = S(q, W) = \left\{ x \in [q]^{\mathbb{Z}} : (x_{i+1}, \dots, x_{i+n}) \in W \quad \forall i \in \mathbb{Z} \right\},$$

for some $W \subset [q]^n \setminus \{(i, \dots, i) : i \in [q]\}$. The **de Bruijn graph** [8] associated to S has vertex set W and edge set

$$E = \left\{ ((x_1, \dots, x_n), (x_2, \dots, x_{n+1})) : (x_1, \dots, x_n), (x_2, \dots, x_{n+1}) \in W \right\}.$$

Shifts of finite type with $n = 2$ correspond to paths in digraphs. More precisely, if $S(q, W)$ is a shift of this form, then $S(q, W)$ is equal to the collection of all edge sets of bi-infinite paths in the digraph $G = ([q], W)$. In this case, the vertices of the de Bruijn graph are the edges of G , and the edges of the de Bruijn graph are pairs of edges in G whose respective tail and head agree. We consider insertion processes associated to de Bruijn graphs; in the case $n = 2$, this may be interpreted in terms of (directed) edge insertion, rather than vertex insertion.

Theorem 4. *Consider a loopless shift of finite type with de Bruijn graph G . Suppose that G has property (EC). The insertion process associated to G extends to a process supported on the shift of finite type, but this process is never finitely dependent.*

Open questions

In our quest to uncover new finitely dependent processes, we considered a class of insertion processes arising from graphs (or weighted and directed variants thereof) with vertex set V . Such an insertion process arose as a limit of a random tuple in V^n as $n \rightarrow \infty$. However, our method of extracting such a limit required delicate combinatorial identities to hold in the graph: the sequence of laws was required to be eventually consistent. Thus it is natural to ask whether similar results continue to hold after weakening this requirement further.

Question 1. *Does the conclusion of Theorem 1 continue to hold if we replace property (EC) with the assumption that sequence of laws converges weakly in the product topology?*

In Theorem 3, we do not even know if property (C) can be weakened to property (EC), let alone to weak limits.

Question 2. *Does the conclusion of Theorem 3 continue to hold for strongly connected weighted digraphs if we replace property (C) with property (EC)?*

Unweighted digraphs already pose challenges for our techniques.

Question 3. *Does the conclusion of Theorem 1 continue to hold for strongly connected unweighted digraphs?*

Recall that Theorem 3 applies to loopless weighted digraphs in which all edges have positive weight. Can this assumption be weakened?

Question 4. *Let G be a loopless weighted digraph satisfying property (C). Suppose that the insertion process associated to G is k -dependent for some integer k . Does it follow that G is unweighted and either $G = K_3$ for $k \geq 2$ or $G = K_4$ for $k \geq 1$?*

Overview

Section 2 presents a combinatorial analysis of insertion processes associated to weighted digraphs. Theorems 1 and 2 are proven in Section 3. Uniform weight graphs appear in Subsection 3.1, complete multipartite graphs in Subsection 3.2, and the proofs of Theorems 1 and 2 are in Subsection 3.3. Section 4 combines Theorem 2 with some additional arguments to deduce Theorem 3. Lastly, we deduce Theorem 4 in Section 5 from Lemma 8 in Subsection 2.2.

2 Weighted insertion

Let V be a finite alphabet. A **word** (of length n) is a finite sequence $x = (x_1, x_2, \dots, x_n) \in V^n$, which we sometimes abbreviate to $x_1x_2 \cdots x_n$. The word of length 0 is denoted by \emptyset . Let S_n be the symmetric group of all permutations of $1, \dots, n$. Let $x \in V^n$ be a word and let $\sigma \in S_n$ be a permutation. We interpret σ as meaning that at time $t = 1, \dots, n$ symbol $x_{\sigma(t)}$ arrives in (relative) position $\sigma(t)$.

Let $\text{Min}(\sigma) := \{1 \leq t \leq n : \sigma(t) = \min_{s \leq t} \sigma(s)\}$ denote the running minima of σ . Let

$$\text{Max}(\sigma) := \{1 \leq t \leq n : \sigma(t) = \max_{s \leq t} \sigma(s)\}$$

denote the set of running maxima. If $\sigma(t)$ is neither a running minimum nor maximum, then at time t the symbol $x_{\sigma(t)}$ is inserted between $x_{\sigma(t^-)}$ and $x_{\sigma(t^+)}$ where

$$\sigma(t^-) = \max_{s < t} \{\sigma(s) : \sigma(s) < \sigma(t)\} \text{ and } \sigma(t^+) = \min_{s < t} \{\sigma(s) : \sigma(s) > \sigma(t)\}.$$

A **weighted digraph** with vertex set V is a function $w : V^2 \rightarrow \mathbb{R}_{\geq 0}$. We define the **weight** of the pair $(x, \sigma) \in V^n \times S_n$ to be

$$w(x; \sigma) := \prod_{t \notin \text{Min}(\sigma)} w(x_{\sigma(t^-)}, x_{\sigma(t)}) \prod_{t \notin \text{Max}(\sigma)} w(x_{\sigma(t)}, x_{\sigma(t^+)}). \quad (1)$$

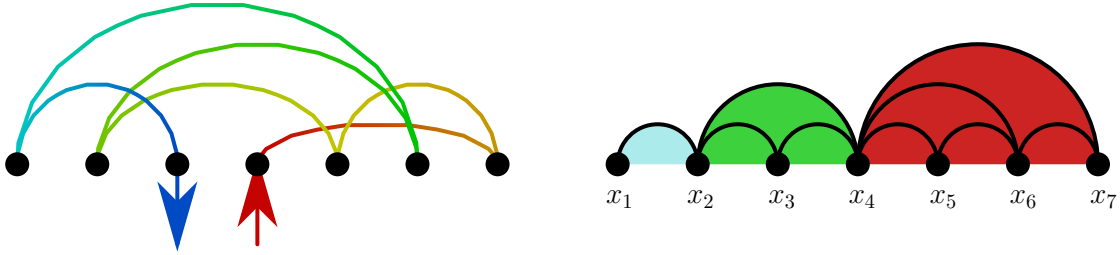


FIGURE 1. Left: The arrival order for $\sigma = 4752613$. Right: Multiply the edge weights to get $w(x_1 \cdots x_7; \sigma)$.

For example, when $\sigma = \text{id}$ is the identity permutation of $1, \dots, n$ we have $w(x; \text{id}) = w(x_1, x_2) \cdots w(x_{n-1}, x_n)$. We denote this quantity by $w(x)$. When the length of x is at most 1, we have that $w(x) = 1$.

If we imagine building the word dynamically using σ , then when $x_{\sigma(t)}$ is inserted between $x_{\sigma(t-)}$ and $x_{\sigma(t+)}$ a multiplicative weight of $w(x_{\sigma(t-)}, x_{\sigma(t)}) w(x_{\sigma(t)}, x_{\sigma(t+)})$ is incurred in (1).

Definition 1. Given a weighted digraph with vertex set V and a word $x \in V^n$, define

$$B(x) := \sum_{\sigma \in S_n} w(x; \sigma).$$

This is a generalization of the building number defined in [13], which is the special case consisting of the weighted digraph w with vertex set $\{1, \dots, q\}$ and weight function $w(i, j) = \mathbf{1}[i \neq j]$.

Clearly $B(x) > 0$ if and only if there exists $\sigma \in S_n$ such that $w(x; \sigma) > 0$. In this case $\sigma = \text{id}$ works. Word x has **positive weight** if either of the equivalent conditions $B(x) > 0$ or $w(x) > 0$ holds.

For a word $x = x_1 x_2 \cdots x_n$, we write $\widehat{x}_i = x_1 \cdots x_{i-1} x_{i+1} \cdots x_n$.

Lemma 5. Suppose that $x \in V^n$. Then $B(x)$ is equal to

$$B(\widehat{x}_1) w(x_1, x_2) + \sum_{i=2}^{n-1} w(x_{i-1}, x_i) B(\widehat{x}_i) w(x_i, x_{i+1}) + w(x_{n-1}, x_n) B(\widehat{x}_n).$$

Proof. Fix an index i . For any σ which satisfies $i = \sigma(n)$,

$$w(x; \sigma) = w(x_{i-1}, x_i) w(\widehat{x}_i; \widehat{\sigma}_i) w(x_i, x_{i+1}),$$

with the obvious modifications when $i = 1, n$. Summing over all $\sigma \in S_n$ yields the result. \square

2.1 Eventual Consistency

Let G be a weighted digraph with (finite) vertex set V . Say that G is **recurrent** if there exist arbitrarily long words of positive weight. We remark that this is equivalent to there existing positive-weight words of *every* length, since if $w(x_1x_2 \cdots x_n) > 0$ then also $w(x_1x_2 \cdots x_k) > 0$ for all $1 \leq k \leq n$. Furthermore, it is easy to see that G is recurrent if and only if

$$\sum_{y \in V^n} B(y) > 0 \quad \text{for all } n \geq 0. \quad (2)$$

If G is a recurrent digraph, then for all $n \geq 0$ we define a probability mass function P_n on V^n by setting

$$P_n(x) = \frac{B(x)}{\sum_{y \in V^n} B(y)}, \quad x \in V^n. \quad (3)$$

It is a simple consequence of Lemma 5 that the random element of V^n computed by Algorithm 1 (from the introduction) has mass function P_n . Moreover, Algorithm 1 halts if and only if G is not recurrent.

If G is recurrent and the mass functions $\{P_n\}_{n=0}^\infty$ satisfy

$$P_n(x) = \sum_{v \in V} P_{n+1}(xv) = \sum_{v \in V} P_{n+1}(vx) \quad \text{for all } n \geq 0, \quad (4)$$

we say that G satisfies **property (C)** and that the mass functions are **consistent**. Likewise if G is recurrent and (4) holds for all sufficiently large n , we say that G satisfies **property (EC)** and that the mass functions are **eventually consistent**.

Clearly, equation (4) is equivalent to the existence of constants $C_n > 0$ such that (5) holds for all $x \in V^n$:

$$\sum_{v \in V} B(xv) = \sum_{v \in V} B(vx) = C_n B(x). \quad (5)$$

Thus, a recurrent graph satisfies property (C) if and only if there exists $C_n > 0$ such that (5) holds for all $n \geq 0$. Similarly, a recurrent graph satisfies property (EC) if and only if for all sufficiently large n , there exists $C_n > 0$ such that (5) holds.

When the mass functions $\{P_n\}$ are eventually consistent, it follows from the Kolmogorov extension theorem [10] that there is a unique process $(X_i)_{i \in \mathbb{Z}}$ satisfying

$$\mathbb{P}((X_{i+1}, \dots, X_{i+n}) = x) = P_n(x)$$

for all sufficiently large n . This is the **insertion process** associated to an (EC) weighted digraph. It follows by construction that the insertion process is stationary.

Lemma 6. *Suppose that G has property (EC). Then the associated insertion process is k -dependent if and only if for all $n, m \in \mathbb{N}$ sufficiently large, there are positive constants $C_{n,m} > 0$ for which*

$$\sum_{W \in V^k} B(x W y) = C_{n,m} B(x) B(y), \quad x \in V^n \text{ and } y \in V^m. \quad (6)$$

Proof. By normalizing (6) we see it is equivalent to

$$\sum_{W \in V^k} P_{n+k+m}(x W y) = P_n(x) P_m(y).$$

This is equivalent to the associated insertion process $(X_n)_{n \in \mathbb{Z}}$ having the property that $(X_i)_{i \in I}$ and $(X_j)_{j \in J}$ are independent for any pair of sufficiently large intervals I and J separated by a distance of k or greater. Since restriction preserves independence, we may remove the adjective ‘sufficiently large’ from the previous sentence. \square

Holroyd and Liggett [13] proved that the (unweighted, undirected) complete graph K_q has property (C) for all $q \geq 2$. Furthermore, they showed that the associated insertion process is k -dependent if and only if $q = 3$ and $k \geq 2$ or $q = 4$ and $k \geq 1$. In Section 3 we show that apart from minor modifications, these are the only unweighted and undirected graphs with property (EC) for which the associated insertion process is finitely dependent.

2.2 Necessary conditions for k -dependence

Consider a weighted digraph G satisfying property (EC). In this subsection, we present necessary conditions for the associated insertion process to be finitely dependent. We show that such digraphs must contain directed triangles (Lemma 8) and are not too far from being strongly connected (Lemma 9).

A **directed triangle on** (a, b) is a triple of vertices $(a, b, c) \in V^3$ satisfying

$$w(a, b) w(b, c) w(a, c) > 0.$$

In the next lemma, we consider weighted digraphs that lack directed triangles. Equivalently, all insertions in Algorithm 1 must occur at the endpoints.

Lemma 7. *For a weighted digraph lacking directed triangles, any $x \in V^n$ satisfies $B(x) = 2^{n-1} w(x)$.*

Proof. For any $1 < i < n$, since $w(x_{i-1}, x_i) > 0$ and $w(x_i, x_{i+1}) > 0$ we must have that $w(x_{i-1}, x_{i+1}) = 0$. Applying Lemma 5 yields that

$$B(x) = w(x_1, x_2) \cdot B(\widehat{x}_1) + w(x_{\ell-1}, x_\ell) \cdot B(\widehat{x}_\ell),$$

which implies the claim by induction on the length of the word. \square

Lemma 8. *If a weighted digraph satisfies property (EC) and the associated insertion process is finitely dependent, then it contains a directed triangle.*

Proof. Suppose there was a weighted digraph G lacking directed triangles that satisfies property (EC) and whose associated insertion process is k -dependent for some $k \geq 0$. Let V denote its vertex set. Applying Lemma 6, it follows that for all sufficiently large m and n and for all vertices $i, j \in V$ we have that

$$\sum_{P \in V^m, W \in V^k, Q \in V^n} B(iPWQj) = C_{m,n} \sum_{P \in V^m} B(iP) \sum_{Q \in V^n} B(Qj). \quad (7)$$

Let A denote the adjacency matrix of G , which is the $V \times V$ matrix with $A_{ij} = w(i, j)$. Using Lemma 7 to express the left hand side of (7) in terms of the adjacency matrix we obtain that

$$(A^{m+k+n+1})_{ij} = 2^{-m-k-n+1} \sum_{P \in V^m, W \in V^k, Q \in V^n} B(iPWQj). \quad (8)$$

Combining (7) with (8) implies that $\text{rank}(A^N) \leq 1$ for $N = m + k + n + 1$. Let $\lambda_1, \dots, \lambda_{|V|}$ denote the multiset of (possibly complex) eigenvalues of A , listed according to algebraic multiplicity. Then $\lambda_1^N, \dots, \lambda_{|V|}^N$ comprise the multiset of eigenvalues of A^N . But since $\text{rank}(A^N) \leq 1$, there can be at most 1 nonzero eigenvalue.

In fact, all eigenvalues must vanish. To see why, consider $\text{Tr } A = \sum_{i \in V} w(i, i)$. Since (i, i, i) is a directed triangle, our hypothesis that G lacks directed triangles implies that $w(i, i) = 0$ and thus $\text{Tr } A = 0$. Combined with the previous paragraph, this implies that all eigenvalues vanish and thus A is nilpotent.

From this it follows that all but finitely many words in $\bigcup_{\ell \geq 0} V^\ell$ have weight zero. Thus G is not recurrent, contradicting our assumption that G satisfies property (EC). \square

A weighted digraph G is **strongly connected** if for every pair of vertices $i, j \in V$, there is a positive-weight word $x \in \bigcup_{\ell \geq 1} V^\ell$ that begins with i and ends with j . We denote the existence of such an x by writing $i \rightarrow j$. The **strongly connected components** (SCCs) of a weighted digraph are the equivalence classes of the relation

$$\{(i, j) \in V^2 : i \rightarrow j \text{ and } j \rightarrow i\}.$$

By convention $i \rightarrow i$ since singleton words are defined to have weight 1 (see the remark prior to Definition 1). Also, recall from our discussion following Definition 1 that a word has positive weight if and only if $w(x) > 0$, which occurs if and only if $B(x) > 0$.

Recall that a weighted digraph is a function $w: V^2 \rightarrow \mathbb{R}_{\geq 0}$. A subset $U \subseteq V$ induces a subdigraph by restriction of w to U^2 . Say that an SCC is **recurrent** if the induced subdigraph of G is recurrent. By our remarks at the beginning of Subsection 2.1, an SCC with vertex set \mathcal{C} is recurrent if and only if there are arbitrarily large n such that there exists a positive-weight word $x \in \mathcal{C}^n$.

Note that any SCC with at least two vertices is recurrent, for if a and b are any two such vertices, then each of the words $(ab)^n \in \mathcal{C}^{2n}$ have positive weight. Also observe that for a directed acyclic graph, the SCCs are singletons, none of which is recurrent. Moreover, a singleton SCC is recurrent if and only if its vertex has a self-loop. In general, each directed cycle is contained in a recurrent SCC.

For the remainder of this section, we restrict attention to weighted digraphs satisfying property (EC) and whose associated insertion process is finitely dependent. We will show in Lemma 9 that such digraphs are not too far from being strongly connected.

Lemma 9. *Let G be a weighted digraph with property (EC) whose associated insertion process is finitely dependent.*

- (i) *G has a unique recurrent SCC.*
- (ii) *Vertices not in the recurrent SCC belong to singleton SCCs that have no directed path to or from the recurrent SCC.*
- (iii) *The subdigraph of G induced by the recurrent SCC also has property (EC). Moreover, its associated insertion process coincides with that of G .*
- (iv) *If moreover G has property (C), then it is strongly connected.*

In particular, it follows from Lemma 9 that the family of finitely dependent insertion processes associated to weighted digraphs is unchanged if one restricts attention to strongly connected weighted digraphs.

Proof. Let $(X_i)_{i \in \mathbb{Z}}$ denote the insertion process on G . Suppose without loss of generality that it is $(k+1)$ -dependent, for some $k \geq 0$. In particular, we have the i.i.d. subsequence $(X_{ik})_{i \in \mathbb{Z}}$.

Let $\mathcal{C}_1, \dots, \mathcal{C}_\ell$ denote the strongly connected components of G . The weighted digraph G induces the weighted digraph G_{SCC} on the set of SCCs, denoted $V_{\text{SCC}} := \{\mathcal{C}_1, \dots, \mathcal{C}_\ell\}$, given by the weight

$$w_{\text{SCC}}(\mathcal{C}_i, \mathcal{C}_j) := \sum_{u \in \mathcal{C}_i, v \in \mathcal{C}_j} w(u, v).$$

Observe that $\mathcal{C}_i \rightarrow \mathcal{C}_j$ in G_{SCC} if and only if for some $u \in \mathcal{C}_i$ and $v \in \mathcal{C}_j$, there is a positive-weight word x composed of vertices of G that begins with u and ends with v . Thus if $\mathcal{C}_i \rightarrow \mathcal{C}_j$ and $\mathcal{C}_j \rightarrow \mathcal{C}_i$, then $\mathcal{C}_i \cup \mathcal{C}_j$ is also an equivalence class, whereupon $\mathcal{C}_i = \mathcal{C}_j$.

Consider the function $f: V \rightarrow V_{\text{SCC}}$ that assigns to each vertex its strongly connected component. Then the sequence $(f(X_i))_{i \in \mathbb{Z}}$ has the property that for all $i < j$, we have $f(X_i) \rightarrow f(X_j)$ in G_{SCC} almost surely. But if $f(X_i)$ occurs once in the sequence, it a.s. occurs infinitely many times (by passing to an i.i.d. subsequence). Thus there exists $\ell > j$ such that $f(X_\ell) = f(X_i)$. Consequently the sequence

$$(f(X_i))_{i \in \mathbb{Z}} = (\dots, \mathcal{C}_{\text{rec}}, \mathcal{C}_{\text{rec}}, \dots) \quad (9)$$

is a.s. constant. Furthermore, the unique value \mathcal{C}_{rec} that it takes must be a recurrent SCC, since the marginals of the process $(X_i)_{i \in \mathbb{Z}}$ then yield arbitrarily long positive-weight words in $\bigcup_{\ell \geq 1} \mathcal{C}_{\text{rec}}^\ell$.

On the other hand, each vertex $v \in V$ belonging to a recurrent SCC occurs infinitely often in the sequence $(X_i)_{i \in \mathbb{Z}}$. Indeed, we can fix a sufficiently long positive-weight word containing v and observe that it occurs with positive density in $(X_i)_{i \in \mathbb{Z}}$ by finite dependence. Thus by the previous paragraph, it follows that \mathcal{C}_{rec} is the a unique recurrent SCC of G , establishing property (i).

Property (ii) follows by combining our previous observation that an SCC containing distinct vertices is recurrent with the following argument. If there was a directed path joining \mathcal{C}_{rec} to another strongly connected component denoted \mathcal{C}' , then there would be arbitrary long positive-weight words that contain vertices in \mathcal{C}' . But then \mathcal{C}' appears in the sequence $(f(X_i))_{i \in \mathbb{Z}}$, by consideration of a sufficiently long marginal of the insertion process. This contradicts (9), proving (ii).

For all $n > |V|$, every word $x \in V^n$ contains some vertex at least twice. Thus by (ii), it follows that if such a word satisfies $B(x) > 0$, then necessarily $x \in \mathcal{C}_{\text{rec}}^n$. Hence

$$\sum_{x \in V^n} B(x) = \sum_{x \in \mathcal{C}_{\text{rec}}^n} B(x),$$

and it follows that all sufficiently long marginals of the insertion processes associated to G and the subdigraph induced by \mathcal{C}_{rec} coincide. This yields (iii).

Finally, (iv) follows from the observation that $B(x) = 1$ for all words of length 1. Thus if G satisfies property (C), the random variable X_0 in the associated insertion process is uniformly random on the vertex set V . But by (9) we have that $X_0 \in \mathcal{C}_{\text{rec}}$ almost surely, implying that $V = \mathcal{C}_{\text{rec}}$ and therefore G is strongly connected. \square

3 Uniform weight graphs

We establish Theorem 2 in this section, and deduce Theorem 1 as a corollary. The plan of attack is as follows:

- (i) Subsection 3.1 defines the special class of weighted digraphs to which Theorem 2 applies, which we call uniform weight graphs. We deduce necessary structural properties for a uniform weight graph to satisfy property (EC).
- (ii) Subsection 3.2 is devoted to complete multipartite graphs of uniform weight. It is shown that understanding finite dependence for such graphs reduces to the analysis of complete graphs, which has already been undertaken in [13].
- (iii) Subsection 3.3 proves Theorem 2 by using a combinatorial argument to show that the structural properties in Subsection 3.1 imply that the graph is complete multipartite, then applying the results of Subsection 3.2.

3.1 Uniform weights

Recall that a weighted digraph is of uniform weight w if $w(i, j) \in \{0, w\}$ for all vertices i and j . In this section we consider only loopless and undirected weighted digraphs of uniform weight, or for short, **uniform weight graphs**.

Since uniform weight graphs are undirected, we refer to their strongly connected components as connected components (when applying Lemma 9, for instance). For uniform weight graphs, we say that vertices i and j are **adjacent** if and only if $w(i, j) > 0$, in which case $w(i, j) = w$.

Applying (5) to a positive-weight word in a uniform weight graph yields that

$$B(x) = wB(\hat{x}_1) + w^2 \sum_{i=2}^{n-1} B(\hat{x}_i) + wB(\hat{x}_n), \quad x \in V^n. \quad (10)$$

For any positive integer n , we denote by $(ab)^{n/2}$ the unique alternating word in $\{a, b\}^n$ beginning with a (in particular, note that n may be odd).

Lemma 10. *If $w(a, b) > 0$ and $w(b, v) > 0$, then*

$$B\left((ab)^{n/2} v\right) = \begin{cases} (2w)^n, & w(a, v) = 0 \\ [2w^2(w + w^2)^{n-1} - (2w)^n]/(w - 1), & w(a, v) = w \text{ and } w \neq 1 \\ 2^{n-1}(n + 1), & w(a, v) = w \text{ and } w = 1. \end{cases} \quad (11)$$

Proof. First suppose that $w(a, v) = 0$. By (10) we have that

$$B\left((ab)^{n/2} v\right) = w \cdot B\left((ba)^{(n-1)/2} v\right) + w \cdot B\left((ab)^{n/2}\right).$$

Applying this inductively yields the first case of (11).

Next, suppose that $w(a, v) = w$ and $w \neq 1$. Again by (10),

$$\begin{aligned} B\left((ab)^{n/2} v\right) &= w \cdot B\left((ba)^{(n-1)/2} v\right) + w^2 \cdot B\left((ab)^{(n-1)/2} v\right) + w \cdot B\left((ab)^{n/2}\right) \\ &= (w + w^2) \cdot B\left((ba)^{(n-1)/2} v\right) + w \cdot (2w)^{n-1}. \end{aligned}$$

Now the second case of (11) follows by induction, starting from the base case $B(av) = 2w$. The final case of (11) follows from a simplification of the previous calculation. \square

Lemma 11. *If G is a non-empty connected uniform weight graph satisfying property (EC), then the following conditions hold.*

- (i) *The graph G is d -regular for some $d \geq 1$.*
- (ii) *For some $t \geq 0$, there are t triangles on every edge of G .*

Proof. Let w be the common weight of the edges in G . Fix a vertex b . Let d denote the degree of b . By hypothesis, there exists an edge (a, b) . Let t denote the number of triangles on the edge (a, b) . First suppose that $w \neq 1$. By Lemma 10,

$$\sum_v B\left((ab)^{n/2} v\right) = (2w)^n(d - t) + \left[\frac{2w^2(w + w^2)^{n-1} - (2w)^n}{w - 1} \right] t.$$

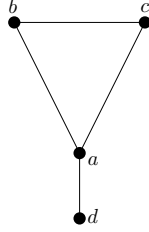
Recall that $B((ab)^{n/2}) = (2w)^{n-1}$. Thus by (5), it follows that

$$C_n = 2w(d - t) + \left[\frac{w\left(\frac{w+1}{2}\right)^{n-1} - 1}{w - 1} \right] (2wt) \tag{12}$$

for all sufficiently large n if $w \neq 1$. Likewise if $w = 1$, for all sufficiently large n we have that $C_n = 2(d - t) + (n + 1)t$. In either case, both d and t are determined by the values of C_n for n sufficiently large. Thus d and t take the same value, for all vertices b and all edges (a, b) . \square

Consider the graph that appears in Figure 2.

Definition 2. An undirected graph on the vertices a, b, c, d is a **kite** if (a, b, c) form a triangle and d is adjacent to a and to no other vertex.

FIGURE 2. The kite $(abc; d)$

Lemma 12. *Suppose that G has uniform weight and satisfies property (EC). Then there is no kite which occurs as an induced subgraph of G .*

Proof. Suppose to the contrary that G contains a kite $(abc; d)$. For all $n \geq 3$, we will construct words x and y satisfying

$$\sum_{v \in V} \frac{B(xv)}{B(x)} > \sum_{v \in V} \frac{B(yv)}{B(y)},$$

from which it will follow that G cannot satisfy property (EC).

Let $x = (ba)^{n/2}$ and let $y = d(ab)^{(n-1)/2}$. Simple casework reveals that for each $\sigma \in S_n$ we have $w(x; \sigma) = w(y; \sigma)$, and consequently $B(x) = B(y)$. Note that this common value is positive (by Lemma 10, for instance). Thus we need only establish that

$$\sum_v B(xv) > \sum_v B(yv).$$

First observe that we have the weaker inequality

$$\sum_v B(xv) \geq \sum_v B(yv). \quad (13)$$

Indeed, each $\sigma \in S_n + 1$ with $w(yv; \sigma) > 0$ is seen to satisfy $w(xv; \sigma) > 0$ as well. Since the graph has uniform weight, it therefore follows that $w(xv; \sigma) \geq w(yv; \sigma)$ and the claim follows upon summation.

To obtain strict inequality, observe that if $\sigma \in S_{n+1}$ satisfies $\sigma(\{1, 2\}) = \{1, n+1\}$ then $w(xc; \sigma) > w(yc; \sigma) = 0$. When combined with the previous inequality, it follows that $B(xc) > B(yc)$. \square

3.2 Complete multipartite graphs

In this section we relate property (EC) for complete multipartite graphs to property (EC) for complete graphs. This allows us to extend results of Holroyd and Liggett [13] to handle complete multipartite graphs.

We use the following notation: K_q denotes the complete graph on q vertices, $K_{r,\dots,r}$ denotes the complete multipartite graph with q parts each of size r , and $w \cdot K_{r,\dots,r}$ denotes the uniform weight graph in which each edge of the corresponding complete multipartite graph has weight $w > 0$. In all cases, we take the vertex set to be $[qr] := \{1, \dots, qr\}$ (with $r = 1$ in the case of K_q), and we take the edge set to be $\{(i, j) : i \not\equiv j \pmod{q}\}$. In the case of $w \cdot K_{r,\dots,r}$, the weight function is given by $w(i, j) = w \cdot \mathbf{1}[i \not\equiv j \pmod{q}]$.

Note that the graph $K_{r,\dots,r}$ is a Turán graph [4].

Lemma 13. *For any $r \geq 1$ and $k \geq 0$, the graph $w \cdot K_{r,\dots,r}$ satisfies property (EC) and its associated insertion process is k -dependent if and only if the same holds for $w \cdot K_q$.*

Proof. Consider the mapping $f: [qr] \rightarrow [q]$ given by $f(i) = i \pmod{q}$. For any word $x \in [qr]^n$, let $f(x)$ denote the word $f(x_1)f(x_2) \cdots f(x_n) \in [q]^n$. Observe that

$$w(i, j) = w(f(i), f(j)) \quad \forall i, j \in [qr],$$

from which it follows that $B(x) = B(f(x))$. Thus the following are equivalent:

- (i) $\sum_{v \in [qr]} B(xv) = rC_n B(x)$
- (ii) $\sum_{v' \in [q]} B(f(x)v') = C_n B(f(x))$

By (5), it follows that $w \cdot K_q$ satisfies property (EC) if and only if $w \cdot K_{r,\dots,r}$ satisfies property (EC).

Similarly, the following are equivalent:

- (i) $\sum_{W \in [qr]^k} B(x W y) = r^k C_{nm} B(x) B(y)$
- (ii) $\sum_{W' \in [q]^k} B(f(x) W' f(y)) = C_{nm} B(f(x)) B(f(y))$

The result now follows by Lemma 6. □

Lemma 14. *Suppose that $w \neq 1$ and $q \geq 3$. Then the graph $w \cdot K_q$ does not satisfy property (EC).*

Proof. Consider an integer $n \geq 3$. Fix distinct vertices $a, b, c \in K_q$. Keeping in mind our notation for alternating words, set

$$x = (ba)^{n/2} \in \{a, b\}^n, \quad y = c (ab)^{(n-1)/2} \in \{a, b, c\}^n$$

(regardless of the parity of n). Consider the quantities

$$Q_n = \sum_{v \in K_q} \frac{B(xv)}{B(x)}, \quad R_n = \sum_v \frac{B(yv)}{B(y)}.$$

We will prove by induction on n that $Q_n > R_n$ for all $w > 1$ and $Q_n < R_n$ for all $w \in (0, 1)$.

From (10) we deduce the recurrences $B(xv) = wB(\hat{x}_1v) + w^2B(\hat{x}_nv) + wB(x)$ and $B(yv) = wB(\hat{y}_1v) + w^2B(\hat{y}_2v) + w^2B(\hat{y}_nv) + wB(y)$, yielding that

$$\begin{aligned} Q_n &= \left(\frac{w+1}{2} \right) Q_{n-1} + qw - w(w+1) \\ R_n &= \frac{wB(\hat{x}_1)}{B(y)} Q_{n-1} + \frac{2w^2B(\hat{y}_2)}{B(y)} R_{n-1} + qw - w(w+1). \end{aligned}$$

Subtracting the previous equations and substituting the recurrence $B(y) = wB(\hat{y}_1) + w^2B(\hat{y}_2) + wB(\hat{y}_n)$ yields that $2B(y)(Q_n - R_n)/w$ equals

$$[(w-1)B(\hat{x}_1) + (w+1)^2B(\hat{y}_2)] Q_{n-1} - 4wB(\hat{y}_2)R_{n-1} \quad (14)$$

When $w > 1$, we leave out the first term, using $(w+1)^2 \geq 4w$ to obtain

$$2B(y)(Q_n - R_n)/w > 4w(Q_{n-1} - R_{n-1}).$$

Since the right side vanishes for $n = 3$, it follows that $Q_n > R_n$ by induction.

Next suppose that $w < 1$. This case requires a tighter bound. We begin by establishing that for all $m \geq 3$,

$$B\left((ba)^{m/2}\right) > (1-w)B\left(c(ba)^{(m-1)/2}\right).$$

Indeed, we deduce the bound inductively from

$$\begin{aligned} B\left((ba)^{m/2}\right) &= 2wB\left((ab)^{(m-1)/2}\right) \\ &= [(1-w)w + w^2 + w]B\left((ab)^{(m-1)/2}\right) \\ \text{(induction)} \quad &> (1-w) \left[wB\left((ab)^{(m-1)/2}\right) + w^2B\left(c(ba)^{(m-2)/2}\right) \right. \\ &\quad \left. + wB\left(c(ba)^{(m-2)/2}\right) \right] \\ &= (1-w)B\left(c(ab)^{(m-1)/2}\right). \end{aligned}$$

Consequently

$$(1-w)B(\hat{x}_1) > (1-w)^2B(\hat{y}_2) \quad (15)$$

Now we plug the bound (15) into (14) to obtain a bound which is suitable for induction:

$$\begin{aligned} 2B(y)(R_n - Q_n)/w &= 4wB(\hat{y}_2)R_{n-1} \\ &\quad + [(1-w)B(\hat{x}_1) - (w+1)^2B(\hat{y}_2)]Q_{n-1} \\ \text{using (15)} \quad &> 4wB(\hat{y}_2)(R_{n-1} - Q_{n-1}). \end{aligned}$$

As before, when $n = 3$ the right side vanishes. Thus it follows by induction that if $0 < w < 1$, we have $R_n > Q_n$ for all $n \geq 3$. Combined with the previous case, we conclude that when $w \neq 1$ and $q \geq 3$, the graph $w \cdot K_q$ does not satisfy property (EC). \square

Combining several results allows us to determine which uniform weight complete multipartite graphs satisfy property (EC) and have a finitely dependent associated insertion process.

Lemma 15. *The graph $w \cdot K_{r,\dots,r}$ for $q \geq 3$ and $r \geq 1$ has property (EC) and has a k -dependent associated insertion process if and only if $w = 1$ and either: $q = 3$ and $k \geq 2$; or $q = 4$ and $k \geq 1$.*

Proof. By Lemma 13, it suffices to consider $w \cdot K_q$. By Lemma 14, $w = 1$. Applying Propositions 10 and 13 of [13], the insertion process on K_q is k -dependent if and only if $q = 3$ ($k \geq 2$) or $q = 4$ ($k \geq 1$). \square

3.3 Extension to uniform weight graphs

We now combine the results of Subsections 3.1 and 3.2 to extend the conclusion of Lemma 15 to all uniform weight graphs. We establish that the only uniform weight graphs satisfying property (EC) that have a k -dependent associated insertion process are $K_{r,r,r}$ and $K_{r,r,r,r}$, and unions thereof with isolated vertices. Note that these graphs have weight $w = 1$, even though we allow $w > 0$ to be arbitrary a priori. These graphs are closely related to the graphs K_3 and K_4 corresponding to the colorings discovered by Holroyd and Liggett; in fact, the graphs $K_{r,r,r}$ and $K_{r,r,r,r}$ are obtained from K_3 (resp. K_4) by replacing each vertex with r copies of itself. Similarly, the insertion process on $K_{r,r,r}$ (resp. $K_{r,r,r,r}$) is obtained from the corresponding process on K_3 (resp. K_4) by replacing each instance of vertex i with an i.i.d. choice of one of its r copies in $K_{r,r,r}$ (resp. $K_{r,r,r,r}$).

The following graph-theoretic lemma allows us to reduce the general uniform weight case to that of the complete multipartite graphs treated in Subsection 3.2. We will use Lemma 16 to show that any uniform weight graph either contains a kite, or is complete multipartite.

Lemma 16. *Let G be a kite-free connected loopless graph containing a triangle abc . Then every vertex $d \in V$ is adjacent to at least two of $\{a, b, c\}$.*

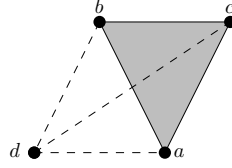


FIGURE 3. The problematic triangle abc in Lemma 16

Proof. By Lemma 12, no vertex $d \in V$ is adjacent to exactly one of abc . Hence it suffices to show that every vertex d is adjacent to abc .

Suppose to the contrary that some $d \in V$ is non-adjacent to abc (Figure 3). Choose a minimal path joining d to abc . We show that a kite is present near the intersection of the path with abc , from which we obtain the desired contradiction.

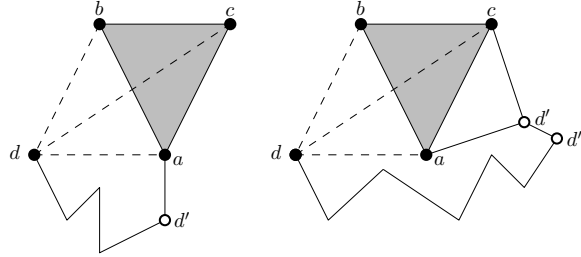


FIGURE 4. Reaching a contradiction

Let d' denote the path vertex adjacent to abc . There are two cases to consider: either d' is adjacent to a single vertex of abc (left half of Figure 4), or it is adjacent to more than one vertex (right half of Figure 4).

If d' is adjacent to a single vertex, then $(abc; d')$ is a kite. Now suppose that d' is adjacent to more than one vertex. Without loss of generality, suppose that d' is adjacent to a and c . Let d'' denote a neighbor of d' on the path from d to d' . By minimality of the path, d'' is non-adjacent to abc . Consequently $(d'ac; d'')$ is a kite. \square

We are now in a position to prove Theorem 2.

Proof of Theorem 2. Suppose that G is a uniform weight graph with property (EC) and suppose that the insertion process associated to G is k -dependent. We will deduce that either $G = K_{r,r,r}$ and $k \geq 2$, or $G = K_{r,r,r}$ and $k \geq 1$, or G is a disjoint union of one of these graphs with a collection of isolated vertices.

Since uniform weight graphs are undirected, Lemma 9 takes on a simpler form in the present context. Indeed, it implies that both the (EC) property and the insertion process are unchanged by deletion of isolated vertices, and moreover it implies that the resulting graph is connected. Hence it suffices to consider the connected case.

We show in this case that G is complete multipartite. Consider the relation

$$\{(i, j) \in V^2: w(i, j) = 0\}. \quad (16)$$

This relation is reflexive and symmetric by definition of a uniform weight graph. Once we establish transitivity, it will follow that the graph G is complete multipartite with partite sets given by the equivalence classes of this relation.

Suppose to the contrary that transitivity did not hold. Then there would exist vertices a, b, d such that $w(a, b) > 0$ yet $w(a, d) = w(b, d) = 0$. By Lemma 8 and Lemma 11, there are $t \geq 1$ triangles on every edge. In particular, we may complete the edge (a, b) into a triangle abc . By Lemma 12, G lacks kites, and by assumption it is connected. Moreover by definition of uniform weight graphs, G is loopless.

Thus we have verified the conditions of Lemma 16. Consequently, the vertex d is adjacent to at least two of a, b, c . This contradicts the hypothesis that $w(a, d) = w(a, b) = 0$. Thus the relation (16) is transitive, and we deduce that it is an equivalence relation.

Decompose the vertex set into equivalence classes of (16). Then G is complete multipartite, with partite sets are given by the equivalence classes of (16). By Lemma 11, G is a regular graph and thus the parts have equal sizes. Thus $G = w \cdot K_{r, \dots, r}$, for some $w > 0$ and $r \geq 1$.

Applying Lemma 8 again, we see that G contains a triangle and therefore $q \geq 3$. The result now follows when G has property (EC) by Lemma 15.

Finally, observe that if G also satisfies property (C), then there can be no isolated vertices by Lemma 9(iv). \square

4 Complete weighted digraphs with property (C)

The results in this section apply to loopless complete weighted digraphs satisfying property (C). That is, the digraphs under consideration satisfy $w(i, i) = 0$ and $w(i, j) > 0$ for all distinct vertices i and j , as well as property (C). We will establish Theorem 3, which states that the only such graphs for which the associated insertion process is k -dependent are the (unweighted) graphs K_3 (for $k \geq 1$) and K_4 (for $k \geq 2$). We will establish this result by reducing it to the case of a uniform weight graph and applying results from Section 3.

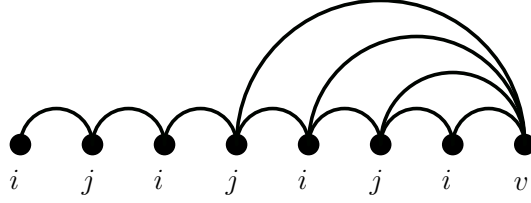


FIGURE 5. Let $x \in \{i, j\}^n$ be the unique alternating word ending in i . For $\sigma \in S_{n+1}$, let ℓ be as in (19). Consider the nearest neighbor graph on $\{1, \dots, n+1\}$, except we modify vertex $n+1$ to be adjacent to its nearest ℓ neighbors. This graph is drawn above with vertex i labeled with the i^{th} symbol of the word xv . Then $w(xv; \sigma) = \prod_{e=(i,j)} w((xv)_i, (xv)_j)$, where the product is taken over the edges of the graph we have just described.

For distinct indices $i, j \in V$, we define the quantity

$$T_n(i, j) = \sum_{v \in V} w(i, v)^{\lceil n/2 \rceil} w(j, v)^{\lfloor n/2 \rfloor},$$

where we use the convention $0^0 := 1$.

Lemma 17. *Fix a loopless complete weighted digraph satisfying property (C) and fix an integer $n \geq 1$. Then the value of $T_n(i, j)$ is constant over all vertices i and j with $w(i, j) > 0$.*

Proof. Let $x \in \{i, j\}^n$ be the unique alternating word ending in i . By Definition 1,

$$B(xv) = \sum_{\sigma \in S_n} w(xv; \sigma). \quad (17)$$

Since $w(i, i) = w(j, j) = 0$ by assumption, for all permutations with $w(xv; \sigma) > 0$ we have

$$w(xv; \sigma) = w(x)w(i, v)^{\lceil \ell/2 \rceil} w(j, v)^{\lfloor \ell/2 \rfloor}, \quad (18)$$

where ℓ is given by the formula

$$\ell = n + 1 - \max_{t < \sigma^{-1}(n+1)} \sigma(t). \quad (19)$$

This is straightforward to verify from the definitions, as indicated in Figure 5.

By (19), we see that ℓ ranges over $\{1, \dots, n\}$ as σ ranges over S_{n+1} . Thus substituting (18) into (17) implies that there are integers $d_1, \dots, d_n > 0$ whose values depend only on n such that

$$B(xv) = w(x) \sum_{\ell=1}^n d_\ell w(i, v)^{\lceil \ell/2 \rceil} w(j, v)^{\lfloor \ell/2 \rfloor}.$$

Summing over all vertices $v \in V$, we obtain

$$\sum_{v \in V} B(xv) = w(x) \sum_{\ell=1}^n d_\ell T_\ell(i, j).$$

By (5) we have that $\sum_{v \in V} B(xv) = C_n B(x)$. Furthermore $B(x) = 2^{n-1}w(x)$, by Lemma 7 applied to the subgraph induced by $\{i, j\}$. As $w(x) > 0$, we have that

$$\sum_{\ell=1}^n d_\ell T_\ell(i, j) = 2^{n-1}C_n.$$

Since $d_1, \dots, d_n > 0$, we may explicitly solve this system of equations to obtain that

$$\begin{pmatrix} T_1(i, j) \\ \vdots \\ T_n(i, j) \end{pmatrix} = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ d_1 & d_2 & \cdots & 0 \\ d_1 & d_2 & \cdots & 0 \\ d_1 & d_2 & \cdots & d_n \end{pmatrix}^{-1} \begin{pmatrix} C_1 \\ 2C_2 \\ \vdots \\ 2^{n-1}C_n \end{pmatrix}.$$

In particular, it follows that $T_n(i, j)$ is independent of the pair (i, j) . \square

In light of Lemma 17, we write T_n in place of $T_n(i, j)$ from now on.

Lemma 18. *For all pairs of distinct indices (i, j) and (i', j') , we have $w(i, j) = w(i', j')$.*

Proof. Let $z_1 > z_2 > \cdots$ denote the set of distinct positive values attained by $w(i, v)w(j, v)$ as v ranges over V . Let V_ℓ denote the vertices which contribute to z_ℓ , given by

$$V_\ell = \{v \in V : w(i, v)w(j, v) = z_\ell\},$$

and let $a_\ell = \sum_{v \in V_\ell} w(i, v)$. Rewriting the expression for T_{2n+1} yields

$$T_{2n+1} = \sum_{\ell} a_\ell z_\ell^n.$$

Next, observe that $z_1 = \inf_{n \rightarrow \infty} (T_{2n+1})^{1/n}$ and $a_1 = \inf_{n \rightarrow \infty} \frac{T_{2n+1}}{z_1^n}$. Hence the parameters a_1 and z_1 can be reconstructed given the sequence $\{T_n\}$. Applying the same procedure to $T_{2n+1} - a_1 z_1^n$ allows us to iteratively reconstruct all of the parameters. Thus a_ℓ and z_ℓ are uniquely determined by the $\{T_n\}$, so they are independent of the choice (i, j) . Finally, observe that $\sum_{\ell} a_\ell = \sum_{v \neq j} w(i, v)$. Therefore

$$w(i, j) = T_1 - \sum_{\ell} a_\ell,$$

which shows that for $i \neq j$, the value of $w(i, j)$ is constant. \square

Combining Lemma 18 with the results of Section 3 allows us to deduce Theorem 3.

Proof of Theorem 3. Suppose that G is a loopless weighted digraph such that $w(i, j) > 0$ for all distinct vertices i and j . Moreover, suppose that G has property (C) and that the associated insertion process is k -dependent process. We show that either: $G = K_3$ and $k \geq 2$; or $G = K_4$ and $k \geq 1$.

By Lemma 18, the graph G has uniform weight, and by Lemma 9 it is strongly connected. Applying Theorem 2, we conclude that either $G = K_{r,r,r}$ and $k \geq 2$, or $G = K_{r,r,r,r}$ and $k \geq 1$. Since $w(i, j) > 0$ for all $i \neq j$, it follows that $r = 1$. Thus G is a complete graph, and applying [13, Proposition 13] we deduce that G is either K_3 (for $k \geq 2$) or K_4 (for $k \geq 1$).

Conversely, by the main result of [13] the graphs K_3 and K_4 are k -dependent for $k \geq 2$ and $k \geq 1$ respectively. \square

5 Shifts of finite type

We turn to the proof of Theorem 4, which states that if the de Bruijn graph of a loopless shift of finite type satisfies property (EC), then the associated insertion process on the shift of finite type is not finitely dependent.

Proof of Theorem 4. Let G denote the de Bruijn graph of the shift of finite type and let $\{Y_\ell\}_{\ell \in \mathbb{Z}}$ denote the insertion process associated to G . We write

$$Y_\ell = (x_\ell, \dots, x_{\ell+n-1}) \in [q]^n.$$

Since $\{Y_\ell\}$ is almost surely a path in G , the overlapping elements in adjacent tuples Y_ℓ and $Y_{\ell+1}$ almost surely coincide. We extend the insertion process from G to the shift of finite type by considering the random sequence $\{x_\ell\}_{\ell \in \mathbb{Z}}$.

Suppose to the contrary that for some k , this process is k -dependent. Then the tuples $Y_\ell = (x_\ell, \dots, x_{\ell+n-1})$ form an $(n+k-1)$ -dependent sequence. Applying Lemma 8 to the insertion process on G , it follows that G has a directed triangle (a, b, c) . Hence we may write

$$a = (x_1, x_2, \dots, x_n), \quad b = (x_2, \dots, x_{n+1}), \quad c = (x_3, \dots, x_{n+2}).$$

Since the edge (a, c) is present in G , we must have

$$(x_2, x_3, \dots, x_n) = (x_3, \dots, x_{n+1}).$$

Thus $x_2 = x_3 = \dots = x_{n+1}$, so b is a constant sequence. This contradicts our assumption that the vertex set of G lacks elements of the form (i, \dots, i) . Therefore the associated insertion process is not k -dependent. \square

References

- [1] J. Aaronson, D. Gilat, and M. Keane. On the structure of 1-dependent Markov chains. *J. Theoret. Probab.*, 5(3):545–561, 1992. MR 1176437
- [2] J. Aaronson, D. Gilat, M. Keane, and V. de Valk. An algebraic construction of a class of one-dependent processes. *Ann. Probab.*, 17(1):128–143, 1989. MR 972778
- [3] N. Alon and O. N. Feldheim. A note on general sliding window processes. *Electron. Commun. Probab.*, 19:no. 66, 7, 2014. MR 3262072
- [4] B. Bollobás. *Modern graph theory*, volume 184 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998. MR 1633290
- [5] C. Borgs, J. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Counting graph homomorphisms. In *Topics in discrete mathematics*, volume 26 of *Algorithms Combin.*, pages 315–371. Springer, Berlin, 2006. MR 2249277
- [6] A. Borodin, P. Diaconis, and J. Fulman. On adding a list of numbers (and other one-dependent determinantal processes). *Bull. Amer. Math. Soc. (N.S.)*, 47(4):639–670, 2010. MR 2721041
- [7] R. M. Burton, M. Goulet, and R. Meester. On 1-dependent processes and k -block factors. *Ann. Probab.*, 21(4):2157–2168, 1993. MR 1245304
- [8] N. G. de Bruijn. A combinatorial problem. *Nederl. Akad. Wetensch., Proc.*, 49:758–764 = *Indagationes Math.* 8, 461–467 (1946), 1946. MR 0018142
- [9] V. de Valk. *One-dependent processes: two-block factors and non-two-block factors*, volume 85 of *CWI Tract*. Stichting Mathematisch Centrum, Centrum voor Wiskunde en Informatica, Amsterdam, 1994. MR 1269540
- [10] R. Durrett. *Probability: theory and examples*, volume 31 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, fourth edition, 2010. MR 2722836
- [11] A. E. Holroyd. One-dependent coloring by finitary factors, arXiv:1411.1463
- [12] A. E. Holroyd and T. M. Liggett. Symmetric 1-dependent colorings of the integers. *Electron. Commun. Probab.*, 20:no. 31, 8, 2015. MR 3327870

- [13] A. E. Holroyd and T. M. Liggett. Finitely dependent coloring. *Forum Math. Pi*, 4:e9, 43, 2016. MR 3570073
- [14] A. E. Holroyd, O. Schramm, and D. B. Wilson. Finitary coloring, arXiv:1412.2725
- [15] I. A. Ibragimov and Y. V. Linnik. *Independent and stationarily connected variables*. Izdat. “Nauka”, Moscow, 1965. MR 0202176
- [16] I. A. Ibragimov and Y. V. Linnik. *Independent and stationary sequences of random variables*. Wolters-Noordhoff Publishing, Groningen, 1971. With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, Translation from the Russian edited by J. F. C. Kingman. MR 0322926
- [17] S. Janson. Runs in m -dependent sequences. *Ann. Probab.*, 12(3):805–818, 1984. MR 744235
- [18] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002. MR 1876169
- [19] D. Lind and B. Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, Cambridge, 1995. MR 1369092

Acknowledgements

I thank Alexander Holroyd for posing the questions which inspired this paper. I am grateful to Alexander Holroyd, Christopher Hoffman, and Omer Angel for valuable discussions as well as their patience. Furthermore I thank the feedback of an anonymous referee for providing the inspiration of the (EC) condition, for suggesting Question 4, and for numerous helpful suggestions that greatly improved the prose.

Chapter 3

STABLE MATCHING WITH MALLOWS PREFERENCES

This is part of joint work with Christopher Hoffman and Elchanan Mossel.

3.1 Introduction

Suppose there are n men and n women, each of whom ranks every member of the opposite gender in such a way that there are no ties. A matching (in which every woman is partnered with a distinct man) is stable if there does not exist a woman and a man, each of whom prefers the other to their partner in the matching.

The Gale-Shapley algorithm [8] constructs a stable matching for any set of rankings, proving the non-trivial result that there always exists such a matching. This is a fundamental result in the economic theory of stable allocations, a theory which appeared in the citation of the 2012 Nobel Prize in Economics awarded to Lloyd Shapley and Alvin E. Roth.

A natural question is: how many stable matchings do typical rankings have? When the rankings are uniformly random and as $n \rightarrow \infty$, it is known that the expected number of stable matchings is asymptotic to $e^{-1}n \log n$ [21], that there are at least $n^{1/2-o(1)}$ stable matchings with high probability [22], and that the probability that there are at least $cn \log n$ stable matchings is bounded below by

$$\frac{2e}{2e+1} \approx .84$$

in the limit as $c \rightarrow 0$ [17]. However, it is unknown whether there exists a constant C such that the number of stable matchings is asymptotic to $Cn \log n$ with high probability.

On the other hand, it is known that there are highly structured instances with an exponential number of stable matchings. In their book on stable marriage [12], Gusfield and Irving

show that there are instances with 2^n stable matchings, for instance. In fact, one expects this type of exponential growth for rankings that are globally, but not locally, correlated. The basic idea is that the stable matchings decompose into approximately independent blocks of constant size, and each block carries a constant number of stable matchings.

We make this intuition rigorous for a ranking model in which the rankings are independent and *Mallows*-distributed. Each ranking is identified with a permutation π of $\{1, \dots, n\}$ by declaring $\pi(i) < \pi(j)$ if j is more desirable than i . The Mallows measure $\text{Mal} = \text{Mal}_{q,n}$ with parameter $q \in (0, 1)$ assigns to each π a probability proportional to $q^{\text{inv}(\pi)}$, where $\text{inv}(\pi)$ is the **inversion number**

$$\text{inv}(\pi) = \#\{1 \leq i < j \leq n: \pi(i) > \pi(j)\}.$$

This probability is maximized for the identity permutation, and decays exponentially in the inversion number; thus a Mallows random permutation may be regarded as a perturbation of the identity. Hence we regard the rankings in our model as being ‘globally correlated’ with the identity permutation, even though the rankings are independent of one another.

We use the following notation to describe the model formally. An **interval of integers** is a set of the form $I \cap \mathbb{Z}$ where $I \subseteq \mathbb{R}$ is an interval. We write $\llbracket a, b \rrbracket := [a, b] \cap \mathbb{Z}$ and similarly for other intervals.

For a (possibly infinite) interval $I \subseteq \mathbb{Z}$, let $\Omega_I := I \times \{\sigma, \varphi\}$ and let S_I be the set of permutations of I (i.e., bijections from I to itself). A **preference structure** \mathcal{P} on Ω_I is a map from Ω_I to S_I . Let $\text{PrefStruc}_I := (S_I)^{\Omega_I}$ denote the set of preference structures on Ω_I . For $i \in I$, we regard $\mathcal{P}((i, \varphi))$ as a ranking of $I \times \{\sigma\}$, and we say that (i, φ) **prefers** (j, σ) to (k, σ) if

$$\mathcal{P}((i, \varphi))(j) > \mathcal{P}((i, \varphi))(k).$$

Similarly, we say that (i, σ) prefers (j, φ) to (k, φ) if

$$\mathcal{P}((i, \sigma))(j) > \mathcal{P}((i, \sigma))(k).$$

A **match** is an unordered pair $\{(i, \sigma), (j, \varphi)\}$ with $i, j \in I$, and the **partner** of an element of a match is the other element. A **matching** of Ω_I is a set of matches for which every element

of Ω_I appears in exactly one match. (That is, we consider only perfect matchings, and not partial matchings.) Finally, a matching is **stable** with respect to \mathcal{P} if there do not exist integers i and j such that (i, σ) and (j, φ) prefer one another to their respective partners in the matching. When there is no risk of confusion, we refer to a matching of Ω_I that is stable with respect to \mathcal{P} as a **stable matching of \mathcal{P}** . We write $\text{StabMatch } \mathcal{P}$ for the set of stable matchings of \mathcal{P} .

In the special case when $I = \llbracket 1, n \rrbracket$, we write S_n , Ω_n , and PrefStruc_n in place of S_I , Ω_I , and PrefStruc_I , respectively.

For a parameter $q \in (0, 1)$ and an integer $n \geq 1$, let $\text{MalPref}_{q,n}$ be a random element of PrefStruc_n assigning independent Mal_q -distribution permutations to all individuals in Ω_n . In other words, the law of $\text{MalPref}_{q,n}$ is the product measure $(\text{Mal}_q)^{\Omega_n}$. We refer to $\text{MalPref}_{q,n}$ as being a **Mallows preference structure**.

We prove that the number of stable matchings of $\text{MalPref}_{q,n}$ grows exponentially in n .

Theorem 1. *For all $q \in (0, 1)$, there exists a constant $0 < \text{gr}_q < \infty$ such that*

$$\frac{\log \# \text{StabMatch } \text{MalPref}_{q,n}}{n} \xrightarrow[n \rightarrow \infty]{d} \text{gr}_q.$$

Furthermore, $\text{gr}_q \rightarrow 0$ as $q \rightarrow 0$.

The difficulty in analyzing stable matchings is that they are globally defined. Given a preference structure on Ω_n , we define an equivalence relation decomposing $\llbracket 1, n \rrbracket$ into intervals of integers in such a way that the stable matchings on the larger interval are determined by those on each equivalence class. We obtain results about the stable matchings by coupling preference structures on Ω_n with preference structures on $\Omega_{\mathbb{Z}}$. The latter is shift invariant which allows us to use the ergodic theorem.

3.2 Coupling with an ergodic system

In this section we couple the sequence of random preference structures $(\text{MalPref}_{q,n})_{n \geq 1}$ with a random infinite preference structure. The latter is stationary and ergodic, and applying the ergodic theorem to it will imply the existence of the constant gr_q in the main theorem.

For a (finite or infinite) set of integers $U \subseteq \mathbb{Z}$, a permutation of U is defined to be a bijection from U to itself. If T is a finite subset of U and π is a permutation of U , the permutation π_T **induced** by T is the unique permutation of T satisfying

$$\pi_T(s) < \pi_T(t) \iff \pi(s) < \pi(t), \quad s, t \in T.$$

In other words, the elements of T are given the same relative rankings under π and π_T .

The Mallows measure $\text{Mal}_{q,I}$ on permutations of a finite interval $I \subset \mathbb{Z}$ assigns a probability proportional to $q^{\text{inv}(\pi)}$ to each permutation π of I , where

$$\text{inv}(\pi) := \#\{i, j \in I : i < j, \pi(i) > \pi(j)\}.$$

We write $\text{MalPref}_{q,I}$ for a random preference structure assigning independent $\text{Mal}_{q,I}$ -distributed permutations to the elements of Ω_I .

There is an infinite extension of the Mallows measure having the following properties.

Theorem ([10]). *For all $q \in [0, 1)$ there exists a measure $\text{Mal}_{q,\mathbb{Z}}$ on permutations of \mathbb{Z} such that if π has law $\text{Mal}_{q,\mathbb{Z}}$, then for all finite intervals $I \subset \mathbb{Z}$ the induced permutation π_I has law $\text{Mal}_{q,I}$. Moreover, the following stationarity property holds:*

$$(\pi(i))_{i \in \mathbb{Z}} \stackrel{d}{=} (\pi(i+1) - 1)_{i \in \mathbb{Z}}.$$

We define $\text{MalPref}_{q,\mathbb{Z}}$ to be a random preference structure assigning to each individual in $\Omega_{\mathbb{Z}}$ an independent $\text{Mal}_{q,\mathbb{Z}}$ -distributed permutation. The law of $\text{MalPref}_{q,\mathbb{Z}}$ is $(\text{Mal}_{q,\mathbb{Z}})^{\Omega_{\mathbb{Z}}}$.

Let $I \subset \mathbb{Z}$ be a finite interval of integers. For a preference structure \mathcal{P} on $\Omega_{\mathbb{Z}}$, we define the preference structure \mathcal{P}_I **induced** by I to be the preference structure of Ω_I given by

$$\mathcal{P}_I((k, \varphi^\natural)) = \mathcal{P}((k, \varphi^\natural))_I, \quad k \in I, \varphi^\natural \in \{\sigma^\natural, \varphi^\natural\}.$$

Observe that for all finite intervals $I \subset \mathbb{Z}$,

$$(\text{MalPref}_{q,\mathbb{Z}})_I \stackrel{d}{=} \text{MalPref}_{q,I}.$$

In particular, taking $I = \llbracket 1, n \rrbracket$ yields a coupling of $(\text{MalPref}_{q,n})_{n \geq 1}$ with $\text{MalPref}_{q,\mathbb{Z}}$.

Let Θ map $\text{PrefStruc}_{\mathbb{Z}}$ to itself by relabeling person i to person $i + 1$. More precisely,

$$\Theta\mathcal{P}((i, \varphi))(j) = \mathcal{P}((i + 1, \varphi))(j + 1) - 1, \quad i, j \in \mathbb{Z}, \varphi \in \{\sigma, \varphi\}.$$

Lemma 2. *The triple $(\text{PrefStruc}_{\mathbb{Z}}, \Theta, (\text{Mal}_q)^{\Omega_{\mathbb{Z}}})$ is an ergodic measure-preserving system.*

That the system is measure-preserving follows from stationarity of $\text{Mal}_{q,\mathbb{Z}}$. The proof of ergodicity is a variant of the well-known proof of Kolmogorov's 0-1 law [15].

Proof. Let \mathcal{P} be a random preference structure with law $(\text{Mal}_{q,\mathbb{Z}})^{\Omega_{\mathbb{Z}}}$. Let \mathcal{I} be the invariant σ -algebra, consisting of all events $A \subseteq \text{PrefStruc}_{\mathbb{Z}}$ satisfying $A = \Theta^{-1}A$. We wish to show that \mathcal{I} is trivial. For a finite interval of integers $I \subset \mathbb{Z}$, consider the σ -algebra

$$\mathcal{F}_I = \sigma\left(\mathcal{P}((i, \varphi)) : i \in I, \varphi \in \{\sigma, \varphi\}\right).$$

Fix $n \geq 1$. If $A \in \mathcal{F}_{[-n,n]}$, then for all $i \in \mathbb{Z}$ we have that $\Theta^i A \in \mathcal{F}_{[-n+i,n+i]}$. In particular, if $i > 2n$ then A is independent of $\Theta^i A$. Thus the σ -algebra $\mathcal{I} \cap \mathcal{F}_{[-n,n]}$ is trivial. The result now follows by taking $n \rightarrow \infty$ and using a standard approximation argument. \square

A map $F: \text{PrefStruc}_{\mathbb{Z}} \rightarrow \mathbb{R}^{\mathbb{Z}}$ satisfying

$$F(\Theta\mathcal{P})_i = F(\mathcal{P})_{i+1}, \quad i \in \mathbb{Z}, \mathcal{P} \in \text{PrefStruc}_{\mathbb{Z}},$$

is called a **factor map**. It intertwines Θ with the shift map on $\mathbb{R}^{\mathbb{Z}}$.

Corollary 3. *If $F: \text{PrefStruc}_{\mathbb{Z}} \rightarrow \mathbb{R}^{\mathbb{Z}}$ is a factor map and $\text{MalPref}_{q,\mathbb{Z}}$ is a random variable with law $(\text{Mal}_q)^{\Omega_{\mathbb{Z}}}$, then the random sequence $F(\text{MalPref}_{q,\mathbb{Z}})$ is stationary and ergodic.*

Proof. This is an immediate consequence of the previous lemma. \square

Combining the pointwise ergodic theorem [15] with Lemma 2 implies the following. For any (product Borel) measurable function $f: \text{PrefStruc}_{\mathbb{Z}} \rightarrow \mathbb{R}$ for which $f(\text{MalPref}_{q,\mathbb{Z}})$ is integrable, we have that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f(\Theta^i \text{MalPref}_{q,\mathbb{Z}}) = \mathbb{E}f(\text{MalPref}_{q,\mathbb{Z}}) \quad a.s. \quad (1)$$

In particular, taking $f = \mathbf{1}[A]$ for A a measurable subset of $\text{PrefStruc}_{\mathbb{Z}}$ yields that

$$\lim_{n \rightarrow \infty} \frac{\#\{1 \leq i \leq n: \Theta^i \text{MalPref}_{q,\mathbb{Z}} \in A\}}{n} = \mathbb{P}(\text{MalPref}_{q,\mathbb{Z}} \in A) \quad a.s. \quad (2)$$

Preference structures whose stable matchings decompose at a given location play a crucial role in our argument. Fix integers i, j , and k with $i \leq j \leq k$ and a preference structure \mathcal{P} on $\Omega_{[[i,k]]}$. We say that j is a **lattice cutpoint** of \mathcal{P} if the set of matchings of $\Omega_{[[i,k]]}$ that are stable with respect to \mathcal{P} coincides with the set of matchings arising as the union of a stable matching of $\mathcal{P}_{[[i,j]]}$ with a stable matching of $\mathcal{P}_{(j,k]}$.

Lattice cutpoints are our tool for counting the number of stable matchings, since if $i \leq j \leq k$ are integers and j is a lattice cutpoint of $\mathcal{P}_{[[i,k]]}$ then

$$\# \text{StabMatch } \mathcal{P}_{[[i,k]]} = \# \text{StabMatch } \mathcal{P}_{[[i,j]]} \cdot \# \text{StabMatch } \mathcal{P}_{(j,k]}. \quad (3)$$

(In fact, there is a natural lattice structure on the set of stable matchings [4], and (3) generalizes to a lattice isomorphism between $\text{StabMatch } \mathcal{P}_{[[i,k]]}$ and the product of the lattices $\text{StabMatch } \mathcal{P}_{[[i,j]]}$ and $\text{StabMatch } \mathcal{P}_{(j,k]}$; hence the name.)

For a (possibly infinite) interval $I \subseteq \mathbb{Z}$ and for a preference structure \mathcal{P} on Ω_I , we say that $j \in I$ is an **essential lattice cutpoint** of \mathcal{P} if it is a lattice cutpoint of $\mathcal{P}_{[[i,k]]}$ for all $i, k \in I$ with $i \leq j \leq k$. Note that only finite preference structures appear in the definition. Denote by $\text{ELat}(\mathcal{P})$ the set of essential lattice cutpoints of \mathcal{P} .

The random set $\text{ELat}(\text{MalPref}_{q,\mathbb{Z}})$ plays a fundamental role in our argument. We define

$$\rho_q := \mathbb{P}(0 \in \text{ELat}(\text{MalPref}_{q,\mathbb{Z}})).$$

Observe that $\text{ELat}(\text{MalPref}_{q,\mathbb{Z}})$ is stationary and ergodic.

Lemma 4. *For all $q \in (0, 1)$,*

$$\lim_{n \rightarrow \infty} \frac{\#(\text{ELat}(\text{MalPref}_{q,\mathbb{Z}}) \cap [[1, n]])}{n} = \rho_q \quad a.s.$$

Proof. Apply the ergodic theorem in the form (2), with $A = \{0 \in \text{ELat}(\text{MalPref}_{q,\mathbb{Z}})\}$. \square

Essential lattice cutpoints allow us to define the following local quantities, for a preference structure \mathcal{P} on $\Omega_{\mathbb{Z}}$ and an integer i . Let $\text{Nbhd}(\mathcal{P}, i) := (\ell, r]$ where ℓ and r are largest and smallest elements of the sets

$$\llbracket -\infty, i \rrbracket \cap (\text{ELat}(\mathcal{P}) \cup \{-\infty\}) \quad \text{and} \quad \llbracket i, \infty \rrbracket \cap (\text{ELat}(\mathcal{P}) \cup \{\infty\}),$$

respectively. That is, $\text{Nbhd}(\mathcal{P}, i)$ is the interval whose endpoints are the essential lattice cutpoints surrounding i , inclusive on the right and exclusive on the left. We write $\text{LocPref}(\mathcal{P}, i)$ for the induced preference structure $\mathcal{P}_{\text{Nbhd}(\mathcal{P}, i)}$, and we set

$$\text{gr}(\mathcal{P}, i) := \frac{\log \# \text{StabMatch LocPref}(\mathcal{P}, i)}{\# \text{Nbhd}(\mathcal{P}, i)}, \quad (4)$$

provided $\text{Nbhd}(\mathcal{P}, i)$ is finite, and otherwise $\text{gr}(\mathcal{P}, i) := 0$. This quantity is a local approximation at i to the exponential growth rate of the number of stable matchings. Observe that $\text{gr}(\mathcal{P}, i) = \text{gr}(\Theta^{-i}\mathcal{P}, 0)$, from which it follows by earlier general remarks that the sequence $(\text{gr}(\text{MalPref}_{q, \mathbb{Z}}, i))_{i \in \mathbb{Z}}$ is stationary and ergodic. We also point out that

$$\text{gr}(\mathcal{P}, i) \leq \log \# \text{Nbhd}(\mathcal{P}, i), \quad (5)$$

since the number of (perfect) matchings is equal to $(\# \text{Nbhd}(\mathcal{P}, i))!$.

The next result is crucial for the proof of the main theorem.

Proposition 5. *For all $q \in (0, 1)$ such that $\text{gr}(\text{MalPref}_{q, \mathbb{Z}}, 0)$ is integrable and $\rho_q > 0$,*

$$\frac{\log \# \text{StabMatch MalPref}_{q, n}}{n} \xrightarrow[n \rightarrow \infty]{d} \mathbb{E} \text{gr}(\text{MalPref}_{q, \mathbb{Z}}, 0).$$

Thus, the constant gr_q in the main theorem is given by $\mathbb{E} \text{gr}(\text{MalPref}_{q, \mathbb{Z}}, 0)$, whenever the above hypotheses hold. Once we show that $\rho_q > 0$ and $0 < \text{gr}_q < \infty$ for all $q \in (0, 1)$, the main theorem will follow. The proofs of these assertions are non-trivial and they will occupy the remaining sections in this chapter.

Proof of Proposition 5. Fix $q \in (0, 1)$ satisfying the hypotheses. Let $\mathcal{P} = \text{MalPref}_{q, \mathbb{Z}}$, let $X_n = \log \# \text{StabMatch } \mathcal{P}_{\llbracket 1, n \rrbracket}$, and let $Y_n = \sum_{i=1}^n \text{gr}(\mathcal{P}, i)$. By the ergodic theorem, $\lim_{n \rightarrow \infty} \frac{1}{n} Y_n = \text{gr}_q$ a.s.

We show that $\frac{1}{n}|Y_n - X_n|$ converges to 0 in probability. Let $N = \min(\text{ELat}(\mathcal{P}) \cap \llbracket 1, \infty \rrbracket)$. Define the random variables $(M_n)_{n \geq 1}$ via

$$M_n := \begin{cases} n - \max(\text{ELat}(\mathcal{P}) \cap \llbracket 1, n \rrbracket), & N \leq n \\ 0, & N > n. \end{cases}$$

On the event $N \leq n$, the integers N and $n - M_n$ are the minimal and maximal elements of $\text{ELat}(\mathcal{P}) \cap \llbracket 1, n \rrbracket$, respectively. The quantity $\text{gr}(\mathcal{P}, i)$ is constant for $i \in \llbracket 1, N \rrbracket$ and for $i \in \llbracket n - M_n, n \rrbracket$. Set $X'_n = \log \# \text{StabMatch } \mathcal{P}_{\llbracket N, n - M_n \rrbracket}$. By (3),

$$\begin{aligned} X_n &= \log \# \text{StabMatch } \mathcal{P}_{\llbracket 1, N \rrbracket} + X'_n + \log \# \text{StabMatch } \mathcal{P}_{\llbracket n - M_n, n \rrbracket} \\ Y_n &= N \text{gr}(\mathcal{P}, 1) + X'_n + M_n \text{gr}(\mathcal{P}, n). \end{aligned}$$

Let $f(x) = x \log x$ for $x \geq 1$. By the last equations and (5),

$$|X_n - Y_n| \leq \max(f(N), f(M_n)) + \max\left(f(\# \text{Nbhd}(\mathcal{P}, 1)), f(\# \text{Nbhd}(\mathcal{P}, n))\right)$$

As $N \leq \# \text{Nbhd}(\mathcal{P}, 1)$ and $M_n \leq \# \text{Nbhd}(\mathcal{P}, n)$ and f is increasing, it follows that

$$|X_n - Y_n| \leq 2 \max\left(f(\# \text{Nbhd}(\mathcal{P}, 1)), f(\# \text{Nbhd}(\mathcal{P}, n))\right). \quad (6)$$

While the bound (6) was obtained on the event $N \leq n$, it holds trivially on the event $N > n$ as well. Thus for any $\epsilon > 0$, we have that

$$\begin{aligned} \mathbb{P}(|X_n - Y_n| > \epsilon n) &\leq \mathbb{P}\left(f(\# \text{Nbhd}(\mathcal{P}, 1)) > \frac{\epsilon n}{2}\right) + \mathbb{P}\left(f(\# \text{Nbhd}(\mathcal{P}, n)) > \frac{\epsilon n}{2}\right) \\ &= 2\mathbb{P}\left(f(\# \text{Nbhd}(\mathcal{P}, 1)) > \frac{\epsilon n}{2}\right). \end{aligned} \quad (7)$$

Combining the hypothesis $\rho > 0$ with Lemma 4, we see that $\# \text{Nbhd}(\mathcal{P}, 1)$ is a.s. finite. Thus as $n \rightarrow \infty$, the probability in (7) tends to 0, implying that $\frac{1}{n}|X_n - Y_n|$ converges to 0 in probability as $n \rightarrow \infty$. When combined with the a.s. convergence of $\frac{1}{n}Y_n$ to gr_q , it follows that $\frac{1}{n}X_n$ converges in distribution to gr_q . \square

3.3 Cutpoint criteria

Essential lattice cutpoints of preference structures were introduced for the purpose of splitting the task of enumerating stable matchings into smaller independent parts. A preference structure $\mathcal{P} \in \text{PrefStruc}_{\mathbb{Z}}$ induces an equivalence relation on \mathbb{Z} by taking the equivalence classes to be the intervals between consecutive essential lattice cutpoints; this is the equivalence relation referred to in the introduction. For any finite interval of integers I that is a union of equivalence classes, the number of stable matchings of the induced preference structure \mathcal{P}_I factors as a product of the number of stable matchings on each of the equivalence classes contained in I . Thus if we can locate essential lattice cutpoints, we will be able to relate the exponential growth rate of the number of stable matchings to a local quantity.

We show that essential lattice cutpoints can be detected by a certain family of inequalities for the preference structure. In fact, we start by showing that simpler inequalities yield a type of cutpoint weaker than the essential lattice cutpoint, defined as follows. Given integers $i \leq j \leq k$ and a matching M of $\Omega_{[i,k]}$, we say that j is a **cutpoint** of M if there do not exist integers a and b with $i \leq a \leq j$ and $j+1 \leq b \leq k$ such that either (a, φ) and (b, σ) are matched, or (a, σ) and (b, φ) are matched. It follows from the definitions that if j is an essential lattice cutpoint of a preference structure \mathcal{P} on some interval $I \ni j$, then for all $i, k \in I$ with $i \leq j \leq k$ we have that j is a cutpoint of the induced preference structure $\mathcal{P}_{[i,k]}$.

Lemma 6. *Let $I \subset \mathbb{Z}$ be a finite interval, let \mathcal{P} be a preference structure on Ω_I , and let $s \in \mathbb{Z} + \frac{1}{2}$ be a half-integer. If for all $i, j \in I$ and both $\varphi \in \{\sigma, \varphi\}$*

$$\left| \mathcal{P}((i, \varphi))(j) - j \right| < \frac{|i - s| + |j - s|}{20}, \quad (8)$$

then $\lfloor s \rfloor$ is a cutpoint of every stable matching of \mathcal{P} .

We say that \mathcal{P} satisfies the **cutpoint bound** at location s whenever (8) holds for all $i, j \in I$ and both $\varphi \in \{\sigma, \varphi\}$. Before presenting the proof, we record a simple result.

Lemma 7. *Let $I \subset \mathbb{Z}$ be a finite interval and let M be a (perfect) matching of Ω_I . Given a match $\{(a, \sigma), (b, \varphi)\}$ in M with $a < b$ and given any real number $t \in (a, b)$, there exist integers a' and b' with $a' < t < b'$ such that the match $\{(a', \varphi), (b', \sigma)\}$ is also in M .*

Proof. If no such a' and b' existed, then the matching would furnish a bijection between sets whose cardinalities differ by one. This is impossible. \square

The proof of Lemma 6 is rather subtle. In words, the cutpoint bound (8) says that preferences do not stray too far from being in order, with the amount of variation increasing at most linearly in the distance from the purported cutpoint location. Therefore if a long match appears in some stable matching, it is not caused by exotic preferences but rather by an even longer match going in the opposite direction and involving a person even higher up in the ranking. Since we are working with a finite system, this cannot go on forever. Therefore there are, in fact, no long matches.

Here ‘longness’ is measured relative to the distance from the cutpoint location. In fact, finding the correct definition is a delicate balancing act. Fortunately this task rests with the authors rather than with the reader, and in the end the definition is not too complicated.

Proof of Lemma 6. Fix I , s , and \mathcal{P} as in the lemma, as well as a stable matching M of \mathcal{P} . We claim that the set

$$\mathcal{S} := \left\{ \begin{array}{l} (i, j) \in I \times I: i > s \text{ and } i - s > 3(j - s) \text{ and either} \\ (i, \varphi) \text{ and } (j, \sigma) \text{ are matched, or } (i, \sigma) \text{ and } (j, \varphi) \text{ are matched} \end{array} \right\}$$

is empty, from which the result will follow. That is, there are no long matches.

Suppose to the contrary that \mathcal{S} is non-empty. Since \mathcal{S} is finite, we may then select a pair $(i, j) \in \mathcal{S}$ with i maximal. Suppose, for the sake of concreteness, that it is (i, φ) which is matched with (j, σ) , rather than the other way around (the argument is unmodified apart from interchanging the roles of σ and φ). The equation $i - s > 3(j - s)$ is equivalent to $j < \frac{1}{3}(i + 2s)$, and therefore when $(i, j) \in \mathcal{S}$ we have that

$$j < \frac{1}{3}i + \frac{2}{3}s < \frac{2}{3}i + \frac{1}{3}s < i, \tag{9}$$

the latter two inequalities being equivalent to $i > s$. Applying Lemma 7 with $t = \frac{2}{3}i + \frac{1}{3}s$, it follows that there exist integers i' and j' such that (i', σ) is matched with (j', φ) and

$$j' < \frac{2}{3}i + \frac{1}{3}s < i'. \quad (10)$$

In particular, it follows that from (10) that $i' > s$.

We claim that $\pi = \mathcal{P}((i, \varphi))$ ranks i' ahead of j . By (8), we have that

$$\pi(i') \geq i' - \frac{1}{20}(|i - s| + |i' - s|) \quad \text{and} \quad \pi(j) \leq j + \frac{1}{20}(|i - s| + |j - s|),$$

so it suffices to verify that

$$i' - \frac{1}{20}(|i - s| + |i' - s|) > j + \frac{1}{20}(|i - s| + |j - s|). \quad (11)$$

The left side of this inequality is increasing in i' , and the right side is increasing in j . Thus by (10) it suffices to verify it with $\frac{2}{3}i + \frac{1}{3}s$ in place of i' , and by (9) it suffices to verify it with $\frac{1}{3}i + \frac{2}{3}s$ in place of j , i.e. (11) is implied by

$$\left(\frac{2}{3}i + \frac{1}{3}s\right) - \frac{|i - s|}{20} - \frac{\left|\left(\frac{2}{3}i + \frac{1}{3}s\right) - s\right|}{20} > \left(\frac{1}{3}i + \frac{2}{3}s\right) + \frac{|i - s|}{20} + \frac{\left|\left(\frac{1}{3}i + \frac{2}{3}s\right) - s\right|}{20}. \quad (12)$$

The expressions inside the absolute values in (12) are all positive, and upon expansion the inequality reduces to $\frac{11}{60}i > \frac{11}{60}s$. Thus $\pi(i') > \pi(j)$.

By stability of the matching, it follows that $\mathcal{P}((i', \sigma))$ ranks j' ahead of i . Thus, by another application of (8), we have that

$$j' + \frac{1}{20}(|i' - s| + |j' - s|) > i - \frac{1}{20}(|i' - s| + |i - s|).$$

Since the left side is increasing in j' and $j' < \frac{2}{3}i + \frac{1}{3}s$, this inequality implies that

$$\frac{2}{3}i + \frac{1}{3}s + \frac{1}{20}(|i' - s| + \left|\left(\frac{2}{3}i + \frac{1}{3}s\right) - s\right|) > i - \frac{1}{20}(|i' - s| + |i - s|).$$

All quantities in the absolute values are seen to be positive, and after some algebra this inequality reduces to $i' - s > \frac{5}{2}(i - s)$. In particular, $i' > i$. Moreover, since $i - s > \frac{3}{2}(j' - s)$ by (10), we have that

$$i' - s > \frac{5}{2} \cdot \frac{3}{2}(j' - s) > 3(j' - s), \quad (13)$$

implying that $(i', j') \in \mathcal{S}$. This contradicts maximality of the pair (i, j) , showing that the set \mathcal{S} is in fact empty. Thus s is a cutpoint of the stable matching M . The result follows since M was chosen arbitrary from the stable matchings of \mathcal{P} . \square

Having succeeded in locating *cutpoints*, we now strengthen the criterion in order to locate *essential lattice cutpoints*. Recall that the cutpoint bound (8) holds if $|\pi(j) - j|$ is sufficiently small for all relevant π and j . One may regard $|\pi(j) - j|$ as a local measure of deviation from the identity. However, this quantity is non-monotonic with respect to induced permutations: it may decrease or increase upon passage to an induced permutation. We start by introducing quantities that do not suffer from this issue.

For integers $a \leq j \leq b$ and a permutation π of $\llbracket a, b \rrbracket$, we set

$$\begin{aligned} L_{++} &= \#\{j < i \leq b: \pi(i) > \pi(j)\}, & L_{+-} &= \#\{j < i \leq b: \pi(i) < \pi(j)\}, \\ L_{-+} &= \#\{a \leq i < j: \pi(i) > \pi(j)\}, & L_{--} &= \#\{a \leq i < j: \pi(i) < \pi(j)\}. \end{aligned} \quad (14)$$

In particular,

$$L_{++}, L_{+-} \in \llbracket 0, b - j \rrbracket \quad \text{and} \quad L_{-+}, L_{--} \in \llbracket 0, j - a \rrbracket. \quad (15)$$

As appropriate, we write $L_{++} = L_{++}(\pi, j) = L_{++}^{\llbracket a, b \rrbracket}(\pi, j)$ and so on. We set

$$\text{offset}(\pi, j) := \max(L_{+-}(\pi, j), L_{-+}(\pi, j)). \quad (16)$$

The quantity $\text{offset}(\pi, j)$ serves as a monotonic replacement for $|\pi(j) - j|$.

Lemma 8. *Fix a finite interval $J \subset \mathbb{Z}$ and a permutation π of J .*

- (i) *For all $j \in J$, we have that $|\pi(j) - j| \leq \text{offset}(\pi, j)$.*
- (ii) *For any subinterval $I \subseteq J$ and any $i \in I$, we have that $\text{offset}(\pi_I, i) \leq \text{offset}(\pi, i)$.*
- (iii) *If $i, j \in J$ satisfy $i \geq j$ and $\pi(i) \leq \pi(j)$, then*

$$\max(\text{offset}(\pi, i), \text{offset}(\pi, j)) \geq i - j.$$

Proof. Adding the equations $L_{++} + L_{+-} = b - j$ and $L_{+-} + L_{--} = \pi(j) - a$ yields

$$L_{++} + 2L_{+-} + L_{--} = \pi(j) - j + b - a, \quad (17)$$

whereas

$$L_{++} + L_{+-} + L_{-+} + L_{--} = \#\{a \leq i \leq b: \pi(i) \neq \pi(j)\} = b - a. \quad (18)$$

Subtracting (18) from (17) yields

$$\pi(j) - j = L_{+-}(\pi, j) - L_{-+}(\pi, j),$$

which implies (i).

Property (ii) is immediate from the definitions.

To establish (iii), apply (i) twice to obtain that

$$\max(|\pi(i) - i|, |\pi(j) - j|) \leq \max(\text{offset}(\pi, i), \text{offset}(\pi, j)).$$

The bound now follows since $|\pi(j) - j - \pi(i) + i| \leq \max(|\pi(i) - i|, |\pi(j) - j|)$ and

$$|\pi(j) - j - \pi(i) + i| = \pi(j) - \pi(i) + i - j \geq i - j. \quad \square$$

The next lemma is the main result of this section.

Lemma 9. *Let $s \in \mathbb{Z} + \frac{1}{2}$ be a half-integer, let $I \subseteq \mathbb{Z}$ be a (possibly infinite) interval, and let \mathcal{P} be a preference structure on Ω_I . If for all $i, j \in I$, for all finite intervals J with $\{i, j\} \subseteq J \subseteq I$, and for both $\varphi \in \{\sigma, \varphi\}$,*

$$\text{offset}\left(\mathcal{P}_J((i, \varphi)), j\right) < \frac{|i - s| + |j - s|}{20}, \quad (19)$$

then $\lfloor s \rfloor$ is an essential lattice cutpoint of \mathcal{P} .

We say that \mathcal{P} satisfies the **lattice cutpoint bound** at s whenever (19) holds for all i, j, J , and φ as above. By Lemma 8 (i) and (ii), if a preference structure \mathcal{P} on Ω_I satisfies the lattice cutpoint bound at s then \mathcal{P}_J satisfies the cutpoint bound at s for every finite interval J satisfying $s \in J \subseteq I$.

Proof of Lemma 9. Fix s and \mathcal{P} satisfying the conditions of the lemma. We must show that for all integers i and j with $i < s < j$, the induced preference structure $\mathcal{P}_{\llbracket i, j \rrbracket}$ has a lattice cutpoint at $\lfloor s \rfloor$. By Lemma 8 (i), (19), and Lemma 6, every stable matching of $\mathcal{P}_{\llbracket i, j \rrbracket}$ has a cutpoint at $\lfloor s \rfloor$. To verify that it is a *lattice* cutpoint, it remains to show that the union of any stable matching of $\mathcal{P}_{\llbracket i, k \rrbracket}$ with any stable matching of $\mathcal{P}_{\llbracket k, j \rrbracket}$ is stable under $\mathcal{P}_{\llbracket i, j \rrbracket}$.

Fix any two such stable matchings and suppose to the contrary that the union M is not stable with respect to $\mathcal{P}_{\llbracket i, j \rrbracket}$. Then there exist integers $a, b \in \llbracket i, j \rrbracket$ such that (a, φ) and (b, σ) each prefer the other (under $\mathcal{P}_{\llbracket i, j \rrbracket}$) to their respective partners under M . Let $s = k + \frac{1}{2}$. By stability of the restricted matchings, either $a > s > b$ or $a < s < b$. Assume, without loss of generality, that

$$a > s > b. \quad (20)$$

Let (a', φ) and (b', σ) be the partners of (b, σ) and (a, φ) under M , respectively. Since M is a union of matchings of $\Omega_{\llbracket i, k \rrbracket}$ and $\Omega_{\llbracket k, j \rrbracket}$, it follows that $b' > s > a'$.

By definition of a, b , and induced preference structures, the permutation $\pi = \mathcal{P}((a, \varphi))$ has $\pi(b) > \pi(b')$. By Lemma 8, the induced preference structure $\mathcal{P}_{\llbracket k, j \rrbracket}$ satisfies the cutpoint bound (8) at s . Hence by the proof of Lemma 6, the match $\{(a, \varphi), (b', \sigma)\}$ satisfies

$$a - s \leq 3(b' - s),$$

for otherwise it would be contained in the set shown to be empty in the proof. That is,

$$b' \geq \frac{1}{3}a + \frac{2}{3}s. \quad (21)$$

Combining the cutpoint bound with the inequality $\pi(b) > \pi(b')$ yields

$$b + \frac{1}{20}(|a - s| + |b - s|) > b' - \frac{1}{20}(|a - s| + |b' - s|). \quad (22)$$

The left and right sides of (22) are increasing in b and b' , respectively. By (20), replacing each occurrence of b with s on the left side increases its value. By (21), replacing each occurrence of b' with $\frac{1}{3}a + \frac{2}{3}s$ on the right decreases its value. Thus (22) implies that

$$s + \frac{|a - s|}{20} > \left(\frac{1}{3}a + \frac{2}{3}s\right) - \frac{1}{20}(|a - s| + |(\frac{1}{3}a + \frac{2}{3}s) - s|).$$

The expressions inside the absolute values are all positive, and upon expansion the inequality simplifies to $\frac{17}{30}s > \frac{17}{30}a$, contradicting (20). \square

3.4 Two positive probabilities

In this section we show that two key probabilities are positive for all $q \in (0, 1)$:

- the probability ρ_q that 0 is an essential lattice cutpoint of $\text{MalPref}_{q,\mathbb{Z}}$, and
- the probability that $\text{LocPref}(\text{MalPref}_{q,\mathbb{Z}}, 0)$ has multiple stable matchings.

Proposition 10. *For all $q \in (0, 1)$ we have that $\rho_q > 0$, and $\rho_q \rightarrow 1$ as $q \rightarrow 0$.*

Proposition 11. *For all $q \in (0, 1)$ we have that $\text{gr}(\text{MalPref}_{q,\mathbb{Z}}, 0) > 0$ with positive probability.*

These results are established by careful analysis of the lattice cutpoint bound (19). We begin by estimating some fundamental Mallows probabilities.

Lemma 12. *Let $I \subset \mathbb{Z}$ be a finite interval and let π be a random permutation of I with law $\text{Mal}_{q,I}$. For all real $t \geq 0$ and all $j \in I$,*

$$\mathbb{P}(\text{offset}(\pi, j) \geq t) \leq 2q^t. \quad (23)$$

Furthermore for all $i, j \in I$ with $i \geq j$,

$$\mathbb{P}(\pi(i) \leq \pi(j)) \leq 4q^{i-j}. \quad (24)$$

Recall the definitions of $L_{+-}(\pi, j)$ and $L_{-+}(\pi, j)$ in (14). Both count inversions involving j . Also recall that $\text{offset}(\pi, j)$ is the maximum of these two quantities (16).

Proof. Inequality (24) follows from (23). Indeed, by Lemma 8 (iii), since $i \geq j$ we have that

$$\mathbb{P}(\pi(i) \leq \pi(j)) \leq \mathbb{P}(\text{offset}(\pi, i) \geq i - j) + \mathbb{P}(\text{offset}(\pi, j) \geq i - j).$$

We will deduce (23) from the stronger inequalities

$$\mathbb{P}(L_{+-}(\pi, j) \geq k) \leq q^k \quad (25)$$

and

$$\mathbb{P}(L_{-+}(\pi, j) \geq k) \leq q^k, \quad (26)$$

where $k \geq 0$ is an integer.

We prove (25), and the proof of (26) is similar. Consider the map from permutations σ of I to sequences given by $\sigma \mapsto (L_{+-}(\sigma, \ell))_{\ell \in I}$. It is straightforward to verify by induction on $\#I$ that this map is injective. Furthermore by (15), the image is contained in the set of sequences $(s_\ell)_{\ell \in I}$ satisfying $0 \leq s_\ell \leq \max I - \ell$ for all $\ell \in I$. There are $(\#I)!$ such sequences. Since this equals the number of permutations, it follows that the map is a bijection onto this set of sequences.

Consider the pushforward of $\text{Mal}_{q,I}$ under this map. Since $\text{inv}(\sigma) = \sum_{\ell \in I} L_{+-}(\sigma, \ell)$, it follows that if π has law $\text{Mal}_{q,I}$ then $(L_{+-}(\pi, \ell))_{\ell \in I}$ is a sequence of independent truncated geometric random variables. In particular,

$$\mathbb{P}(L_{+-}(\pi, j) \geq k) = \frac{q^k + q^{k+1} + \dots + q^{\max I - j}}{1 + q + \dots + q^{\max I - j}} \leq q^k,$$

establishing (25). Replacing the bounds for L_{+-} with those for L_{-+} in the argument above yields (26). Lastly, (24) follows from (25) and (26) by the union bound. \square

For finite intervals of integers $I \subseteq J$, denote the set of elements of J that are smaller than $\min I$ by $(J \setminus I)^-$, and denote the elements larger than $\max I$ by $(J \setminus I)^+$. Note that $J = (J \setminus I)^- \cup I \cup (J \setminus I)^+$. For a permutation τ of I , denote by $\text{Res}(\tau, I, J)$ the set of permutations of J which fix the sets $(J \setminus I)^\pm$ and whose restriction to I is τ . For each τ , there is an evident bijection from $\text{Res}(\tau, I, J)$ to pairs consisting of a permutation of $(J \setminus I)^-$ and a permutation of $(J \setminus I)^+$. Define $\text{Fix}(I, J) := \text{Res}(\text{id}, I, J)$, where id is the identity permutation fixing every element of I .

We estimate the probability that a Mallows-distributed permutation lies in $\text{Res}(\tau, I, J)$. In fact, we show that it is bounded below by a quantity exponentially small in $\#I$ and

independent of J . During this calculation, the function

$$\phi(q) := \prod_{k=1}^{\infty} (1 - q^k)$$

will appear. Note that $\phi(q) > 0$ for all $q \in (0, 1)$, since $\sum_{k \geq 1} q^k$ is finite.

Lemma 13. *Fix finite intervals of integers $I \subseteq J$, fix a permutation τ of I , and fix $q \in [0, 1)$. Let π be a random permutation of J with law $\text{Mal}_{q,J}$. Conditional on the event that π fixes each of $(J \setminus I)^-$, I , and $(J \setminus I)^+$, the restrictions of π to each of $(J \setminus I)^-$, I , and $(J \setminus I)^+$ are independent, Mal_q -distributed permutations. Furthermore,*

$$\mathbb{P}(\pi \in \text{Res}(\tau, I, J)) \geq q^{\text{inv}(\tau)} (1 - q)^{\#I} \phi(q). \quad (27)$$

In the proof we will use the following well-known fact (Cor. 1.3.13, p. 36 in [24]):

$$\sum_{\sigma \in S_n} q^{\text{inv}(\sigma)} = \prod_{k=1}^n (1 + q + \dots + q^{k-1}) = (1 - q)^{-n} \prod_{k=1}^n (1 - q^k). \quad (28)$$

This is straightforward to verify by induction on n . Set $Z(q, n) := \sum_{\sigma \in S_n} q^{\text{inv}(\sigma)}$.

Proof of Lemma 13. We start by proving conditional independence. If σ is a permutation of J fixing the sets $(J \setminus I)^-$, I , and $(J \setminus I)^+$, then the restriction of σ to each of these sets is equal to the induced permutation, and

$$\text{inv}(\sigma) = \text{inv}(\sigma_{(J \setminus I)^-}) + \text{inv}(\sigma_I) + \text{inv}(\sigma_{(J \setminus I)^+}).$$

From the form of the Mallows measure, it follows that the restrictions of σ to each of $(J \setminus I)^-$, I , and $(J \setminus I)^+$ are independent Mallows-distributed permutations of the respective intervals.

Further conditioning π_I to equal τ implies that, conditional on $\pi \in \text{Res}(\tau, I, J)$, the random permutations $\pi_{(J \setminus I)^-}$ and $\pi_{(J \setminus I)^+}$ are conditionally independent and Mallows-distributed.

Therefore

$$\begin{aligned} \mathbb{P}(\pi \in \text{Res}(\tau, I, J)) &= q^{\text{inv}(\tau)} \frac{Z(q, \#(J \setminus I)^-) \cdot Z(q, \#(J \setminus I)^+)}{Z(q, \#J)} \\ &= q^{\text{inv}(\tau)} (1 - q)^{\#I} \frac{\prod_{i=1}^a (1 - q^i) \prod_{i=1}^b (1 - q^i)}{\prod_{i=1}^c (1 - q^i)}, \end{aligned}$$

where $a = \#(J \setminus I)^-$ and $b = \#(J \setminus I)^+$ and $c = \#J$. The claimed lower bound on $\mathbb{P}(\pi \in \text{Res}(\tau, I, J))$ now follows from

$$\frac{\prod_{i=1}^a (1 - q^i) \prod_{i=1}^b (1 - q^i)}{\prod_{i=1}^c (1 - q^i)} = \frac{\prod_{i=1}^a (1 - q^i)}{\prod_{i=b+1}^c (1 - q^i)} \geq \prod_{i=1}^a (1 - q^i) \geq \phi(q). \quad \square$$

Having developed the requisite bounds for probabilities involving permutations, we now turn to bounds for entire preference structures. Our aim is to show that the lattice cutpoint bound occurs with positive probability when $\mathcal{P} = \text{MalPref}_{q, \mathbb{Z}}$. For this, we introduce a cutoff parameter N , show that there is a positive (but tiny) probability that

$$\text{offset}(\mathcal{P}((i, \varphi^{\circ}), j)) = 0$$

whenever $|i - s| + |j - s| < N$, and show that there is a large conditional probability given this event that the bound is satisfied for all i, j with $|i - s| + |j - s| \geq N$.

Let N be a positive integer. Let I be a finite interval of integers, and let $s \in \mathbb{Z} + \frac{1}{2}$ be a half-integer. Denote by $\mathcal{B}_I(s, N)$ the set of preference structures \mathcal{P} on Ω_I such that, for all $i \in I$ with $|i - s| < N$ and for both $\varphi^{\circ} \in \{\sigma^{\circ}, \varphi^{\circ}\}$,

$$\mathcal{P}((i, \varphi^{\circ})) \in \text{Fix}(\{j \in I : |j - s| + |i - s| < N\}, I).$$

Observe that if $\mathcal{P} \in \mathcal{B}_I(s, N)$, then \mathcal{P} satisfies the lattice cutpoint bound (19) for all $i, j \in I$ with $|i - s| + |j - s| < N$, since $\text{offset}(\mathcal{P}((i, \varphi^{\circ}), j)) = 0$ for all such i and j .

We now show that ρ_q , the probability that 0 is an essential lattice cutpoint of $\text{MalPref}_{q, \mathbb{Z}}$, is positive for all $q \in (0, 1)$.

Lemma 14. *For all $q \in (0, 1)$, for all integers $N \geq 1$, and for all half-integers $s \in \mathbb{Z} + \frac{1}{2}$, the probability that $\text{MalPref}_{q, \mathbb{Z}}$ satisfies the lattice cutpoint bound (19) at s is at least*

$$(1 - q)^{8N^2} \phi(q)^{4N} \left(1 - \left(\frac{4q^{N/40}}{1 - q^{1/20}} \right)^2 \right).$$

In particular, this expression is a lower bound for ρ_q .

Proof. Fix q , N , and let $\mathcal{P} = \text{MalPref}_{q,\mathbb{Z}}$. For all integers $n \geq 1$, let $\mathcal{A}_{\llbracket -n, n \rrbracket}$ be the event that $\mathcal{P}_{\llbracket -n, n \rrbracket}$ satisfies the lattice cutpoint bound (19) at $s = \frac{1}{2}$. By Lemma 8 (ii), the lattice cutpoint bound is inherited by induced permutations. Thus the events $\mathcal{A}_{\llbracket -n, n \rrbracket}$ are decreasing, and if they all occur then by Lemma 9, the preference structure \mathcal{P} has an essential lattice cutpoint at $\lfloor s \rfloor$. Consequently,

$$\rho_q \geq \inf_{n \geq 1} \mathbb{P}(\mathcal{A}_{\llbracket -n, n \rrbracket}). \quad (29)$$

We proceed to find a uniform lower bound for $\mathbb{P}(\mathcal{A}_{\llbracket -n, n \rrbracket})$.

The preferences $\mathcal{P}((i, \varphi^\sigma))$ for $(i, \varphi^\sigma) \in \Omega_{\mathbb{Z}}$ are independent. Therefore

$$\mathbb{P}(\mathcal{P}_{\llbracket -n, n \rrbracket} \in \mathcal{B}_{\llbracket -n, n \rrbracket}(s, N)) = \prod_{\substack{i \in A_n, \\ \varphi^\sigma \in \{\sigma^\circ, \varphi\}}} \mathbb{P}(\mathcal{P}_{\llbracket -n, n \rrbracket}((i, \varphi^\sigma)) \in \text{Fix}(I_{n,i}, \llbracket -n, n \rrbracket)), \quad (30)$$

where $A_n = \{i \in \llbracket -n, n \rrbracket : |i - s| < N\}$, $I_{n,i} = \{j \in \llbracket -n, n \rrbracket : |j - s| + |i - s| < N\}$, and we take n sufficiently large such that A_n is non-empty. By Lemma 13,

$$\mathbb{P}(\mathcal{P}_{\llbracket -n, n \rrbracket}((i, \varphi^\sigma)) \in \text{Fix}(I_{n,i}, \llbracket -n, n \rrbracket)) \geq (1 - q)^{\#I_{n,i}} \phi(q). \quad (31)$$

Substitute (31) into (30) and use $\#A_n \leq 2N$ and $\sum_{i \in A_n} \#I_{n,i} \leq 4N^2$ to obtain that

$$\begin{aligned} \mathbb{P}(\mathcal{P}_{\llbracket -n, n \rrbracket} \in \mathcal{B}_{\llbracket -n, n \rrbracket}(s, N)) &\geq \left(\prod_{i \in A_n} (1 - q)^{\#I_{n,i}} \phi(q) \right)^2 \\ &\geq (1 - q)^{8N^2} \phi(q)^{4N}. \end{aligned} \quad (32)$$

The conditional probability that $\mathcal{A}_{\llbracket -n, n \rrbracket}$ does *not* occur given that $\mathcal{P}_{\llbracket -n, n \rrbracket}$ belongs to $\mathcal{B}_{\llbracket -n, n \rrbracket}(s, N)$ is bounded from above by

$$\sum_{\varphi^\sigma \in \{\sigma^\circ, \varphi\}} \sum_{\substack{i, j \in \llbracket -n, n \rrbracket : \\ |i - s| + |j - s| \geq N}} \mathbb{P}\left(\text{offset}(\mathcal{P}_{\llbracket -n, n \rrbracket}((i, \varphi^\sigma)), j) \geq \frac{|i - s| + |j - s|}{20}\right). \quad (33)$$

By Lemma 13, the conditional law of $\mathcal{P}_{\llbracket -n, n \rrbracket}$ given that $\mathcal{P}_{\llbracket -n, n \rrbracket} \in \mathcal{B}_{\llbracket -n, n \rrbracket}(s, N)$ is such that for all $(i, \varphi^\sigma) \in \Omega_{\llbracket -n, n \rrbracket}$ the restrictions of $\mathcal{P}_{\llbracket -n, n \rrbracket}((i, \varphi^\sigma))$ to $(\llbracket -n, n \rrbracket \setminus I_n)^\pm$ are conditionally Mallows-distributed. Thus (23) in Lemma 12 applies, so (33) is bounded from above by

$$2 \sum_{\substack{i, j \in \llbracket -n, n \rrbracket : \\ |i - s| + |j - s| \geq N}} 2q^{\frac{|i - s| + |j - s|}{20}} \leq 4 \left(\frac{2q^{N/40}}{1 - q^{1/20}} \right)^2,$$

and therefore

$$\mathbb{P}\left(\mathcal{A}_{[-n,n]} \mid \mathcal{P}_{[-n,n]} \in \mathcal{B}_{[-n,n]}(s, N)\right) \geq 1 - \left(\frac{4q^{N/40}}{1 - q^{1/20}}\right)^2.$$

The result follows by combining this with (32) and substituting into (29). \square

Proof of Proposition 10. That $\rho_q > 0$ for all $q \in (0, 1)$ follows by taking $N = N_q$ sufficiently large in Lemma 14. Taking $N = 1$ and $q \rightarrow 0$ in the lemma yields that $\lim_{q \rightarrow 0} \rho_q = 1$. \square

Having established the first of the two positive probabilities promised in this section, we turn to the second. Our aim is to show that there is a positive probability that the random preference structure $\text{LocPref}(\text{MalPref}_{q,\mathbb{Z}}, 0)$ induced by a neighborhood of the origin has multiple stable matchings. By the celebrated Gale-Shapley algorithm [8], $\text{LocPref}(\text{MalPref}_{q,\mathbb{Z}}, 0)$ always has at least one stable matching. That is,

$$\text{gr}(\text{MalPref}_{q,\mathbb{Z}}, 0) \geq 0. \quad (34)$$

Consider the unique preference structure \mathcal{Q} on $\Omega_{\mathbb{Z}}$ satisfying the following properties:

- individuals in $\Omega_{\mathbb{Z}} \setminus \Omega_2$ rank everyone in order, and
- individuals in Ω_2 rank individuals in $\Omega_{\mathbb{Z}} \setminus \Omega_2$ in order, and
- the induced preference structure \mathcal{T} on Ω_2 is given by

$$\begin{aligned} \mathcal{T}((1, \varphi)) &= 12, & \mathcal{T}((2, \varphi)) &= 21, \\ \mathcal{T}((1, \sigma)) &= 21, & \mathcal{T}((2, \sigma)) &= 12. \end{aligned}$$

In other words, preferences in \mathcal{Q} are in order except that $(2, \varphi)$ prefers $(1, \sigma)$ to $(2, \sigma)$ and $(1, \sigma)$ prefers $(1, \varphi)$ to $(2, \varphi)$. Observe that for all $m \geq 2$, the induced preference structure $\mathcal{Q}_{[-m,m]}$ has exactly 2 stable matchings: the in-order matching $\bigcup_{i \in [-m,m]} \{(i, \varphi), (i, \sigma)\}$ and

$$\{(1, \varphi), (2, \sigma)\} \cup \{(2, \varphi), (1, \sigma)\} \cup \bigcup_{i \in [-m,m] \setminus \{1,2\}} \{(i, \varphi), (i, \sigma)\}.$$

The proof of Proposition 11 is quite similar to the proof of Proposition 10, since both amount to obtaining lower bounds on the probability that the lattice cutpoint bound (or a close analogue thereof) is satisfied. In light of this, we have presented certain steps of the proof of Proposition 11 in a complete yet concise manner, with the understanding that a lengthier exposition of the steps in question appears in the proof of Proposition 10.

We recall the following notation for a preference structure \mathcal{P} on $\Omega_{\mathbb{Z}}$.

- $\text{Nbhd}(\mathcal{P}, 0)$ is the interval whose endpoints are the two essential lattice cutpoints surrounding 0 (inclusive for the right endpoint and exclusive for the left).
- $\text{LocPref}(\mathcal{P}, 0)$ is the preference structure induced by $\text{Nbhd}(\mathcal{P}, 0)$.
- $\text{gr}(\mathcal{P}, 0)$ is defined via

$$\text{gr}(\mathcal{P}, 0) = \frac{\log \#\text{StabMatch LocPref}(\mathcal{P}, 0)}{\#\text{Nbhd}(\mathcal{P}, 0)}$$

if $\text{Nbhd}(\mathcal{P}, 0)$ is finite and $\text{gr}(\mathcal{P}, 0) = 0$ otherwise.

For an integer $i \in \mathbb{Z}$ and a non-empty set $S \subseteq \mathbb{Z}$, the **distance** from i to S is

$$d(i, S) := \min_{j \in S} |i - j|.$$

Proof of Proposition 11. Fix $q \in (0, 1)$ and let $\mathcal{P} = \text{MalPref}_{q, \mathbb{Z}}$ be a random Mallows-distributed preference structure on $\Omega_{\mathbb{Z}}$. By the previous proposition $\rho > 0$, and thus by Lemma 4 the set of essential lattice cutpoints has positive density. Let I be the a.s. finite interval whose endpoints are the essential lattice cutpoints of \mathcal{P} surrounding the origin, exclusive on the left and inclusive on the right. We show that $\text{gr}(\mathcal{P}, 0) > 0$ with positive probability for all $q \in (0, 1)$, which is equivalent showing that \mathcal{P}_I has multiple stable matchings with positive probability.

Choose $m \geq 2$ sufficiently large such that the induced matching $\mathcal{Q}_{[-m, m]}$ satisfies the cutpoint bound (8) at $\pm m$ (for the sake of concreteness, $m = 50$ suffices). Let \mathcal{E} be the

event that $\mathcal{P}_{\llbracket -m, m \rrbracket} = \mathcal{Q}_{\llbracket -m, m \rrbracket}$ and that for every finite interval of integers J containing $\llbracket -m, m \rrbracket$, the induced preference structure \mathcal{P}_J satisfies the cutpoint bound at $\pm m$ as well.

We claim that $\mathbb{P}(\mathcal{E}) > 0$. More precisely, we show that for all integers $N \geq 1$,

$$\mathbb{P}(\mathcal{E}) \geq q^2(1-q)^{8(N+m)^2} \phi(q)^{4(N+m)} \left(1 - \left(\frac{4q^{N/40}}{1-q^{1/20}} \right)^2 \right), \quad (35)$$

by a calculation similar to the proof of Lemma 14. This is positive for $N = N_q$ large enough.

For all integers $n \geq m$, let $\mathcal{A}'_{\llbracket -n, n \rrbracket}$ be the event that $\mathcal{P}_{\llbracket -n, n \rrbracket}$ satisfies the lattice cutpoint bound (19) at $\pm m$ and that $\mathcal{P}_{\llbracket -m, m \rrbracket} = \mathcal{Q}_{\llbracket -m, m \rrbracket}$ (c.f. $\mathcal{A}_{\llbracket -n, n \rrbracket}$ in the proof of Lemma 14.)

For all integers $n \geq m$, let $A'_n = \{i \in \llbracket -n, n \rrbracket : d(i, \llbracket -m, m \rrbracket) < N\}$. For all $i \in A'_n$, let

$$I'_{n,i} = \{j \in \llbracket -n, n \rrbracket : d(i, \llbracket -m, m \rrbracket) + d(j, \llbracket -m, m \rrbracket) < N\}$$

(c.f. the sets A_n and $I_{n,i}$ defined in the proof of Lemma 14).

Let $\mathcal{B}'_{\llbracket -n, n \rrbracket}$ denote the set of preference structures \mathcal{R} on $\Omega_{\llbracket -n, n \rrbracket}$ such that, for all $i \in A_n$ and for both $\varphi \in \{\sigma, \varphi\}$,

$$\mathcal{R}((i, \varphi)) \in \text{Res}\left(\mathcal{Q}_{\llbracket -n, n \rrbracket}((i, \varphi)), I_{n,i}, \llbracket -n, n \rrbracket\right).$$

By independence of the preferences $\mathcal{P}((i, \varphi))$ and Lemma 13,

$$\mathbb{P}(\mathcal{P}_{\llbracket -n, n \rrbracket} \in \mathcal{B}'_{\llbracket -n, n \rrbracket}) \geq \left(q \prod_{i \in A'_n} (1-q)^{\#I'_{n,i}} \phi(q) \right)^2.$$

(There is an extra factor of q , when compared with the calculation in Lemma 14, since the restricted preference lists are taken to have a total of one inversion for each gender.) Since $\#A'_n \leq 2(N+m)$ and $\sum_{i \in A'_n} \#I'_{n,i} \leq 4(N+m)^2$,

$$\mathbb{P}(\mathcal{P}_{\llbracket -n, n \rrbracket} \in \mathcal{B}'_{\llbracket -n, n \rrbracket}) \geq q^2(1-q)^{8(N+m)^2} \phi(q)^{4(N+m)}. \quad (36)$$

As in the proof of Lemma 14, it follows from Lemma 13 that the conditional probability that $\mathcal{A}'_{\llbracket -n, n \rrbracket}$ does *not* occur given that $\mathcal{P}_{\llbracket -n, n \rrbracket}$ belongs to $\mathcal{B}'_{\llbracket -n, n \rrbracket}$ is bounded from above by

$$2 \sum_{\substack{i, j \in \llbracket -n, n \rrbracket : \\ d(i, \llbracket -m, m \rrbracket) + d(j, \llbracket -m, m \rrbracket) \geq N}} 2q^{\frac{d(i, \llbracket -m, m \rrbracket) + d(j, \llbracket -m, m \rrbracket)}{20}} \leq 4 \left(\frac{2q^{N/40}}{1-q^{1/20}} \right)^2,$$

and therefore

$$\mathbb{P}\left(\mathcal{A}'_{[-n,n]} \mid \mathcal{P}_{[-n,n]} \in \mathcal{B}'_{[-n,n]}\right) \geq 1 - 4\left(\frac{2q^{N/40}}{1 - q^{1/20}}\right)^2.$$

Combining this with (36) yields (35). Now taking $N = N_q$ sufficiently large yields that $\mathbb{P}(\mathcal{E}) > 0$, as claimed.

On \mathcal{E} , the lattice cutpoint bound holds at $\pm m$ and therefore every stable matching of \mathcal{P}_I restricts to a stable (perfect) matching of $\mathcal{P}_{[-m,m]}$, yielding

$$\#\text{StabMatch } \mathcal{P}_I \geq \#\text{StabMatch } \mathcal{P}_{[-m,m]} = \#\text{StabMatch } \mathcal{Q}_{[-m,m]} = 2.$$

Consequently,

$$\mathbb{P}(\#\text{StabMatch } \mathcal{P}_I \geq 2) \geq \mathbb{P}(\mathcal{E}) > 0. \quad (37)$$

Finally, we deduce from (37) that $\text{gr}(\mathcal{P}, 0) > 0$ with positive probability. Indeed, suppose to the contrary that $\text{gr}(\mathcal{P}, 0) = 0$ a.s. Then by stationarity, $\text{gr}(\mathcal{P}, i) = 0$ a.s. for all $i \in \mathbb{Z}$ as well, whereupon it holds with probability 1 that

$$\log \#\text{StabMatch } \mathcal{P}_I \leq \#I \sum_{i \in [-n,n]} \text{gr}(\mathcal{P}, i) = 0.$$

This contradicts (37), and thus $\text{gr}(\mathcal{P}, 0) > 0$ with positive probability. \square

3.5 Proof of main theorem

In this section we prove the main theorem, which states that for all $q \in (0, 1)$, the number of stable matchings of $\text{MalPref}_{q,n}$ grows exponentially in n , and the growth rate tends to 0 as $q \rightarrow 0$. Recall the quantity $\text{gr}(\text{MalPref}_{q,\mathbb{Z}}, 0)$, previously introduced as a local approximation of the growth rate. In previous sections, we used the ergodic theorem to show that the exponential growth rate of the number of stable matchings is given by

$$\text{gr}_q = \mathbb{E} \text{gr}(\text{MalPref}_{q,\mathbb{Z}}, 0),$$

for those values of $q \in (0, 1)$ for which $\text{gr}(\text{MalPref}_{q,\mathbb{Z}}, 0)$ is integrable. Specifically, this follows by combining the results of propositions 5, 10, and 11. To complete the proof of the main

theorem, it remains to show that $\text{gr}(\text{MalPref}_{q,\mathbb{Z}}, 0)$ is integrable for all $q \in (0, 1)$, and that its mean tends to 0 as $q \rightarrow 0$. Throughout this section, functions of q on $(0, 1)$ tending to 0 as $q \rightarrow 0$ are denoted by $o_q(1)$.

Proposition 15. *For all $q \in (0, 1)$ we have that $\text{gr}(\text{MalPref}_{q,\mathbb{Z}}, 0)$ is integrable. Furthermore, $\text{gr}_q = \mathbb{E} \text{gr}(\text{MalPref}_{q,\mathbb{Z}}, 0)$ tends to 0 as $q \rightarrow 0$.*

We prove this by obtaining tail bounds on the distance from the origin to the closest essential lattice cutpoints. Since we already know that the set of essential lattice cutpoints has positive density by Proposition 14, the only way for this distance to have heavy tails is if the essential lattice cutpoints are highly correlated.

Working with essential lattice cutpoints directly is challenging, so as in the previous section we instead consider the set of locations at which the lattice cutpoint bound (19) holds. This is a subset of the set of essential lattice cutpoints, and along the way to proving Proposition 14 we showed that it, too, has positive density. Here we establish that for fixed locations s and t , the covariance between the event that the lattice cutpoint bound holds at s and that it holds at t decays exponentially in $|s - t|$. This is a straightforward (if technical) computation, since all relevant probabilities in the Mallows model decay exponentially in the separation distance. Summing this bound over all pairs of locations yields a variance bound, and we use the second moment method to conclude.

Lemma 16. *Let $I \subseteq J$ be finite intervals of integers and let π be a random permutation of J with law $\text{Mal}_{q,J}$, for $q \in [0, 1)$. Then for all $j \in I$,*

$$\mathbb{P}(\text{offset}(\pi, j) > \text{offset}(\pi_I, j)) \leq (c_1 + o_q(1))q^{c_2 d(j, \mathbb{Z} \setminus I)},$$

where $c_1, c_2 > 0$ are absolute constants and $o_q(1)$ tends to 0 as $q \rightarrow 0$.

Proof. If $\text{offset}(\pi, j) > \text{offset}(\pi_I, j)$, then either $L_{+-}(\pi, j) > L_{+-}(\pi_I, j)$ or $L_{-+}(\pi, j) > L_{-+}(\pi_I, j)$. Suppose that the former occurs. By definition of L_{+-} , it follows that there is

some $i \in J \setminus I$ with $i > j$ and $\pi(i) < \pi(j)$. Thus by the union bound and Lemma 8 (iii),

$$\begin{aligned} \mathbb{P}(L_{+-}(\pi, j) > L_{+-}(\pi_I, j)) &\leq \sum_{\substack{i \in J \setminus I: \\ i > j}} \mathbb{P}(\pi(i) < \pi(j)) \\ &\leq \sum_{i \in (J \setminus I)^+} 4q^{i-j} = \frac{4q^{\max I + 1 - j}}{1 - q}. \end{aligned} \quad (38)$$

By a similar argument,

$$\mathbb{P}(L_{-+}(\pi, j) > L_{-+}(\pi_I, j)) \leq \frac{4q^{j - \min I - 1}}{1 - q}. \quad (39)$$

The result now follows by (38), (39), and the union bound. \square

Fix $q \in (0, 1)$. For a finite interval $I \subset \mathbb{Z}$ and a preference structure \mathcal{P} on Ω_I , let $\text{Bnd}(\mathcal{P}) \subseteq \mathbb{Z} + \frac{1}{2}$ denote the set of half-integers at which the lattice cutpoint bound (19) holds for \mathcal{P} . If \mathcal{P} is a preference structure on $\Omega_{\mathbb{Z}}$, then by definition of the lattice cutpoint bound we have that

$$\text{Bnd}(\mathcal{P}) = \bigcap_{\substack{i, j \in \mathbb{Z}: \\ i \leq j}} \text{Bnd}(\mathcal{P}_{\llbracket i, j \rrbracket}).$$

Lemma 17. *For all $q \in (0, 1)$, all $s \in \mathbb{Z} + \frac{1}{2}$, and all finite intervals $I \subset \mathbb{Z}$,*

$$\mathbb{P}\left(s \in \text{Bnd}((\text{MalPref}_{q, \mathbb{Z}})_I) \setminus \text{Bnd}(\text{MalPref}_{q, \mathbb{Z}})\right) \leq \#I \cdot (c_1 + o_q(1)) \cdot q^{c_2 d(s, \mathbb{Z} \setminus I)},$$

where $c_1, c_2 > 0$ are absolute constants and $o_q(1)$ tends to 0 as $q \rightarrow 0$.

Proof. Fix q , I , and let $\mathcal{P} = \text{MalPref}_{q, \mathbb{Z}}$. If $s \notin \text{Bnd}(\mathcal{P})$, then there exist $i, j \in \mathbb{Z}$ and $\varphi \in \{\sigma^\circ, \varphi\}$ such that

$$\text{offset}\left(\mathcal{P}((i, \varphi), j)\right) \geq \frac{|i - s| + |j - s|}{20}. \quad (40)$$

Let φ , i , and j be a triple satisfying (40). Either

(i) $i \in \mathbb{Z} \setminus I$ or $|j - s| > d(s, \mathbb{Z} \setminus I)/2$, or

(ii) $i \in I$ and $|j - s| \leq d(s, \mathbb{Z} \setminus I)/2$.

By the union bound and (23) from Lemma 12, we have that

$$\mathbb{P}(\exists i \in \mathbb{Z} \setminus I, j \in \mathbb{Z}, \varphi^\natural \in \{\sigma^\natural, \varphi^\natural\} \text{ satisfying (40)}) \leq 2 \sum_{\substack{i \in \mathbb{Z} \setminus I \\ j \in \mathbb{Z}}} 2q^{\frac{|i-s|+|j-s|}{20}}.$$

Similarly, the probability that there exist $i \in \mathbb{Z}$, $j \in \mathbb{Z}$, and $\varphi^\natural \in \{\sigma^\natural, \varphi^\natural\}$ satisfying $|j-s| > d(s, \mathbb{Z} \setminus I)/2$ and (40) is bounded from above by

$$2 \sum_{\substack{i, j \in \mathbb{Z}: \\ |j-s| > d(s, \mathbb{Z} \setminus I)/2}} 2q^{\frac{|i-s|+|j-s|}{20}} = (c_3 + o_q(1))q^{c_4 d(s, \mathbb{Z} \setminus I)},$$

for some absolute constants $c_3, c_4 > 0$. Adding these two bounds together yields an upper bound on the probability that some i , j , and φ^\natural satisfy (i) and (40).

Finally, we bound the probability that $s \in \text{Bnd}(\mathcal{P}_I)$ and that both (ii) and (40) are satisfied for some φ^\natural , i , and j . If all these conditions are satisfied, then

$$\text{offset}\left(\mathcal{P}((i, \varphi^\natural), j)\right) \geq \frac{|i-s| + |j-s|}{20} > \text{offset}\left(\mathcal{P}_I((i, \varphi^\natural), j)\right), \quad (41)$$

and $|j-s| \leq d(s, \mathbb{Z} \setminus I)/2$. By Lemma 16, the probability that (41) holds for given φ^\natural , i , and j , is exponentially small in $d(j, \mathbb{Z} \setminus I)$. Taking a union bound implies that the probability of (ii), (40), and $s \in \text{Bnd}(\mathcal{P}_I)$ being simultaneously satisfied for some φ^\natural , i , and j , is at most

$$\#I \cdot (c_5 + o_q(1)) \cdot q^{c_6 d(s, \mathbb{Z} \setminus I)},$$

for some absolute constants $c_5, c_6 > 0$. The result follows by combining this with case (i). \square

Lemma 18. *For all $q \in (0, 1)$ and for all half-integers $s, t \in \mathbb{Z} + \frac{1}{2}$,*

$$\text{Cov}\left(\mathbf{1}[s \in \text{Bnd}(\text{MalPref}_{q, \mathbb{Z}})], \mathbf{1}[t \in \text{Bnd}(\text{MalPref}_{q, \mathbb{Z}})]\right) \leq (c_1 + o_q(1)) \sqrt{|s-t|} q^{c_2 |s-t|}, \quad (42)$$

where $c_1, c_2 > 0$ are absolute constants and $o_q(1)$ tends to 0 as $q \rightarrow 0$.

Proof. Fix q , s , and t , and let $\mathcal{P} = \text{MalPref}_{q, \mathbb{Z}}$. Define the neighborhoods

$$I_s = \left\{j \in \mathbb{Z}: |j-s| < \frac{|s-t|}{2}\right\} \quad \text{and} \quad I_t = \left\{j \in \mathbb{Z}: |j-t| < \frac{|s-t|}{2}\right\},$$

and note that I_s and I_t are disjoint and contain s and t , respectively. For $r \in \{s, t\}$, let $X_r = \mathbf{1}[r \in \text{Bnd}(\mathcal{P})]$ and let $Y_r = \mathbf{1}[r \in \text{Bnd}(\mathcal{P}_{I_r})]$. The random variables Y_s and Y_t are independent, since they are defined using preferences of disjoint sets of individuals. Thus $\text{Cov}(Y_s, Y_t) = 0$, and therefore

$$\begin{aligned} \text{Cov}(X_s, X_t) &= \text{Cov}(X_s, X_t) - \text{Cov}(Y_s, Y_t) \\ &\leq |\text{Cov}(Y_s - X_s, X_t)| + |\text{Cov}(Y_s, Y_t - X_t)|. \end{aligned} \quad (43)$$

Since $\text{Var} X_t \leq 1$ and $\text{Var} Y_s \leq 1$, we have by Cauchy-Schwarz that

$$|\text{Cov}(Y_s - X_s, X_t)| + |\text{Cov}(Y_s, Y_t - X_t)| \leq \sqrt{\text{Var}(Y_s - X_s)} + \sqrt{\text{Var}(Y_t - X_t)}. \quad (44)$$

For $r \in \{s, t\}$, the random variable $Y_r - X_r$ is supported in $\{0, 1\}$ and has mean given by

$$\mathbb{E}(Y_r - X_r) = \mathbb{P}(r \in \text{Bnd}(\mathcal{P}_{I_r}) \setminus \text{Bnd}(\mathcal{P})).$$

When combined with (43) and (44), this implies that

$$\text{Cov}(X_s, X_t) \leq \sqrt{\mathbb{P}(s \in \text{Bnd}(\mathcal{P}_{I_s}) \setminus \text{Bnd}(\mathcal{P}))} + \sqrt{\mathbb{P}(t \in \text{Bnd}(\mathcal{P}_{I_t}) \setminus \text{Bnd}(\mathcal{P}))}. \quad (45)$$

The result now follows by substituting the bound from Lemma 17 into (45). \square

Using the covariance bound from the previous lemma, we establish Proposition 15, which states that the random variable $\text{gr}(\text{MalPref}_{q, \mathbb{Z}}, 0)$ is integrable for all $q \in (0, 1)$, and that its mean gr_q tends to 0 as $q \rightarrow 0$.

Proof of Proposition 15. Fix $q \in (0, 1)$. Let $\mathcal{P} = \text{MalPref}_{q, \mathbb{Z}}$. Recall from (5) and (34) that

$$0 \leq \text{gr}(\mathcal{P}, 0) \leq \log \# \text{Nbhd}(\mathcal{P}, 0).$$

We will show that $\log \# \text{Nbhd}(\mathcal{P}, 0)$ is integrable, and a key ingredient in the argument is the second moment method. The form we will use consists of the inequality

$$\mathbb{P}(X = 0) \leq \frac{\text{Var}(X)}{(\mathbb{E}X)^2}, \quad (46)$$

valid for any non-negative random variable X with finite second moment that is not a.s. 0.

Consider $X_n = \#(\text{Bnd}(\mathcal{P}) \cap [0, n])$ for $n \geq 1$. By bilinearity of the variance, we have

$$\text{Var } X_n = \sum_{s, t \in [0, n] \cap (\mathbb{Z} + \frac{1}{2})} \text{Cov}(\mathbf{1}[s \in \text{Bnd}(\mathcal{P})], \mathbf{1}[t \in \text{Bnd}(\mathcal{P})]). \quad (47)$$

By Lemma 18, there is a function f on $(0, 1)$ satisfying

$$f(q) = \frac{c_1 + o_q(1)}{\log(1/q)},$$

such that for all $s, t \in [0, n] \cap (\mathbb{Z} + \frac{1}{2})$ with $|s - t| > f(q) \log n$ we have that

$$\left| \text{Cov}(\mathbf{1}[s \in \text{Bnd}(\mathcal{P})], \mathbf{1}[t \in \text{Bnd}(\mathcal{P})]) \right| \leq \frac{1}{n^2}.$$

Here $c_1 > 0$ denotes an absolute constant. In the following paragraphs, we will also use other absolute constants c_2, c_3, c_4 , and c_5 , each of which is greater than 0.

Thus, the total sum of the terms in (47) with $|s - t| > f(q) \log n$ is bounded as $n \rightarrow \infty$. As there are fewer than $2f(q)n \log n$ terms with $|s - t| \leq f(q) \log n$, each of which is bounded in absolute value, we have that

$$\text{Var } X_n \leq c_2 f(q) n \log n. \quad (48)$$

By stationarity of $\text{Bnd}(\mathcal{P})$, we have that

$$\mathbb{E} X_n = \mathbb{E} \#(\text{Bnd}(\mathcal{P}) \cap [0, n]) = n \mathbb{P}(\frac{1}{2} \in \text{Bnd}(\mathcal{P})),$$

and by Lemma 14 this quantity is positive for all $q \in (0, 1)$ and satisfies

$$\mathbb{E} X_n = n(1 - o_q(1)), \quad q \rightarrow 0. \quad (49)$$

Substituting (48) and (49) into the second moment inequality (46) for X_n yields that

$$\mathbb{P}(\text{Bnd}(\mathcal{P}) \cap [0, n] = \emptyset) \leq \frac{c_3 + o_q(1)}{\log(1/q)} \cdot \frac{\log n}{n},$$

and similarly for $\mathbb{P}(\text{Bnd}(\mathcal{P}) \cap [-n, 0] = \emptyset)$. Thus, we have that

$$\mathbb{P}(\text{gr}(\mathcal{P}, 0) > n) \leq \mathbb{P}(\log \# \text{Nbhd}(\mathcal{P}, 0) > n) \leq \frac{c_4 + o_q(1)}{\log(1/q)} \cdot n e^{-n}. \quad (50)$$

By the tail sum formula for expectation, (50) implies that $\text{gr}(\mathcal{P}, 0)$ is integrable for all $q \in (0, 1)$, and its mean gr_q satisfies

$$\text{gr}_q \leq \frac{c_5 + o_q(1)}{\log(1/q)}.$$

Consequently $\text{gr}_q \rightarrow 0$ as $q \rightarrow 0$. □

Proof of Theorem 1. Combine Propositions 5, 10, 11, and 15. □

Chapter 4

THE MALLOWS PERMUTATION AS A STABLE MATCHING

In this section we show that the Mallows random permutation can be interpreted as the solution of a stable matching problem with random incompatibility. This is part of joint work with Omer Angel, Alexander E. Holroyd, and Tom Hutchcroft.

4.1 *The finite case*

Suppose that there are n heterosexual men and n heterosexual women who seek to be married. Both the men and the women are ranked objectively in order of attractiveness, where 1 is the most attractive and n is the least. Unfortunately, each man-woman pair has a probability $q \in [0, 1)$ of being incompatible, independently of all other pairs, meaning that neither will consider the other for marriage under any circumstances. Attractiveness and compatibility are the only factors affecting preferences, so a woman's first choice for marriage is given by the most attractive man she is compatible with, her second choice is the second most attractive man she is compatible with, and so on.

The graph with vertex set given by the men and women and with edges between compatible individuals bears the following description. Write $p = 1 - q$ and let $K_{n,n}(p)$ be the random bipartite graph with vertex set $\llbracket 1, n \rrbracket \times \{\sigma, \varphi\}$ in which each edge $\{(i, \sigma), (j, \varphi)\}$ is present with probability p , independently of all other edges. In other words, $K_{n,n}(p)$ is the graph obtained by performing Bernoulli bond percolation with parameter p on the complete bipartite graph $K_{n,n}$. This graph encodes compatibility between individuals.

A **matching** of a graph is a set of edges, no two of which share a common vertex. A matching is **perfect** if every vertex is adjacent to an edge of the matching. Permutations of $\llbracket 1, n \rrbracket$ are in bijection with perfect matchings of $K_{n,n}$, and we identify $\sigma \in S_n$ with the

matching containing each of the edges $\{(i, \sigma), (\sigma(i), \varphi)\}$ for $i = 1, \dots, n$. A matching of a subgraph G of $K_{n,n}$ is **stable** (with respect to G) if there does not exist an edge $\{(i, \sigma), (j, \varphi)\}$ of G for which there exist $i' > i$ and $j' > j$ such that $\{(i, \sigma), (j', \varphi)\}$ and $\{(i', \sigma), (j, \varphi)\}$ both belong to the matching.

Recall that the **Mallows distribution** Mal_q on permutations of $\llbracket 1, n \rrbracket$ with parameter $q \in [0, 1]$ is the probability distribution that assigns to each permutation $\sigma \in S_n$ a probability proportional to $q^{\text{inv}(\sigma)}$, where

$$\text{inv}(\sigma) = \#\{(i, j) \in \llbracket 1, n \rrbracket^2 : i < j \text{ but } \sigma(i) > \sigma(j)\}.$$

A **Mallows random permutation** is a random permutation with law Mal_q .

Theorem 1. *Let $p \in (0, 1]$, let $q = 1 - p$, and let n be a positive integer. Every subgraph of $K_{n,n}$ has a unique stable matching. The unique stable matching of the random subgraph $K_{n,n}(p)$ is perfect with probability $\prod_{k=1}^n (1 - q^k)$. Conditional on the event that the unique stable matching of $K_{n,n}(p)$ is perfect, it is Mal_q -distributed.*

The limit as $n \rightarrow \infty$ of the probabilities appearing in Theorem 1, namely $\prod_{k=1}^{\infty} (1 - q^k)$, is positive for all $q \in [0, 1)$. Furthermore, it satisfies the asymptotics

$$\prod_{k=1}^{\infty} (1 - q^k) = \sqrt{\frac{2\pi}{\log(1/q)}} \exp\left(\frac{\pi^2}{6 \log q} - \frac{\log q}{24}\right) + o(1)$$

as $q \rightarrow 1$ [27]. In other words, for all $p \in (0, 1]$, the probability that the unique stable matching of $K_{n,n}(p)$ is perfect is bounded away from zero.

Proof of Theorem 1. Fix a subgraph G of $K_{n,n}$. In any stable matching of G , the most attractive man must be married to the most attractive women with whom he is compatible. Inductively, the i^{th} most attractive man must be married to the most attractive women with whom he is compatible and who is not married to a man more attractive than him. This shows that the stable matching is unique, and gives us an algorithm to compute it. Formally, we set $\min \emptyset = \infty$ and define $\sigma: \llbracket 1, n \rrbracket \rightarrow \llbracket 1, n \rrbracket \cup \{\infty\}$ recursively by setting

$$\sigma(1) = \min\{1 \leq k \leq n : k \text{ is compatible with } 1\}$$

and, for all $1 < i \leq n$,

$$\sigma(i) = \min(\{k: k \text{ is compatible with } i\} \setminus \{\sigma(j): 1 \leq j < i, \sigma(j) \neq \infty\}).$$

It follows by induction on i that σ is the unique stable matching of G , where $\sigma(i) = \infty$ means that i is left unmatched.

Now take $G = K_{n,n}(p)$ to be random. Then the probability that $\sigma(i) \neq \infty$ given $\sigma(1), \dots, \sigma(i-1)$ is equal to the probability that i is compatible with some k in the set

$$[[1, n] \setminus \{\sigma(1), \sigma(2), \dots, \sigma(i-1)\}.$$

This set has cardinality $n - i + 1$ on the event that $\sigma(j) \neq \infty$ for all $1 \leq j < i$, and thus

$$\mathbb{P}(\sigma(i) \neq \infty \mid \sigma(1) \neq \infty, \dots, \sigma(i-1) \neq \infty) = 1 - q^{n-i+1}.$$

Hence, the probability that the unique stable matching is perfect is given by

$$\mathbb{P}(\sigma(1) \neq \infty, \dots, \sigma(n) \neq \infty) = \prod_{i=1}^n (1 - q^{n-i+1}) = \prod_{i=1}^n (1 - q^i).$$

Fix a permutation $\tau \in S_n$. It remains to show that the probability that $\sigma = \tau$ equals a constant multiple of $q^{\text{inv}(\tau)}$. By the recursive formula for σ given above, we have that for all $1 \leq i \leq n$,

$$\mathbb{P}(\sigma(i) = \tau(i) \mid \sigma(1) = \tau(1), \dots, \sigma(i-1) = \tau(i-1)) = q^{\#\{j < \tau(i): \tau^{-1}(j) > i\}} (1 - q).$$

Taking the product of these conditional probabilities and observing that

$$\sum_{j=1}^n \#\{j < \tau(i): \tau^{-1}(j) > i\} = \text{inv}(\tau)$$

yields that $\mathbb{P}(\sigma = \tau) = q^{\text{inv}(\tau)}(1 - q)^n$. □

4.2 Infinite intervals bounded from below

There is a natural extension of the Mallows distribution from finite to infinite intervals of \mathbb{Z} . We use the following notation. An **interval of \mathbb{Z}** is a set of the form $I \cap \mathbb{Z}$ where I is

an interval of real numbers. A **permutation** of an interval I of \mathbb{Z} is a bijection from I to itself. For a permutation σ of I and an integer $i \in I$, we define the code of σ at i to be

$$\mathcal{C}(\sigma)_i := \#\{j < \sigma(i) : \sigma^{-1}(j) > i\}.$$

We write $K_{I,I}$ for the complete bipartite graph with vertex set $I \times \{\sigma, \varphi\}$ and $K_{I,I}(p)$ for the graph obtained by performing Bernoulli bond percolation with parameter p on $K_{I,I}$.

Let I be an infinite interval of \mathbb{Z} that is bounded from below. We define the Mallows distribution Mal_q^I on permutations of I with parameter $q \in [0, 1)$ to be the unique measure such that, if σ is a random permutation of I with law Mal_q^I , we have that $(\mathcal{C}(\sigma)_i)_{i \in I}$ is an iid sequence of geometric random variables with parameter q , i.e.,

$$\mathbb{P}(\mathcal{C}(\sigma)_i \geq k) = q^k, \quad i \in I, k \geq 0.$$

Existence and uniqueness of such a measure was established in [11].

Theorem 2. *Let $p \in (0, 1]$, let $q = 1 - p$, and let $I \subset \mathbb{Z}$ be an infinite interval that is bounded from below. Every subgraph of $K_{I,I}$ has a unique stable matching. The unique stable matching $K_{I,I}(p)$ is a.s. perfect, and is distributed as the Mallows permutation of I with parameter q .*

Proof. Assume, without loss of generality, that $I = [1, \infty) \cap \mathbb{Z}$. Uniqueness of the stable matching $\sigma : I \rightarrow I \cup \{\infty\}$ is established exactly as in the proof of Theorem 1, and it is given recursively by

$$\sigma(1) = \min\{k \in I : k \text{ is compatible with } 1\}$$

and, for all $1 < i < \infty$,

$$\sigma(i) = \min(\{k \in I : k \text{ is compatible with } i\} \setminus \{\sigma(j) : 1 \leq j < i, \sigma(j) \neq \infty\}). \quad (1)$$

Since the set $I \setminus \{\sigma(1), \dots, \sigma(i-1)\}$ is infinite for all $i \geq 1$, the minimum in (1) is taken over a set which is a.s. non-empty. Thus the unique stable matching is a.s. perfect.

Finally, we verify that $(\mathcal{C}(\sigma)_i)_{i \in I}$ is an iid sequence of geometric random variables with parameter q . Observe that $\sigma(i)$ is the $\mathcal{C}(\sigma)_i^{\text{th}}$ -smallest element of the set

$$I \setminus \{\sigma(1), \dots, \sigma(i-1)\},$$

from which it follows that the conditional law of $\mathcal{C}(\sigma)_i$ given $\sigma(1), \dots, \sigma(i-1)$ is the geometric distribution with parameter q . Thus $\mathcal{C}(\sigma)_i$ is geometric with parameter q and independent of $\sigma(1), \dots, \sigma(i-1)$. The result now follows since $\mathcal{C}(\sigma)_1, \dots, \mathcal{C}(\sigma)_{i-1}$ all belong to the sigma-algebra generated by these random variables. \square

4.3 Stable matching of the integers

Gnedin and Olshansky extended the Mallows distribution to a measure $\text{Mal}_q^{\mathbb{Z}}$ on permutations of \mathbb{Z} in [10], which we now consider in relation with stable matching. Both permutations of \mathbb{Z} , and stable matchings of $K_{\mathbb{Z}, \mathbb{Z}}$, can exhibit pathologies not present in the previously considered cases. This complicates matters and also leads to interesting new phenomena.

In forthcoming joint work (not included here), we will establish the following result.

Theorem 3. *Let $q \in [0, 1)$ and $p = 1 - q$. There is a coupling of $K_{\mathbb{Z}, \mathbb{Z}}(p)$ and a $\text{Mal}_q^{\mathbb{Z}}$ -distributed random permutation σ of \mathbb{Z} such that σ is a perfect matching of $K_{\mathbb{Z}, \mathbb{Z}}(p)$ a.s.*

While this result plays an important motivational role in the remainder of this section, we make no use of its conclusion. Thus, the remainder of the section is entirely self-contained.

As we will now show, for $p \in (0, 1)$ the random graph $K_{\mathbb{Z}, \mathbb{Z}}(p)$ a.s. has uncountably many stable matchings. In fact, $\text{Mal}_q^{\mathbb{Z}}$ -distributed matchings of \mathbb{Z} have several key features that are not shared with the vast majority of stable matchings of $K_{\mathbb{Z}, \mathbb{Z}}(p)$. We colloquially refer to these other matchings as being ‘wild’. There is a simple and explicit algorithm capable of producing uncountably many wild matchings, which we present below.

A matching m of $K_{\mathbb{Z}, \mathbb{Z}}$ is identified with a function from $\mathbb{Z} \times \{\varnothing\}$ to $(\mathbb{Z} \times \{\sigma\}) \cup \{\infty\}$, where a value of ∞ indicates an unmatched individual. For $i \in \mathbb{Z}$, we define the quantities

$$L_+(m, i + \frac{1}{2}) = \#\left\{j, k \in \mathbb{Z}: j < i + \frac{1}{2} \text{ and } k > i + \frac{1}{2} \text{ and } m((j, \varnothing)) = (k, \sigma)\right\}$$

and

$$L_-(m, i + \frac{1}{2}) = \#\left\{j, k \in \mathbb{Z}: j > i + \frac{1}{2} \text{ and } k < i + \frac{1}{2} \text{ and } m((j, \varphi)) = (k, \sigma)\right\}$$

The matching m is **locally finite** if $L_+(m, i + \frac{1}{2})$ and $L_-(m, i + \frac{1}{2})$ are both finite for any (and hence every) $i \in \mathbb{Z}$. If m is perfect and locally finite, we define the **flow** of m to be

$$L_+(m, i + \frac{1}{2}) - L_-(m, i + \frac{1}{2}),$$

which is easily seen to be independent of i . We say that m is **balanced** if it is perfect, locally finite, and has flow zero.

As shown by Gneden and Olshansky, the measure $\text{Mal}_q^{\mathbb{Z}}$ constructed in [10] is supported on balanced permutations of \mathbb{Z} . In fact, these authors characterized $\text{Mal}_q^{\mathbb{Z}}$ as the unique measure on balanced permutations whose pushforward under the map

$$\sigma \mapsto (\mathcal{C}(\sigma)_i)_{i \in \mathbb{Z}}$$

is the law of a sequence of iid geometric random variables with parameter q .

Theorem 4. *Let $p \in (0, 1)$. There a.s. exists both a stable matching of $K_{\mathbb{Z}, \mathbb{Z}}(p)$ which is perfect but not locally finite, and one which is neither perfect nor locally finite. In fact, there are a.s. uncountably many of each.*

Recall our convention that smaller integers are more highly preferred. More precisely, if integers i, j, k are such that (i, φ) is compatible with (j, σ) and (k, σ) and if $j < k$, then (i, φ) prefers (j, σ) to (k, σ) .

Proof. For a finite set $A \subset \mathbb{Z}$ and for an integer $i \in \mathbb{Z} \setminus A$, let $\Omega_{A, i}$ be the event that both of the sets

$$\{j \in \mathbb{Z}: (j, \varphi) \text{ is compatible with } (i, \sigma) \text{ but not } (k, \sigma) \text{ for any } k \in A\}$$

and

$$\{j \in \mathbb{Z}: (j, \sigma) \text{ is compatible with } (i, \varphi) \text{ but not } (k, \varphi) \text{ for any } k \in A\}$$

are unbounded from below. Clearly $\mathbb{P}(\Omega_{A,i}) = 1$ for all A and i . Since there are only a countable number of finite subsets of \mathbb{Z} , the event

$$\Omega = \bigcap \{\Omega_{A,i} : A \text{ finite subset of } \mathbb{Z}, i \in \mathbb{Z} \setminus A\}$$

also occurs with probability 1.

We now show that the conclusions of the theorem hold on Ω . Let $(x_n)_{n \geq 0}$ be an enumeration of $\mathbb{Z} \times \{\sigma, \varphi\}$, and let $(a_j)_{j \geq 1}$ be a sequence of positive integers. We iteratively construct a perfect stable matching of $K_{\mathbb{Z}, \mathbb{Z}}(p)$ as follows. At each step of the construction, we choose the next individual in the enumeration, and if they are not already matched, we find a compatible partner for them who is incompatible with everyone previously matched.

More precisely, at each step $n \geq 0$ of the construction, we maintain the following objects.

- Matched_n , $\text{Matched}_n(\varphi)$ and $\text{Matched}_n(\sigma)$ will be subsets of $\mathbb{Z} \times \{\varphi, \sigma\}$, $\mathbb{Z} \times \{\varphi\}$, and $\mathbb{Z} \times \{\sigma\}$, respectively, satisfying

$$\text{Matched}_n = \text{Matched}_n(\varphi) \cup \text{Matched}_n(\sigma) \quad \text{and} \quad \text{Matched}_n \subseteq \{x_m : m \leq n\}.$$

- $\sigma_n : \text{Matched}_n(\varphi) \rightarrow \text{Matched}_n(\sigma) \cup \{\infty\}$ will be a stable matching of Matched_n .
- There will be a positive integer $j_n \geq 1$.

Initialize these quantities when $n = 0$ by setting σ_0 to be the empty matching, $j_0 = 1$, and $\text{Matched}_0 = \emptyset$.

Given σ_n and j_n for some $n \geq 0$, we define σ_{n+1} and j_{n+1} as follows.

- If $x_{n+1} \in \text{Matched}_n$, we set $\text{Matched}_{n+1} = \text{Matched}_n$ and $\sigma_{n+1} = \sigma_n$ and $j_{n+1} = j_n$. (That is, we do nothing.)
- If $x_{n+1} \in \mathbb{Z} \times \{\varphi\} \setminus \text{Matched}_n(\varphi)$, we set $j_{n+1} = j_n + 1$, set $\sigma_{n+1}(x_{n+1})$ to be the $a_{j_{n+1}}^{\text{th}}$ largest element of

$$\{k \leq 0 : (k, \sigma) \text{ is compatible with } x_{n+1} \text{ but not with any element of } \text{Matched}_n(\varphi)\}$$

(on Ω , there exists such an element), and set

$$\begin{aligned}\text{Matched}_{n+1}(\varphi) &= \text{Matched}_n(\varphi) \cup \{(\sigma_{n+1}(x_{n+1}))\} \\ \text{Matched}_{n+1}(\sigma) &= \text{Matched}_n(\sigma) \cup \{x_{n+1}\}.\end{aligned}$$

- Similarly, if $x_{n+1} \in \mathbb{Z} \times \{\sigma\} \setminus \text{Matched}_n(\sigma)$, we set $j_{n+1} = j_n + 1$, set $\sigma_{n+1}(x_{n+1})$ to be the $a_{j_{n+1}}^{\text{th}}$ largest element of

$$\{k \leq 0: (k, \varphi) \text{ is compatible with } x_{n+1} \text{ but not with any element of } \text{Matched}_n(\sigma)\}$$

(on Ω , there exists such an element), and set

$$\begin{aligned}\text{Matched}_{n+1}(\sigma) &= \text{Matched}_n(\sigma) \cup \{(\sigma_{n+1}(x_{n+1}))\} \\ \text{Matched}_{n+1}(\varphi) &= \text{Matched}_n(\varphi) \cup \{x_{n+1}\}.\end{aligned}$$

It is easily seen by induction that σ_n is a perfect stable matching of Matched_n for every $n \geq 1$. We can therefore define a perfect stable matching of $K_{\mathbb{Z}, \mathbb{Z}}(p)$ by setting $\sigma(x_n) = \sigma_n(x_n)$ for every $n \geq 1$.

The construction ensures that for every $k > 0$, the partner of (k, φ) belongs to $\{\dots, -1, 0\} \times \{\sigma\}$, so that σ does not have finite flow.

If we fix the numeration $(x_n)_{n \geq 1}$, then any two distinct sequences of positive integers $(a_n)_{n \geq 1}$ and $(a'_n)_{n \geq 1}$ must yield distinct matchings. Indeed, at the first place the sequences differ, the edge added to the matchings will differ as well. This proves that on Ω , there are uncountably many stable matchings of $K_{\mathbb{Z}, \mathbb{Z}}(p)$ that are perfect but not locally finite.

That there are uncountably many stable matchings of $K_{\mathbb{Z}, \mathbb{Z}}(p)$ that are neither perfect nor locally finite follows by taking $(x_n)_{n \geq 1}$ to be an enumeration of

$$(\{\dots, -1, 0\} \times \{\varphi\}) \cup (\mathbb{Z} \times \{\sigma\})$$

in the previous argument. This yields an uncountable family of stable matchings of $K_{\mathbb{Z}, \mathbb{Z}}(p)$ such that no (k, φ) with $k > 0$ is matched. \square

BIBLIOGRAPHY

- [1] Riddhipratim Basu and Nayantara Bhatnagar, *Limit theorems for longest monotone subsequences in random mallows permutations*, arXiv preprint arXiv:1601.02003 (2016).
- [2] Itai Benjamini, Noam Berger, Christopher Hoffman, and Elchanan Mossel, *Mixing times of the biased card shuffling and the asymmetric exclusion process*, Transactions of the American Mathematical Society **357** (2005), no. 8, 3013–3029.
- [3] Nayantara Bhatnagar and Ron Peled, *Lengths of monotone subsequences in a Mallows permutation*, Probab. Theory Related Fields **161** (2015), no. 3-4, 719–780. MR3334280
- [4] Charles Blair, *The lattice structure of the set of stable matchings with multiple partners*, Mathematics of operations research **13** (1988), no. 4, 619–628.
- [5] Ralph Allan Bradley and Milton E Terry, *Rank analysis of incomplete block designs: I. the method of paired comparisons*, Biometrika **39** (1952), no. 3/4, 324–345.
- [6] Mark Braverman and Elchanan Mossel, *Sorting from noisy information*, CoRR **arXiv preprint arXiv:0910.1191** (2009).
- [7] Persi Diaconis, Arun Ram, et al., *Analysis of systematic scan metropolis algorithms using iwahori-hecke algebra techniques*, Department of Statistics, Stanford University, 2000.
- [8] David Gale and Lloyd S Shapley, *College admissions and the stability of marriage*, The American Mathematical Monthly **69** (1962), no. 1, 9–15.
- [9] Alexey Gladkikh and Ron Peled, *On the Cycle Structure of Mallows Permutations*, arXiv preprint arXiv:1601.06991 (2016).
- [10] Alexander Gnedin and Grigori Olshanski, *The two-sided infinite extension of the Mallows model for random permutations*, Advances in Applied Mathematics **48** (2012), no. 5, 615–639.
- [11] Alexander Gnedin, Grigori Olshanski, et al., *q -exchangeability via quasi-invariance*, The Annals of Probability **38** (2010), no. 6, 2103–2135.
- [12] Dan Gusfield and Robert W Irving, *The stable marriage problem: structure and algorithms*, MIT press, 1989.

- [13] Alexander E. Holroyd and Thomas M. Liggett, *Finitely dependent coloring*.
- [14] ———, *Symmetric 1-dependent colorings of the integers*, Electron. Commun. Probab. **20** (2015), no. 31, 8. MR3327870
- [15] Olav Kallenberg, *Foundations of modern probability*, Second, Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR1876169 (2002m:60002)
- [16] Maurice George Kendall, *Rank correlation methods*. (1948).
- [17] Craig Lennon and Boris Pittel, *On the likely number of solutions for the stable marriage problem*, Combinatorics, Probability and Computing **18** (2009), no. 03, 371–421.
- [18] Avi Levy, *Finitely dependent insertion processes*, Electronic Journal of Probability **22** (2017).
- [19] C. L. Mallows, *Non-null ranking models. I*, Biometrika **44** (1957), 114–130. MR0087267
- [20] Carl Mueller and Shannon Starr, *The length of the longest increasing subsequence of a random Mallows permutation*, Journal of Theoretical Probability **26** (2013), no. 2, 514–540.
- [21] Boris Pittel, *The average number of stable matchings*, SIAM Journal on Discrete Mathematics **2** (1989), no. 4, 530–549.
- [22] ———, *On likely solutions of a stable marriage problem*, The Annals of Applied Probability (1992), 358–401.
- [23] Alvin E Roth and Elliott Peranson, *The effects of the change in the nrmp matching algorithm*, JAMA **278** (1997), no. 9, 729–732.
- [24] Richard P Stanley, *Enumerative Combinatorics. vol. 1, vol. 49 of Cambridge Studies in Advanced Mathematics*, 2nd ed., Cambridge University Press, Cambridge, 1997.
- [25] S. Starr and M. Walters, *Phase Uniqueness for the Mallows Measure on Permutations*, ArXiv e-prints (February 2015), available at 1502.03727.
- [26] Shannon Starr, *Thermodynamic limit for the Mallows model on S_n* , Journal of mathematical physics **50** (2009), no. 9.
- [27] George N Watson, *The final problem: an account of the mock theta functions*, Journal of the London Mathematical Society **1** (1936), no. 1, 55–80.