The Dissertation Committee for James Thomas Murphy III
certifies that this is the approved version of the following dissertation:

Coupling and Unimodularity in Stationary Settings

Committee:

François Baccelli, Supervisor

Lewis Bowen

Lorenzo Sadun

Sanjay Shakkottai

Coupling and Unimodularity in Stationary Settings

by

James Thomas Murphy III

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2019

Dedicated to my parents, Jim and Reta, who taught me
the preparedness to sieze opportunity when it comes,
the discipline to create it when it does not, and
the compassion to share it with others.

Acknowledgments

First and foremost, I thank my advisor, François Baccelli, for his instrumental guidance throughout my doctoral journey. His incredible ability to teach and his enthusiasm to learn about my work were inspirational. There are few things as motivating as one who shows true interest in hearing more. I also thank Abishek, Antonio, Deepanshu, Eliza, Jae Oh, Mayank, and Natasa for their questions and criticism during the earliest stages of my research. Also thanks to Mir-Omid Haji-Mirsadeghi, whose collaborative contributions were vital to the present work. Further thanks go to the many people who asked questions during my presentations at SPA and StochNet 2018, to the anonymous reviewers of my papers, to those who volunteered to read early versions of this thesis, and to my doctoral committee.

I also thank the math and educational psychology graduate student communities at UT Austin for maintaining a collaborative and friendly environment. Thanks in particular to Andrew, Gill, Logan, Max, Sean, and Tom, and to Pierce, Trent, Savannah, and Selim. Thanks to Elisa and Eva for their work keeping the math department running smoothly. Thanks to Bryce, Cannon, Jordan, Katelin, Logan, Moh, and Ren for believing in me. Thanks to Ana for being a great friend. Thanks to my sister Laura and to my parents again for being incredibly supportive.

Thanks to Bill Hrusa and Steve Sheve for helping me open the next chapter in my professional life, and Gautam Iyer, Giovanni Leoni, Po-Shen Loh, John Mackey, and the rest of my professors at Carnegie Mellon for challenging and pushing me to learn the skills that are now invaluable to my career.

Thanks to MIT and Stanford for providing high-quality educational resources for free online. I have used these resources extensively through high school, college, and graduate school. The commitment to making education easy to access has not gone unnoticed.

Thanks to the free and open source community for providing computing infrastructure. My research, simulations, pet projects, and more depend on too many open source projects to count. Thanks in particular to every individual that contributes to the Linux kernel, to the Debian project, to the Electronic Frontier Foundation, and to the Tor project.

The contents of Chapter 3 have been published [Mur17]. The contents of Chapter 4 and Chapter 5 have been submitted for publication. Additionally, François Baccelli and Mir-Omid Haji-Mirsadeghi are coauthors of Chapter 4. This thesis was supported by a grant of the Simons Foundation (#197982 to The University of Texas at Austin).

Coupling and Unimodularity in Stationary Settings

Publication No. _____

James Thomas Murphy III, Ph.D.
The University of Texas at Austin, 2019

Supervisor: François Baccelli

This dissertation studies three applications of the tools of coupling and unimodularity in stationary settings. The first application is to exact coupling of random walks. Conditions for admitting a successful exact coupling are given that are necessary and in the Abelian case also sufficient. This solves a problem posed by H. Thorisson [Tho11]. The second application is centered on the random graph generated by a Doeblin-type coupling of discrete time processes whereby when two paths meet, they merge. This random graph is studied through a novel subgraph, called a bridge graph, generated by paths started in a fixed state. The bridge graph is then made into a unimodular network. The final application focuses on point-shifts of point processes on topological groups. Foliations and connected components generated by point-shifts are studied, and the cardinality classification of connected components is generalized to unimodular groups.

Table of Contents

Chapter 1

Introduction

The heart of this thesis is primarily comprised of three papers in probability theory. The papers, which make up the contents of Chapters 3 to 5, may be read in isolation, but there is value in taking them as a whole. The common thread between the papers is the set of mathematical tools used to create them. Broadly speaking, those tools may be boiled down to coupling, unimodularity, and stationarity. This thesis may be seen as an example of how these tools work together. This introduction will give the essence of each of the three papers and their applicability. Chapter 2 then covers the requisite preliminaries for the rest of the document.

Exact Coupling of Random Walks on Polish Groups

Chapter 3 deals with exact coupling of random walks. In general, a successful exact coupling between two processes (or their distributions) gives a specific way of defining copies of the processes on the same probability space in such a way that almost surely the two processes eventually merge, i.e., that they meet in finite time and move in unison for all time thereafter. One of the primary motivations for constructing an exact coupling of two processes is to obtain bounds on the total variation distance

between the processes after some number of steps. The intuition behind this is that, because copies of the processes can be made to merge almost surely, in which case their tail behaviors are literally identical, their distributions after a certain number of steps must be getting closer together. Thinking of the processes as experiments or simulations, this means that the initial conditions become irrelevant to an outcome observed a long time after the start of the experiment. Moreover, a specific exact coupling can be analyzed to give bounds on the rate at which the total variation distance between the two processes decays.

The first attributable instance of exact coupling is the seminal 1938 work of Doeblin [Doe38]. Doeblin considered the simplest kind of exact coupling, where two irreducible and aperiodic Markov chains on a finite state space are run independently until they meet. However, Doeblin's coupling fails when applied to more general chains. An exact coupling is not "successful" if it does not ensure the two processes meet with probability one. This failure occurs with Doeblin's coupling, e.g., for differently started random walks on $\mathbb{Z}$. In 1968, Ornstein [Orn69] introduced a new coupling that is successful for such random walks on $\mathbb{Z}$. Ornstein's idea was to start with independent walks, but to have them share a common step-length if the difference in their step-lengths were to be larger than some threshold. This ensures that the walks cannot separate from each other too quickly, which, under an aperiodicity assumption, is enough to show the coupling is successful. Eventually, exact coupling of random walks not living on a

lattice were studied. Instead of using aperiodicity conditions to guarantee the existence of a successful exact coupling between random walks started at any two starting points, the question shifted to asking from which pairs of starting points can a successful exact coupling be constructed. This is the subject of Chapter 3.

In Chapter 3, the processes in question are taken to be random walks that jump in an i.i.d. fashion according to a common step-length distribution, but that are started at differing locations. The space on which the walks occur is allowed to be a topological group that is nice enough (a Polish group), though the results are new even on $\mathbb{R}$. This work was inspired by H. Thorisson's paper [Tho11] and Ö. Arnaldsson's thesis [Arn10], which delineated the continuous (or, technically, "spread out") and discrete cases on $\mathbb{R}$. The underlying technique of the general case, which encapsulates both the former cases, is a small extension of what was required in [Arn10]. For the discrete case, Arnaldsson demonstrated in [Arn10] that a successful exact coupling of two random walks started at different locations can be constructed if and only if independent copies of the walks have positive probability of meeting after a finite number of steps. This condition is rephrased in terms of being able to fit a nontrivial distribution underneath different shifts of the random walks viewed after some number of steps. The rephrasing allows for generalizing the answer in the discrete case to all distributions on $\mathbb{R}$, as well as many other topological groups.

The main result of Chapter 3 is to give a condition that is necessary

and, in the Abelian case, also sufficient to guarantee the existence of a successful exact coupling. When the condition is met, a specific exact coupling is given, which leads to a bound on the total variation distance between the two walks after n steps by $O(1/\sqrt{n})$ and, in some special cases, by $O(\rho^n)$ for some $\rho \in (0, 1)$. The coupling itself is similar to the one presented in [Arn10]. Ultimately, regardless of the group on which the random walks live, a trick allows one to turn the difference process between the random walks (or something resembling a difference process in the non-Abelian case) into a random walk on a cyclic group, reducing the entire study to a question of return times of lazy symmetric random walks on $\mathbb{Z}$ or $\mathbb{Z}/d\mathbb{Z}$.

Using the criterion established to determine from which starting points a successful exact coupling with a walk started at the identity can be constructed, the previously known continuous and discrete cases are shown as consequences. Then the set of such good starting points, here referred to as the successful exact coupling set, is studied in its own right. In particular, the successful exact coupling set is shown to be a Borel measurable group. Finally, the weaker notion of possible exact coupling, whereby two random walks are only made to merge with positive probability instead of almost surely, is studied. It is shown in the Abelian case that the ostensibly weaker condition of admitting a possible exact coupling is, in fact, equivalent to admitting a successful exact coupling. The chapter ends with a non-Abelian example where possible and successful exact coupling are not the same notions.

Doeblin Trees

Chapter 4 is an extension of the vast amount of literature on what is now commonly referred to as coupling from the past (CFTP). CFTP traditionally refers to either a mathematical technique or a corresponding algorithm to obtain samples from the stationary distribution of an irreducible, aperiodic, and positive recurrent Markov chain. The mathematical theory, which applies even in a non-Markovian setting, was initially developed by Borovkov and Foss in [BF92], and the algorithmic viewpoint focusing on the Markov case is due to Propp and Wilson in [PW96].

The CFTP algorithm, in certain cases, offers a much better alternative to the naive method of running the Markov chain for a very long time in order to sample from its steady state behavior. The naive method often fails to perform because it is unknown how long "a very long time" should be. For example, if there are 2^{64} states in the system, as is the case when the state space is all possible spin configurations (up/down) of particles aligned in an 8×8 grid, it is unclear whether one million time steps or even one billion time steps would be sufficient. CFTP is more tenable in that certain models can be simulated in significantly less time and space than with the naive algorithm, and, importantly, when the algorithm terminates, one can be sure that the result is indeed a sample from the true stationary distribution. The spin system example was given in the original Propp and Wilson paper [PW96] as one of the systems for which CFTP may be advantageously applied. The property that CFTP returns a sample from the stationary distribution itself

and not from a distribution that is just close to the stationary distribution is what the literature commonly refers to as a "perfect" sample.

When analyzing the CFTP algorithm, one envisions different copies of the Markov chain started in all possible states and at all possible times defined in such a way that when two paths meet, they merge. The trajectories of all these copies of the Markov chain may then be interpreted as a random graph, which is a tree if it has all trajectories eventually coalescing[1]. For this introduction, assume that the random graph in question is a tree. While an impressive amount of literature studies the CFTP algorithm and how it can be extended to more general settings or specialized for specific Markov chains, there has been no explicit treatment or study of the random graph of trajectories, coined a Doeblin graph, in its own right. To the best of the author's knowledge, Chapter 4 is the first such treatment. In Chapter 4, one takes a step back from the CFTP algorithm and obtaining samples from the stationary distribution, and instead one focuses on properties of the Doeblin graph itself. Identifying global properties of the Doeblin graph gives a hands-on kind of geometric insight into why the CFTP algorithm works.

The novel part of the study of Doeblin graphs is a particular kind of subgraph, called a bridge graph, which consists of all trajectories starting

[1]In this case, the setup actually gives many examples of the sort of exact coupling studied in Chapter 3 (any two trajectories started at the same time eventually merge), but Markov chains are used instead of random walks. However, this viewpoint is not elaborated upon in Chapter 4.

at a particular anchor state. The primary result, from which essentially all the other results in this chapter follow, is that, under certain conditions, bridge graphs are unimodularizable. An explicit way of picking a root turns a bridge graph into a unimodular network in the sense of Aldous and Lyons [AL07]. This is used to prove the existence of a special kind of bi-infinite path in the Doeblin graph. The special path is the almost surely unique path in the Doeblin graph that is recurrent both forwards and backwards in time for any (and hence every) state. The existence of this so-called bi-recurrent path in a bridge graph is leveraged in many ways to discern properties of the larger Doeblin graph, and of Markov chains in general.

It is shown that any Markov chain indexed by $\mathbb{Z}$ can be seen as living inside some Doeblin graph. Hence properties of Markov chains in general can be obtained from the theory of Doeblin graphs. For example, a Markov chain that is indexed by $\mathbb{Z}$ and that has an irreducible, aperiodic, and positive recurrent transition matrix is stationary if and only if its trajectories are almost surely bi-recurrent, cf. Theorem 4.2.11. Another example is that, if its transition matrix converges in a uniform sense to its stationary distribution, any Markov chain that is indexed by $\mathbb{Z}$ must be stationary and bi-recurrent. In other words, a bi-infinite Markov chain that mixes uniformly is automatically stationary. This is a partial converse to the well-known fact that a stationary Markov chain indexed by $\mathbb{N}$ can be extended to a stationary Markov chain indexed by $\mathbb{Z}$. Embedding Markov chains into Doeblin graphs

may also be used to give new proofs of well-known properties of Markov chains. In particular, the classical cycle formulas for relating expected return times of a Markov chain to the stationary distribution and relating the average number of visits to a state before returning to another state to the ratio of the stationary distribution on those states are recovered using Doeblin graphs.

Additional questions about Doeblin graphs studied include the following: When is the bi-recurrent path the only bi-infinite path in the Doeblin graph? What is the relationship between bridge graphs using different anchor states? Vertical slices of a bridge graph themselves form a Markov chain, what are its properties? Can the unimodular version of a bridge graph be approximated in the sense of local weak convergence by finite windows? How can the mass-transport principle be leveraged in other ways beyond giving the existence of the bi-recurrent path?

Finally, the flexibility that one has in picking the transition structure that generates the Doeblin graph allows the class of Doeblin graphs to contain a rich set of examples to call upon for study, and which can have major differences in certain qualitative aspects, such as having only one versus infinitely many bi-infinite paths. Additionally, the family of unimodular versions of bridge graphs is a concrete class of examples of unimodular networks that the author hopes will prove useful to check intuition, ideas, and perhaps conjectures against throughout the further development of the theory of unimodular networks in general.

Point-shifts of Point Processes on Topological Groups

Chapter 5 deals with point-shifts of stationary point processes. Point-shifts map points of a point process to other points of the process in a flow-adapted manner. That is, if the whole process were shifted by some amount, then a point-shift must map a shifted point to the shift of where the original point would be mapped to. This condition means that a point-shift must act using only information about the local view from the perspective of the point it is mapping. It has no concept of an origin with which to refer as an absolute anchor point, and two points that have the same local view must be mapped in the same manner. This philosophy imposed upon point processes and point-shifts captures the idea that no point or location in space is special in an absolute sense. Stationary point processes are most often defined on $\mathbb{R}^d$. However, stationarity may be considered with respect to the action of any group. Last's [Las10a, Las10b] are good references on this topic, and the framework for stationary point processes on groups used in this chapter is based on these references.

Chapter 5 studies the behavior and structure of point-shifts of stationary point processes that are stationary with respect to groups. The usefulness of studying such point-shifts is somewhat indirect. Indeed, point-shifts of stationary point processes are in many ways analogous to so-called vertex-shifts of random networks. Vertex-shifts map points of a network to other points of the network in a manner respecting isomorphisms. The advantage of using random networks is that they are more widely applica-

ble and are not limited to living on a particular space like $\mathbb{R}^d$. They also come with a natural notion of convergence that makes it easy to formalize statements like "an $n \times n$ grid converges to $\mathbb{Z}^2$ as $n \to \infty$". Stationary point processes, in a certain sense, act like unimodular networks, i.e., networks where, heuristically, the root has been selected "uniformly at random". In fact, under its Palm measure, a stationary point processes on $\mathbb{R}^d$ can be directly seen as an embedding of a unimodular network, cf. Theorem 5.4.4 and Proposition 5.4.6. However, there is a separation between stationary point processes and unimodular networks when the underlying group with respect to which the process is stationary is itself not unimodular (in the group-theoretic sense of being unimodular). Studying this boundary and relationship between stationarity, unimodularity in the sense of groups, and unimodularity in the sense of random networks is a stepping stone towards a better understanding of random networks. This is the main reason for studying point-shifts of point processes on groups other than $\mathbb{R}^d$, but occasionally it may also be useful to take advantage of specific structure of the underlying group that may be lost in the up-to-isomorphisms framework of random networks.

The primary result of Chapter 5 is an extension of the classification of components of the graph drawn by a point-shift on a stationary point process. This result was originally proved on $\mathbb{R}^d$ by Baccelli and Haji-Mirsadeghi in [BHM18]. The result splits all components into three classes that have drastically different qualitative structure according to conditions that are

often trivial to check. In this chapter, their result is extended verbatim to unimodular groups, and proofs using some properties about the order of $\mathbb{R}$ have been replaced by mass-transport arguments. This is possible because stationary point processes satisfy a kind of mass-transport theorem that, when the underlying group is unimodular, reduces to the standard formula:

"expected mass in" = "expected mass out".

The classification theorem has since been extended to unimodular networks [BHMK18], though, for point-shifts of point processes on non-unimodular groups, the classification theorem fails to hold. This is one of the reasons preventing a general stationary point processes from being seen under its Palm measure as an embedding of a unimodular network, as is possible on $\mathbb{R}^d$. The question of for which groups *can* every stationary point processes be seen as an embedding of a unimodular network is also studied in this chapter. It remains conjectured that all unimodular groups have this property. To better understand the divide between stationary point processes and unimodular networks, more general theory about point-shifts on possibly non-unimodular groups is given that studies generalizations of Mecke's invariance theorem on bijective point-shifts preserving the Palm measure, running a point-shift backwards in time, and separating points of a point process by a function.

Chapter 2

Preliminaries

These sections serve to lay out some of the basic ideas and notation that are shared across multiple chapters.

2.1. Coupling

This section is suggested as a preliminary for Chapters 3 and 4. Standard references on coupling are [Lin02, Tho00]. Often times one wants to compare two probability distributions $\mathcal{P}_1$ and $\mathcal{P}_2$ on some spaces S_1 and S_2. One way to do so is to find a joint distribution $\mathcal{P}_{1,2}$ on $S_1 \times S_2$ that has $\mathcal{P}_1$ and $\mathcal{P}_2$ as its marginals. Such a distribution is called a **coupling** of $\mathcal{P}_1$ and $\mathcal{P}_2$. Typically, a coupling is constructed by defining random variables $X_1 \sim \mathcal{P}_1$ and $X_2 \sim \mathcal{P}_2$ on a common probability space, so that the distribution of the pair (X_1, X_2) is a coupling of $\mathcal{P}_1$ and $\mathcal{P}_2$. Since there is little chance of confusion, one also commonly refers to the random variable (X_1, X_2) itself, together with the probability space on which it is defined, as a **coupling** of X_1 and X_2, or of $\mathcal{P}_1$ and $\mathcal{P}_2$.

Consider the following example of how coupling can give a useful result. Suppose one can construct a coupling so that $X_1 \leqslant X_2$ a.s., then

one can gather that $\mathbf{P}(X_1 \geqslant t) \leqslant \mathbf{P}(X_2 \geqslant t)$ for all t simply by noting that $\mathbf{E}[1_{\{X_2 \geqslant t\}} - 1_{\{X_1 \geqslant t\}}] \geqslant 0$. Note that in order to even write the expression $\mathbf{E}[1_{\{X_2 \geqslant t\}} - 1_{\{X_1 \geqslant t\}}]$ one has already used that X_1 and X_2 are defined on the same probability space.

When $\mathcal{P}_1$ and $\mathcal{P}_2$ are distributions of sequences, i.e., distributions on $S^{\mathbb{N}}$ or $S^{\mathbb{Z}}$ for some space S, it is often useful to construct a coupling for which the two sequences merge. If (X^1, X^2) is a coupling of two sequences such that $X^1_n = X^2_n$ for all n after a random time T, then one can use the fact that X^1 and X^2 are defined on the same space and the fact that they eventually merge in order to bound the total variation distance between the distributions of X^1_n and X^2_n. This type of coupling is studied at length in Chapter 3.

More generally, the question arises: What can be gleaned about the relationship between two distributions by constructing an interesting coupling, and what are methods to construct interesting couplings? In fact, the question is not limited to a coupling of just two distributions. One can define a single probability space that simultaneously houses couplings of many different distributions. For example, one could define a single probability space on which there is a copy of a given Markov chain started in every possible starting location and view the interactions between all of the different versions of the chain. This idea is taken even further in Chapter 4, where a Markov chain is started not just in every location, but also at every time.

2.2. Point Processes

This section is suggested as a preliminary for Chapter 5 and for a single section (see Section 4.2.3) of Chapter 4. The general theory of point processes and random measures has been developed on any locally compact second-countable Hausdorff (LCSH) space G, cf. [DVJ08]. In this document, only point processes are needed. A **point process** on G is a random element Ψ in the space $\mathbf{M}$ of all locally finite counting ($\mathbb{N} \cup \{\infty\}$-valued) measures on G, where $\mathbf{M}$ is endowed with the cylindrical σ-algebra generated by the mappings $\mu \in \mathbf{M} \mapsto \mu(B)$ for each $B \in \mathcal{B}(G)$. All point processes Ψ in this document are simple, meaning every atom of Ψ has mass 1. In this case, Ψ can and will be identified with its support, which is a random discrete subset of G, allowing for notation such as $X \in \Psi$ to denote that $\Psi(\{X\}) = 1$. When both arguments need to be specified, the notation $\Psi(\omega, B)$ will be preferred over the more cumbersome $\Psi(\omega)(B)$. This notation is also consistent with the view of a point process as being transition kernel from Ω to G.

2.3. Haar Measure

This section is suggested as a preliminary for Chapters 3 to 5. The standard text [Coh13] covers all the basics on Haar measures used in this text. Suppose G is a LCSH topological group. Then, for each $x \in G$ and $B \in \mathcal{B}(G)$, consider the shifted set $xB := \{xb : b \in B\}$. Up to a multiplicative constant, there exists a unique Borel measure λ on G that is finite on compact

sets[1] and satisfies $\lambda(xB) = \lambda(B)$ for all $x \in G$ and $B \in \mathcal{B}(G)$. One refers to λ as a (or, when there is no chance of confusion, "the") **left-invariant Haar measure on** G. The same game can be played with right-invariance and the sets $Bx := \{bx : b \in B\}$ for all $x \in G$ and $B \in \mathcal{B}(G)$ to get a right-invariant Haar measure, but in this document a Haar measure will always refer to a left-invariant Haar measure.

2.4. Random Networks

This section is suggested as a preliminary for Chapters 4 and 5. This section reviews the theory of random networks in the sense of Aldous and Lyons. See [AL07, Khe17] for a more thorough review of random networks than what is provided here. A **network** is a graph $\Gamma = (V(\Gamma), E(\Gamma))$ equipped with a complete separable metric space $(\Xi_\Gamma, d_{\Xi_\Gamma})$ called the **mark space** and two maps from $V(\Gamma)$ and $\{(v, e) : v \in V(\Gamma), e \in E(\Gamma), v \sim e\}$ to Ξ_Γ, where $\sim$ is used for adjacency of vertices or edges. The image of v (resp. (v, e)) in Ξ_Γ is called its **mark**, which is extra information associated to the vertex (resp. edge). The mark of (v, e) may also be thought of as the mark of e considering it to be a directed edge with initial vertex v. The graph distance between v and w is denoted $d_\Gamma(v, w)$. Unless explicitly mentioned otherwise, networks are assumed to be nonempty, locally finite, and connected.

An **isomorphism** between two networks with the same mark space

[1]The unnecessary assumption that G is second-countable is used here to avoid discussion of regularity of measures. Finiteness on compact sets implies regularity on a LCSH space.

is a graph isomorphism that also preserves the marks. A **rooted network** is a pair (Γ, o) in which Γ is a network and o is a distinguished vertex of Γ called the **root**. An **isomorphism** of rooted networks is a network isomorphism that takes the root of one network to the root of the other. Similar definitions apply to doubly rooted networks (Γ, o, v). For convenience, from now on throughout the document, consider only networks with mark space $(\Xi_{\text{univ}}, d_{\Xi_{\text{univ}}})$, where Ξ_{univ} is some fixed uncountable complete separable metric space, such as $[0, 1]$, since all possible mark spaces are homeomorphic to a subset of such a Ξ_{univ}. When constructing examples of networks, the marks will be specified on whatever space is convenient, but the reader should be aware that the marks specified are actually being embedded into Ξ_{univ}. Similarly, to avoid uninformative tedium, only the relevant portion of marks will be specified. The most common example of this used in this document is that directed networks use edge marks to specify which direction(s) the edges point. If, additionally, one wants to mark the edges with weights, one might say, e.g., "the network is marked with i.i.d. weights on the edges" when, in actuality, an edge mark would be the embedding into Ξ_{univ} of a pair (edge direction indicator, weight).

Let $\mathcal{G}$ denote the set of isomorphism classes of nonempty, locally finite, connected networks, and let $\mathcal{G}_*$ (resp. $\mathcal{G}_{**}$) be the set of isomorphism classes of singly (resp. doubly) rooted networks of the same kind. The isomorphism class of a network Γ (resp. (Γ, o), or (Γ, o, v)) is denoted by $[\Gamma]$ (resp. $[\Gamma, o]$ or $[\Gamma, o, v]$).

The sets $\mathcal{G}_*$ and $\mathcal{G}_{**}$ are equipped with natural metrics making them complete separable metric spaces, cf. [AL07]. The distance $d_{\mathcal{G}_*}([\Gamma_1, o_1], [\Gamma_2, o_2])$ between the isomorphism classes of (Γ_1, o_1) and (Γ_2, o_2) is $1/(1+\alpha)$, where α is the supremum of those $r > 0$ such that there is a rooted isomorphism of the neighborhoods $N_{\Gamma_1}(o_1, r), N_{\Gamma_2}(o_2, r)$ of graph-distance $\lfloor r \rfloor$ around the roots of Γ_1, Γ_2 such that each pair of corresponding marks has distance less than $1/r$. The distance on $\mathcal{G}_{**}$ is defined similarly and the projections $[\Gamma, o, v] \mapsto [\Gamma, o]$ and $[\Gamma, o, v] \mapsto [\Gamma, v]$ are continuous. Borel measurable functions on $\mathcal{G}_*$ and $\mathcal{G}_{**}$ are, in a certain sense, those functions that can be determined by looking only at (perhaps larger and larger) finite neighborhoods (including marks) of the root or roots. For example, $[\Gamma, o] \mapsto \deg(o)$ and $[\Gamma, o, v] \mapsto d_\Gamma(o, v)$ are Borel functions. A **random (rooted) network** $[\boldsymbol{\Gamma}, \boldsymbol{o}]$ is a random element in $\mathcal{G}_*$ equipped with its Borel σ-algebra $\mathcal{B}(\mathcal{G}_*)$. Weak convergence of probability measures on $\mathcal{G}_*$ is referred to as **local weak convergence**.

2.5. Unimodularity

This section is suggested as a preliminary for Chapters 4 and 5. Chapter 4 only uses unimodularity of random networks explicitly, and Chapter 5 only uses unimodularity of groups explicitly, but the connections between these two types of unimodularity are important to the cohesiveness of this thesis as a whole.

Unimodularity of Random Networks

A random network $[\mathbf{\Gamma}, \boldsymbol{o}]$ is called **unimodular** if it obeys the following **mass-transport principle**: For all measurable $g : \mathcal{G}_{**} \to \mathbb{R}_{\geqslant 0}$,

$$\mathbf{E} \sum_{v \in V(\mathbf{\Gamma})} g[\mathbf{\Gamma}, \boldsymbol{o}, v] = \mathbf{E} \sum_{v \in V(\mathbf{\Gamma})} g[\mathbf{\Gamma}, v, \boldsymbol{o}]. \tag{2.1}$$

Heuristically, the root of a unimodular network is picked uniformly at random from its vertices. However, since there is no uniform distribution on an infinite set of vertices, the mass-transport principle (2.1) is used in lieu of requiring the root to be picked uniformly at random. One should take care to note that the sums in the previous equation depend only on the isomorphism class $[\mathbf{\Gamma}, \boldsymbol{o}]$ and not which representative is used.

Next, the notions of covariant vertex-shifts, foils, connected components, and the cardinality classification of components of a unimodular network are reviewed. See [BHMK18] for a reference on these concepts. A **(covariant) vertex-shift** is a map Φ which associates to each network Γ a function $\Phi_\Gamma : V(\Gamma) \to V(\Gamma)$ such that Φ commutes with network isomorphisms and the function $[\Gamma, o, v] \to 1_{\{\Phi_\Gamma(o)=v\}}$ is measurable on $\mathcal{G}_{**}$. For a vertex-shift Φ, define two equivalence relations on each network Γ by saying $u, v \in V(\Gamma)$ are in the same Φ**-foil** if $\Phi_\Gamma^n(u) = \Phi_\Gamma^n(v)$ for some $n \in \mathbb{N}$, or in the same Φ**-component** if $\Phi_\Gamma^n(u) = \Phi_\Gamma^m(v)$ for some $n, m \in \mathbb{N}$. Two vertices are in the same Φ-component if their forward orbits under Φ intersect, whereas they are in the same Φ-foil if, after some finite number of applications of Φ, the vertices meet. The Φ**-graph of** Γ is the graph drawn on Γ with vertices

$V(\Gamma)$ and edges from each $v \in V(\Gamma)$ to $\Phi_\Gamma(v)$. The following is a special case of the classification theorem appearing in [BHMK18]. The full version tells even more about the structure of each of the classes.

Theorem 2.5.1 (Foil Classification in Unimodular Networks [BHMK18])**.** *Let* $[\boldsymbol{\Gamma}, \boldsymbol{o}]$ *be a unimodular network and* Φ *a vertex-shift. Almost surely, every vertex has finite degree in the* Φ*-graph of* $\boldsymbol{\Gamma}$*. In addition, each component* C *of the* Φ*-graph of* $\boldsymbol{\Gamma}$ *falls in one of the following three classes:*

(i) ***Class F/F****: C and all its foils are finite, and there is a unique cycle in C.*

(ii) ***Class I/F****: C is infinite but all its foils are finite, there are no cycles in C, and there is a unique bi-infinite path in C.*

(iii) ***Class I/I****: C is infinite and all its foils are infinite, and there are no cycles or bi-infinite paths in C.*

The last tool needed from [BHMK18] is the so-called no infinite/finite inclusion lemma, which is used heavily in the proof of Theorem 2.5.1. To state it, the following definitions are needed. A **covariant subset (of the set of vertices)** is a map C that associates to each network Γ a set $C_\Gamma \subseteq V(\Gamma)$ such that C commutes with network isomorphisms, and such that $[\Gamma, o] \mapsto 1_{\{o \in C_\Gamma\}}$ is measurable. A **covariant (vertex) partition** is a map Π which associates to all networks Γ a partition Π_Γ of $V(\Gamma)$ such that Π commutes with network isomorphisms, and such that the (well-defined) subset $\{[\Gamma, o, v] : v \in \Pi_\Gamma(o)\} \subseteq \mathcal{G}_{**}$ is measurable, where $\Pi_\Gamma(o)$ denotes the partition element in Π_Γ containing o. Then one has the following.

Lemma 2.5.2 (No Infinite/Finite Inclusion [BHMK18])**.** *Let* $[\mathbf{\Gamma}, \boldsymbol{o}]$ *be a unimodular network,* Π *a covariant partition, and* C *a covariant subset. Almost surely, there is no infinite element* E *of* $\Pi_{\mathbf{\Gamma}}$ *such that* $E \cap C_{\mathbf{\Gamma}}$ *is finite and nonempty.*

Unimodularity of Groups

Again let λ be a Haar measure on G. Notice that, for each $x \in G$, the measure $\lambda_x(B) := \lambda(Bx)$ defined for $B \in \mathcal{B}(G)$ is also a Haar measure, and therefore differs by a multiplicative constant from λ. That is, for each $x \in G$, $\lambda_x = \Delta(x)\lambda$ for some $\Delta(x) \in (0, \infty)$. The function $\Delta : G \to (0, \infty)$ is called the **modular function** of G, and it depends only on G and not which scalar multiple was chosen to be λ. One has that Δ is, in fact, a continuous homomorphism. One says that G is **unimodular** if $\Delta = 1$ everywhere. In other words, G is unimodular if and only if $\lambda(Bx) = \lambda(B)$ for all $B \in \mathcal{B}(G)$. In this case, λ is bi-invariant.

One may wonder what this type of unimodularity has to do with the unimodularity of random networks. There are many deep connections, but the one that will be explored most in Chapter 5 is a mass-transport principle very similar in form to Equation (2.1), except that it applies to G-stationary point processes instead of unimodular networks. This ultimately gives the main result of the chapter, which is an analogous classification theorem to Theorem 2.5.1.

Chapter 3

Exact Coupling of Random Walks on Polish Groups

Let G be a Polish group with identity e. Recall that a Polish group is a group equipped with a topology under which multiplication and inversion are continuous operations, and such that the topology is separable and completely metrizable. The most poignant examples to keep in mind throughout are $\mathbb{R}^d$ and $\mathbb{Z}^d$. If G is Abelian, additive notation is used instead and the identity is denoted 0. Fix, for the remainder of the chapter, a Borel probability measure μ on G. For each $x \in G$, let $\mathrm{RW}(x, \mu)$ be the law of a (right) random walk on G started at x and with **step-length distribution** μ. That is, $\mathrm{RW}(x, \mu)$ is the law on the product space $G^{\mathbb{N}}$ of a process $S^x = (S_n^x)_{n=0}^{\infty}$ such that

$$S_n^x = xX_1X_2 \cdots X_n, \qquad 0 \leqslant n < \infty, \tag{3.1}$$

where the **step-lengths** $(X_i)_{i=1}^{\infty}$ are i.i.d. random elements in G with distribution μ. Such a process $S^x \sim \mathrm{RW}(x, \mu)$ is called an (x, μ)**-random walk.**

One may be interested in the long-term effects of the choice of the initial location $x \in G$ of a random walk. After a long time, can one distinguish an (x, μ)-random walk from a (y, μ)-random walk in the sense of total

variation? That is, for (x, μ)- and (y, μ)-random walks S^x and S^y, one would like to know whether

$$\left\|\mathbf{P}(S_n^x \in \cdot) - \mathbf{P}(S_n^y \in \cdot)\right\|_{\mathrm{TV}} \to 0, \qquad n \to \infty, \tag{3.2}$$

where $\|\nu\|_{\mathrm{TV}} = \sup_B \nu(B) - \inf_B \nu(B)$ denotes the total variation of a finite signed measure ν. The primary result of this chapter precisely determines, under certain conditions, when (3.2) occurs. Note that for probability measures ν_1 and ν_2, one also has

$$\|\nu_1 - \nu_2\|_{\mathrm{TV}} = 2 \sup_B |\nu_1(B) - \nu_2(B)|. \tag{3.3}$$

An equivalent formulation of (3.2) may be expressed in terms of successful exact couplings.

An **exact coupling** of $\mathrm{RW}(x, \mu)$ and $\mathrm{RW}(y, \mu)$ is a triple (S^x, S^y, T) defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ in such a way that $S^x \sim \mathrm{RW}(x, \mu)$, $S^y \sim \mathrm{RW}(y, \mu)$, and T is a random time, called a **coupling time**, such that

$$S_n^x = S_n^y, \qquad n \geqslant T. \tag{3.4}$$

If T is a.s. finite, the exact coupling is called **successful**. If $\widetilde{S}^x \sim \mathrm{RW}(x, \mu)$ and $\widetilde{S}^y \sim \mathrm{RW}(y, \mu)$ are defined on possibly different spaces, one also calls (S^x, S^y, T) an exact coupling of $\widetilde{S}^x$ and $\widetilde{S}^y$. The condition (3.2) is equivalent to the statement that $\mathrm{RW}(x, \mu)$ and $\mathrm{RW}(y, \mu)$ admit a successful exact coupling. See Theorem 9.4 in Chapter 4, Section 9.5 of [Tho00] for the equivalence of these statements. In essence, successful exact coupling may be achieved if

and only if the initial condition is uniformly forgotten as time progresses. Moreover, the tail probabilities of a coupling time control the speed at which the total variation distance between the two random walks decays. Indeed, if (S^x, S^y, T) is an exact coupling as above, one has for any Borel set $B \subseteq G$,

$$\left|\mathbf{P}(S_n^x \in B) - \mathbf{P}(S_n^y \in B)\right| = \left|\mathbf{E}[1_{\{S_n^x \in B\}} - 1_{\{S_n^y \in B\}}]\right| \leqslant \mathbf{P}(T > n), \qquad n \geqslant 0.$$

Multiplying by 2 and taking the supremum over all Borel B, one finds

$$\left\|\mathbf{P}(S_n^x \in \cdot) - \mathbf{P}(S_n^y \in \cdot)\right\|_{\mathrm{TV}} \leqslant 2\mathbf{P}(T > n), \qquad n \geqslant 0. \tag{3.5}$$

This chapter investigates under what conditions successful exact couplings may be constructed. When successful exact coupling can be achieved, the constructed coupling time is analyzed to give bounds on the rate at which total variation distance decays.

Note that if S^x is an (x, μ)-random walk and $y \in G$, then yS^x is a (yx, μ)-random walk. Hence $\mathrm{RW}(x, \mu)$ and $\mathrm{RW}(y, \mu)$ admit a successful exact coupling if and only if $\mathrm{RW}(e, \mu)$ and $\mathrm{RW}(y^{-1}x, \mu)$ admit a successful exact coupling. It therefore suffices to study only the case when one of the initial locations is the identity. That is, for what initial positions x do $\mathrm{RW}(e, \mu)$ and $\mathrm{RW}(x, \mu)$ admit a successful exact coupling?

Definition 3.0.1. Define the **successful exact coupling set** G_s to be the subset of all $x \in G$ such that there exists a successful exact coupling of $\mathrm{RW}(e, \mu)$ and $\mathrm{RW}(x, \mu)$.

The primary question is then to determine what is the set G_s. This question was posed by Thorisson in [Tho11] for $G := \mathbb{R}$, and in that case the following two special cases were known as early as 1965, cf. [Sta66, Her65], though references using more modern notation are cited here. In the following, recall that when G admits a (left-invariant) Haar measure λ (e.g. Lebesgue measure on $\mathbb{R}^d$ or the cardinality counting measure on $\mathbb{Z}^d$), then μ is called **spread out** if for some $n \geqslant 1$ one has $\mu^n \geqslant \int f \, d\lambda$ for some Borel $f \geqslant 0$ not λ-a.e. zero, where for a Borel measure ν on G, ν^n denotes the n-fold convolution of ν with itself.

Theorem 3.0.2. *[Tho00] Let $G := \mathbb{R}$. Then $G_s = G$ if and only if the step-length distribution μ is spread out.*

Theorem 3.0.3. *[Arn10] Let $G := \mathbb{R}$ and suppose μ is purely atomic with A denoting the set of atoms of μ. Then G_s is the subgroup generated by $A - A = \{a - a' : a, a' \in A\}$.*

Given two Borel measures ν_1 and ν_2 on G, denote $\nu_1 \wedge \nu_2$ to be the largest measure smaller than ν_1 and ν_2. The zero measure is denoted $\mathbf{0}$. For $x \in G$, also define the shift $\theta_x \nu$ by $\theta_x \nu(B) := \nu(x^{-1}B)$ for each Borel $B \subseteq G$. The interpretation of $\theta_x \nu$ is ν with all mass shifted (left-multiplied) by x, and $\theta_x \nu$ satisfies $\int_G f(y) \, \theta_x \nu(dy) = \int_G f(xy) \, \nu(dy)$ for all Borel $f : G \to \mathbb{R}_{\geqslant 0}$.

The resolution to Thorisson's problem and generalizations of the previous theorems may now be stated. The proof is postponed and broken into several separate more general theorems appearing across multiple sections.

Theorem 3.0.4. *Suppose G is Abelian. Then the following hold:*

(a) $G_s = \{x \in G : \exists n \geqslant 1, \mu^n \wedge \theta_x^{-1}\mu^n \neq \mathbf{0}\}$.

(b) For $x \in G_s$ and $n_0 \geqslant 1$ such that $\mu^{n_0} \wedge \theta_x^{-1}\mu^{n_0} \neq \mathbf{0}$, there is $C = C(\mu, x, n_0) > 0$ such that for $S \sim \mathrm{RW}(0, \mu)$ and $S^x \sim \mathrm{RW}(x, \mu)$ under $\mathbf{P}$, one has

- *if x has infinite order,*

$$\|\mathbf{P}(S_n \in \cdot) - \mathbf{P}(S_n^x \in \cdot)\|_{\mathrm{TV}} \leqslant \frac{C}{\sqrt{n}}, \qquad n \geqslant 1, \tag{3.6}$$

- *if x has finite order, there is $\rho = \rho(\mu, x, n_0) \in (0, 1)$ such that*

$$\|\mathbf{P}(S_n \in \cdot) - \mathbf{P}(S_n^x \in \cdot)\|_{\mathrm{TV}} \leqslant C\rho^n \qquad n \geqslant 1. \tag{3.7}$$

(c) Suppose G is locally compact with Haar measure λ. If $G_s = G$, then μ is spread out. If G is connected, the converse holds as well.

(d) Suppose μ is purely atomic with A denoting the set of atoms of μ. Then G_s is the subgroup generated by $A - A$.

(e) G_s is a Borel measurable subgroup of G.

Section 3.1 builds to the main theorem, Theorem 3.1.6, which generalizes Theorem 3.0.4 (a) and (b). It is more technical but also applies in some non-Abelian cases. The reader familiar with the proof of the spread out case [Tho00] or the purely atomic case [Arn10] on $\mathbb{R}$ may recognize the

proof of Proposition 3.1.5, which shows that if μ^n dominates the sum of a measure ν and some shift $\theta_x^{-1}\nu$, then successful exact coupling can be achieved. Similarly, the proof of the main theorem of this chapter follows the spirit of the purely atomic case on $\mathbb{R}$.

With the main theorem proved, Section 3.2 covers parts (a)-(d) of Theorem 3.0.4 as simple corollaries. That is, it resolves the Abelian case and gives an even simpler description of G_s in the spread out and purely atomic cases. Note that Corollary 3.2.2 is more general than claimed in Theorem 3.0.4 (c), as one direction applies in the non-Abelian case without extra restrictions.

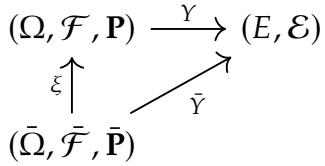
Section 3.3 then investigates the structure of G_s. In particular, it is shown to be Borel measurable. This, together with the fact that G_s is a group, shows part (e) of Theorem 3.0.4.

Finally, in Section 3.4, the weaker notion of possible exact coupling is studied. It is noted that the necessary and sufficient conditions derived for successful exact coupling in the Abelian case are coincidental. It is shown that the conditions derived for admitting a successful exact coupling in the Abelian case are, in the general case, equivalent to the ostensibly weaker notion of admitting a possible exact coupling, but that in the Abelian case admitting a possible exact coupling and admitting a successful exact coupling are equivalent. The chapter ends with an example on a (non-Abelian) free group for which possible exact coupling can be done but successful exact coupling cannot.

3.1. The Successful Exact Coupling Existence Criterion

This section culminates in the main theorem of the chapter, Theorem 3.1.6, which gives necessary and sometimes sufficient conditions for successful exact coupling to occur, even in the non-Abelian case. The transfer and splitting theorems that appear in [Tho00] are used. Less general versions are stated that are sufficient for the current setting. The need of a Polish space in the transfer and splitting theorems is also the primary reason G is assumed to be Polish.

In order to state the transfer and splitting theorems, first the notion of an extension of a probability space is needed. Recall that $(E, \mathcal{E})$ is a measurable space if E is a set and $\mathcal{E}$ is a σ-algebra on E, and that a random element in $(E, \mathcal{E})$ is a measurable map $Y : \Omega \to E$, where $(\Omega, \mathcal{F}, \mathbf{P})$ is some probability space. When the target space is $\mathbb{R}$ or $\mathbb{R}^d$ this is usually called a random variable or random vector. One says $((\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{P}}), \xi)$ is an **extension** of a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ if $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{P}})$ is a probability space and $\xi : \bar{\Omega} \to \Omega$ is a random element with distribution $\mathbf{P}$, that is to say, $\bar{\mathbf{P}}(\xi \in \cdot) = \mathbf{P}$. In this case, for any random element Y defined on Ω and taking values in any measurable space $(E, \mathcal{E})$, ξ induces a canonical copy $\bar{Y} := Y \circ \xi$ of Y that is defined on $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{P}})$, as depicted in the following diagram.



By calling $\bar{Y}$ a **copy** or **version** of Y, one means that $\bar{Y}$ and Y have the same distribution, which follows in this case from the fact that

$$\bar{\mathbf{P}}(\bar{Y} \in \cdot) = \bar{\mathbf{P}}(\xi \in Y^{-1}(\cdot)) = \mathbf{P}(Y \in \cdot).$$

Because $\bar{Y}$ is induced by ξ as described, one refers to $\bar{Y}$ as an **original** random element. That is, $\bar{Y}$ represents a random element coming from $(\Omega, \mathcal{F}, \mathbf{P})$, as opposed to a random element defined on the extended space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{P}})$ that does not come from a random element defined on $(\Omega, \mathcal{F}, \mathbf{P})$.

Now it is possible to state the transfer and splitting theorems.

Theorem 3.1.1 (Transfer Theorem)**.** *[Tho00] Suppose $(\Omega, \mathcal{F}, \mathbf{P})$ is a probability space and Y_1 is a random element in some measurable space. Further suppose that there is a pair (Y_1', Y_2') on some probability space $(\Omega', \mathcal{F}', \mathbf{P}')$ with Y_2' a random element in some Polish space, and Y_1 is a version of Y_1'. Then Y_2' can be **transferred** to $(\Omega, \mathcal{F}, \mathbf{P})$. That is, there exists an extension $((\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{P}}), \xi)$ of $(\Omega, \mathcal{F}, \mathbf{P})$ that contains a copy $\bar{Y}_2$ of Y_2' such that:*

(i) $(\bar{Y}_1, \bar{Y}_2)$ has the same distribution as (Y_1', Y_2'), where $\bar{Y}_1 = Y_1 \circ \xi$ is the induced copy of Y_1.

(ii) If Y_0 is a random element defined on $(\Omega, \mathcal{F}, \mathbf{P})$ and $\bar{Y}_0$ its induced copy, then $\bar{Y}_2$ is conditionally independent of $\bar{Y}_0$ given $\bar{Y}_1$, i.e.,

$$\bar{\mathbf{P}}(\bar{Y}_2 \in \cdot \mid \bar{Y}_1) = \bar{\mathbf{P}}(\bar{Y}_2 \in \cdot \mid \bar{Y}_0, \bar{Y}_1).$$

This transfer procedure can be repeated countably many times.

Theorem 3.1.2 (Splitting Theorem)**.** *[Tho00] Suppose* $(\Omega, \mathcal{F}, \mathbf{P})$ *is a probability space and* Y *is a random element in a measurable space* $(E, \mathcal{E})$*. Let* $(\nu_i)_{i=0}^{\infty}$ *be subprobability measures on* $(E, \mathcal{E})$ *and suppose* $\mathbf{P}(Y \in \cdot) \geqslant \sum_i \nu_i$*. Then there exists an extension* $((\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{P}}), \xi)$ *of* $(\Omega, \mathcal{F}, \mathbf{P})$ *containing a nonnegative integer-valued random variable* $\bar{K}$*, called a* ***splitting variable****, such that:*

(i) $\bar{\mathbf{P}}(\bar{Y} \in \cdot, \bar{K} = i) = \nu_i$ *for each* i*, where* $\bar{Y} = Y \circ \xi$ *is the induced copy of* Y*.*

(ii) If Y_0 *is a random element defined on* $(\Omega, \mathcal{F}, \mathbf{P})$ *and* $\bar{Y}_0$ *its induced copy, then* $\bar{K}$ *is conditionally independent of* $\bar{Y}_0$ *given* $\bar{Y}$*, i.e.,*

$$\bar{\mathbf{P}}(\bar{K} \in \cdot \mid \bar{Y}) = \bar{\mathbf{P}}(\bar{K} \in \cdot \mid \bar{Y}_0, \bar{Y}).$$

This splitting operation can be repeated countably many times.

After an extension $((\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{P}}), \xi)$ is introduced, one typically wants to pretend that one started with the extension in the first place. Therefore, when there is no risk of confusion, one often drops the bar notation and renames the extended space to $(\Omega, \mathcal{F}, \mathbf{P})$, and any induced variable $\bar{Y}$ to Y. This allows one to assume that all the random elements being considered are defined on a common probability space. One signals that this renaming is happening by saying that $(\Omega, \mathcal{F}, \mathbf{P})$ has been "extended" to house new random elements with certain properties. All the results in this chapter concern only the laws or joint distributions of random elements, so there is no loss of precision by replacing a space by an extension when needed.

Theorems 3.1.1 and 3.1.2 are useful for constructing random variables with specific dependencies on a single probability space. A simple example application is the following. If one can construct up to a countable number of successful exact couplings, then in fact they can be made to occur on the same probability space.

Proposition 3.1.3. *If $N \in \mathbb{N} \cup \{\infty\}$ and $\{x_i\}_{i=1}^{N} \subseteq G_s$, then there exists a single probability space $(\Omega, \mathcal{F}, \mathbf{P})$ on which there exist $S \sim \mathrm{RW}(e, \mu)$ and $S^{x_i} \sim \mathrm{RW}(x_i, \mu)$ for each i such that for every i there is an a.s. finite random time T_i with $S_n^{x_i} = S_n$ for all $n \geqslant T_i$. That is, (S, S^{x_i}, T_i) is a successful exact coupling for all i.*

Proof. By assumption, there is a successful exact coupling of $\mathrm{RW}(e, \mu)$ and $\mathrm{RW}(x_1, \mu)$ on some $(\Omega, \mathcal{F}, \mathbf{P})$. Since $x_i \in G_s$ for each i, the extension procedure given by Theorem 3.1.1 can be repeated countably many times, once for each S^{x_i}, to give a single extension of $(\Omega, \mathcal{F}, \mathbf{P})$ on which S and S^{x_i} couple for every i. □

This gives the first structural result about the successful exact coupling set.

Corollary 3.1.4. *G_s is a group.*

Proof. Consider $x, y \in G_s$. By Proposition 3.1.3, respectively define successful exact couplings (S, S^x, T^x) and (S, S^y, T^y) on a common probability space. Then for $n \geqslant \max\{T^x, T^y\}$ it holds that $S_n^y = S_n = S_n^x$. In particular,

$(x^{-1}S^x, x^{-1}S^y, \max\{T^x, T^y\})$ is a successful exact coupling of $\mathrm{RW}(e, \mu)$ and $\mathrm{RW}(x^{-1}y, \mu)$. Thus $x^{-1}y \in G_s$. □

The following works towards determining a specific scenario when a successful exact coupling can be constructed and it is an extension of a result of Ö. Arnaldsson in [Arn10] with nearly identical proof.

Proposition 3.1.5. *Fix $x \in G$ and suppose that $n \geqslant 1$ is such that $\mu^n \geqslant \nu + \theta_x^{-1}\nu$ for a nonzero measure ν. If G is Abelian or, more generally, if there is B with $\mu^n(B) = 1$ such that x commutes with all of B, then $x \in G_s$. In this case,* $\mathrm{RW}(e, \mu)$ *and* $\mathrm{RW}(x, \mu)$ *admit a successful exact coupling with a coupling time T for which T/n has the same distribution as the hitting time of e of a lazy simple symmetric random walk on the cyclic group $\langle x \rangle$ started at x with probability $1 - 2\nu(G)$ of not moving at each step. In particular, $\mathbf{P}(T = n) = \nu(G)$.*

Proof. Begin with an $(\Omega, \mathcal{F}, \mathbf{P})$ on which there is $S \sim \mathrm{RW}(e, \mu)$ with step-lengths $(X_i)_{i=1}^{\infty}$. An extension of $(\Omega, \mathcal{F}, \mathbf{P})$ and an $S^x \sim \mathrm{RW}(x, \mu)$ on that extension are constructed such that successful exact coupling occurs. Let

$$L_i := X_{(i-1)n+1} \cdots X_{in} \tag{3.8}$$

for $i \geqslant 1$ so that $(L_i)_{i=1}^{\infty}$ is an i.i.d. family and $\mathbf{P}(L_i \in \cdot) = \mu^n \geqslant \nu + \theta_x^{-1}\nu$. By Theorem 3.1.2, expand $(\Omega, \mathcal{F}, \mathbf{P})$ to accommodate random variables $(K_i)_{i=1}^{\infty}$ taking values in $\{0, 1, 2\}$ such that $(L_i, K_i)_{i=1}^{\infty}$ is an i.i.d. sequence and, for each i,

$$\mathbf{P}(L_i \in \cdot, K_i = 1) = \nu, \qquad \mathbf{P}(L_i \in \cdot, K_i = 2) = \theta_x^{-1}\nu. \tag{3.9}$$

For $i \geqslant 1$ define

$$L_i' := \begin{cases} L_i, & K_i = 0, \\ x^{-1}L_i, & K_i = 1, \\ xL_i, & K_i = 2. \end{cases} \tag{3.10}$$

It is elementary to check using (3.9) that L_i' has the same distribution as L_i. Let R be the random walk started at e with step-lengths $(L_i)_{i=1}^{\infty}$, and let R' be the random walk started at x with step-lengths $(L_i')_{i=1}^{\infty}$. By construction, $L_i'L_i^{-1} \in \{e, x, x^{-1}\}$ for each i. By assumption, it is possible to choose B with $\mu^n(B) = 1$ such that x commutes with all of B. Thus, a.s. every $L_i, L_i' \in B$ and so a.s. for every i,

$$R_i'R_i^{-1} = xL_1' \cdots L_i'L_i^{-1} \cdots L_1^{-1} = x(L_1'L_1^{-1}) \cdots (L_i'L_i^{-1}) \in \langle x \rangle = \{x^m : m \in \mathbb{Z}\}.$$

Thus, $R'R^{-1}$ is in distribution the same as a lazy simple symmetric random walk started from x with step-lengths $(L_i'L_i^{-1})_{i=1}^{\infty}$. The walk has probability $\nu(G)$ to increase the power of x, $\nu(G)$ to decrease it, and $1 - 2\nu(G) < 1$ to stay put at each time step. Since nontrivial lazy simple symmetric random walks on cyclic groups are recurrent, there is an a.s. finite random time M with $R_M'R_M^{-1} = e$, i.e. the random walks R' and R meet at time M. Theorem 3.1.1 makes it possible to extend $(\Omega, \mathcal{F}, \mathbf{P})$ one final time to accommodate an i.i.d. sequence $(X_i'')_{i=1}^{\infty}$ with each X_i'' having distribution μ and such that $L_i' = X''_{(i-1)n} \cdots X''_{in}$ for $i \geqslant 1$. Define $T := Mn$ and let S^x be the random walk started at x with step-lengths

$$X_i' := \begin{cases} X_i'', & i \leqslant T, \\ X_i, & i > T. \end{cases} \tag{3.11}$$

Then S and S^x witness the definition of successful exact coupling with coupling time T. If $K_1 = 1$, then R and R' meet in one time step, so $T = n$, showing $\mathbf{P}(T = n) = \mathbf{P}(K_1 = 1) = \nu(G)$. □

In the previous proof, the problem is reduced to the case where a difference process (or, in the non-Abelian case, something that resembles a difference process) is a random walk. Since a general random walk may be transient, it is important to the proof that the difference process is made to be a random walk not on all of G, but rather on the cyclic group generated by x, so that the analysis reduces to that of $\mathbb{Z}$ or $\mathbb{Z}/d\mathbb{Z}$. This highlights the fact that the joint distribution of S and S^x required to cause successful exact coupling is very special, and could not, except in trivial cases, be achieved with S and S^x being independent before the coupling time T.

The main theorem of the chapter follows.

Theorem 3.1.6 (Successful Exact Coupling Existence Criterion)**.** *Fix $x \in G$. If $x \in G_s$, then there is $n \geqslant 1$ such that $\mu^n \wedge \theta_x^{-1}\mu^n \neq \mathbf{0}$. Conversely, if $n_0 \geqslant 1$ is such that $\mu^{n_0} \wedge \theta_x^{-1}\mu^{n_0} \neq \mathbf{0}$ and there exists B with $\mu^{n_0}(B) = 1$ such that x commutes with all of B, then $x \in G_s$. In this case, there exists a successful exact coupling of $\mathrm{RW}(e, \mu)$ and $\mathrm{RW}(x, \mu)$ with a coupling time T satisfying $\mathbf{P}(T = n_0) > 0$. Moreover, there is $C = C(\mu, x, n_0) > 0$ such that for $S \sim \mathrm{RW}(0, \mu)$ and $S^x \sim \mathrm{RW}(x, \mu)$ under $\mathbf{P}$, one has*

- *if x has infinite order,*

$$\|\mathbf{P}(S_n \in \cdot) - \mathbf{P}(S_n^x \in \cdot)\|_{\mathrm{TV}} \leqslant \frac{C}{\sqrt{n}}, \qquad n \geqslant 1, \tag{3.12}$$

- *if x has finite order, there is $\rho = \rho(\mu, x, n_0) \in (0,1)$ such that*

$$\|\mathbf{P}(S_n \in \cdot) - \mathbf{P}(S_n^x \in \cdot)\|_{\mathrm{TV}} \leqslant C\rho^n \qquad n \geqslant 1. \tag{3.13}$$

Proof. Suppose that $S \sim \mathrm{RW}(e, \mu)$ and $S^x \sim \mathrm{RW}(x, \mu)$ witness the definition of successful exact coupling with coupling time T and respective step-lengths $(X_i)_{i=1}^{\infty}$ and $(X_i')_{i=1}^{\infty}$. Choose n such that $\mathbf{P}(T = n) > 0$, which is possible since T is a.s. finite. Then one has the following comparisons of measures,

$$\begin{aligned}
\mathbf{0} \neq \mathbf{P}(T = n, S_n = S_n^x \in \cdot) \\
&\leqslant \mathbf{P}(S_n = S_n^x \in \cdot) \\
&\leqslant \mathbf{P}(S_n \in \cdot) \wedge \mathbf{P}(S_n^x \in \cdot) \\
&= \mu^n \wedge \theta_x \mu^n.
\end{aligned}$$

Applying θ_x^{-1} to both sides of the previous inequality then gives $\mathbf{0} \neq \theta_x^{-1}\mu^n \wedge \mu^n$.

Conversely, suppose n_0 is such that $\xi := \mu^{n_0} \wedge \theta_x^{-1}\mu^{n_0} \neq \mathbf{0}$ and that there exists B with $\mu^{n_0}(B) = 1$ such that x commutes with all of B. In case $x = e$, $\mathrm{RW}(e, \mu)$ and $\mathrm{RW}(x, \mu)$ clearly admit a successful exact coupling with a coupling time $T := 0$, so assume $x \neq e$. Choose $y \in \operatorname{supp} \xi$. Since $y \neq xy$,

it is possible to choose a neighborhood U of y small enough that $U \cap xU = \emptyset$. Consider

$$\nu := \xi((x^{-1}\cdot) \cap U) \neq \mathbf{0}. \tag{3.14}$$

Then

$$\nu \leqslant \mu^{n_0}\left(x\left((x^{-1}\cdot) \cap U\right)\right) = \mu^{n_0}(\cdot \cap xU)$$

and

$$\theta_x^{-1}\nu = \xi(\cdot \cap U) \leqslant \mu^{n_0}(\cdot \cap U).$$

It follows that

$$\nu + \theta_x^{-1}\nu \leqslant \mu^{n_0}(\cdot \cap (U \cup xU)) \leqslant \mu^{n_0}. \tag{3.15}$$

Proposition 3.1.5 then shows $\mathrm{RW}(e, \mu)$ and $\mathrm{RW}(x, \mu)$ admit a successful exact coupling with a coupling time T satisfying $\mathbf{P}(T = n_0) = \nu(G) > 0$ and such that $\tau := T/n_0$ has the distribution of the hitting time to e of a symmetric lazy random walk on $\langle x \rangle$ with $1 - 2\nu(G)$ chance of not moving at each step.

Suppose that x has finite order d. In this case, τ also has the distribution of the hitting time to 0 of the symmetric lazy random walk on $\mathbb{Z}/d\mathbb{Z}$ that is started at 1 and absorbed when it hits 0. Call P the transition kernel of the absorbing walk. The absorbing walk converges geometrically quickly to its stationary distribution δ_0, cf. [LP17]. Choose $\rho \in (0, 1)$ and $C > 0$ such that $\|P^n(1, \cdot) - \delta_0\|_{\mathrm{TV}} \leqslant C\rho^n$ for all $n \geqslant 1$. By (3.5) it suffices to show that

$\mathbf{P}(T > n)$ decays geometrically as $n \to \infty$. Indeed, for all $n \geqslant 1$,

$$\begin{aligned}
\mathbf{P}(T > n) &= \mathbf{P}(\tau > n/n_0) \\
&= P^{\lfloor n/n_0 \rfloor + 1}(1, \{0\}^c) \\
&\leqslant \left\| P^{\lfloor n/n_0 \rfloor + 1}(1, \cdot) - \delta_0 \right\|_{\mathrm{TV}} + \delta_0(\{0\}^c) \\
&\leqslant C\rho^{\lfloor n/n_0 \rfloor + 1} + 0 \\
&\leqslant \widetilde{C} \left(\widetilde{\rho}\right)^n
\end{aligned}$$

for new constants $\widetilde{C} > 0$ and $\widetilde{\rho} \in (0, 1)$, as desired.

Next suppose that x has infinite order, so $\langle x \rangle \simeq \mathbb{Z}$. The tail decay of symmetric lazy random walks on $\mathbb{Z}$ are known, see, for example, Corollary 2.28 in [LP17]. In [LP17], lazy random walks are defined to have chance $1/2$ of staying still at each step, but allowing a $1 - 2\nu(G) \in (0, 1)$ chance of staying still at each step does not modify the result beyond giving a different leading constant in the decay rate. Hence, by Corollary 2.28 in [LP17], choose $C > 0$ such that $\mathbf{P}(\tau > n) \leqslant \frac{C}{\sqrt{n}}$ for integers $n \geqslant 1$. Thus, for all $n \geqslant 1$,

$$\begin{aligned}
\mathbf{P}(T > n) &= \mathbf{P}(\tau > n/n_0) \\
&= \mathbf{P}(\tau > \lfloor n/n_0 \rfloor + 1) \\
&\leqslant \frac{C}{\sqrt{\lfloor n/n_0 \rfloor + 1}} \\
&\leqslant \frac{\widetilde{C}}{\sqrt{n}}
\end{aligned}$$

for some new constant $\widetilde{C} > 0$, as desired. □

3.2. The Abelian Case

In this section, parts (a)-(d) of Theorem 3.0.4 are derived as simple corollaries of the main theorem. Firstly, determining G_s can be resolved entirely for Abelian G. This is parts (a) and (b) of Theorem 3.0.4.

Corollary 3.2.1. *Suppose G is Abelian. Then $G_s = \{x \in G : \exists n \geqslant 1, \mu^n \wedge \theta_x^{-1}\mu^n \neq \mathbf{0}\}$. Moreover, for $x \in G_s$ and $n_0 \geqslant 1$ such that $\mu^{n_0} \wedge \theta_x^{-1}\mu^{n_0} \neq \mathbf{0}$, there is $C = C(\mu, x, n_0) > 0$ such that for $S \sim \mathrm{RW}(0, \mu)$ and $S^x \sim \mathrm{RW}(x, \mu)$ under $\mathbf{P}$, one has*

- *if x has infinite order,*

$$\|\mathbf{P}(S_n \in \cdot) - \mathbf{P}(S_n^x \in \cdot)\|_{\mathrm{TV}} \leqslant \frac{C}{\sqrt{n}}, \qquad n \geqslant 1,$$

- *if x has finite order, there is $\rho = \rho(\mu, x, n_0) \in (0, 1)$ such that*

$$\|\mathbf{P}(S_n \in \cdot) - \mathbf{P}(S_n^x \in \cdot)\|_{\mathrm{TV}} \leqslant C\rho^n \qquad n \geqslant 1.$$

Proof. Since G is Abelian, the condition in Theorem 3.1.6 that there is B with $\mu^n(B) = 1$ such that x commutes with all of B is automatic. □

Next, a generalization of part (c) of Theorem 3.0.4 is covered. That is, for connected spaces step-lengths are spread out if and only if a successful exact coupling can always be achieved. The only if direction is essentially the same as in [Ber79], Theorem 5.3.2, and it also applies in the non-Abelian setting.

Corollary 3.2.2. *Suppose G is locally compact with Haar measure λ. If $G_s = G$, then then μ is spread out. If G is connected and Abelian, the converse holds as well. More generally, if G is Abelian but not necessarily connected, then G_s is clopen.*

Proof. Suppose $\mathrm{RW}(e, \mu)$ and $\mathrm{RW}(x, \mu)$ admit a successful exact coupling for all $x \in G$. Then for all $x \in G$, $\left\|\mu^n - \theta_x^{-1}\mu^n\right\|_{\mathrm{TV}} \to 0$ as $n \to \infty$. Consequently,

$$G = \bigcup_{n=1}^{\infty} \{x \in G : \left\|\mu^n - \theta_x^{-1}\mu^n\right\|_{\mathrm{TV}} \leqslant 1\}. \tag{3.16}$$

The measurability of the sets $B_n := \{x \in G : \left\|\mu^n - \theta_x^{-1}\mu^n\right\|_{\mathrm{TV}} \leqslant 1\}$ for $n \geqslant 1$ is taken for granted here. This fact is proved in the upcoming Corollary 3.3.2. Choose n large enough that $\lambda(B_n) > 0$. Suppose for contradiction that a Borel set $N \subseteq G$ is such that $\mu^n(N) = 1$ but $\lambda(N) = 0$. Then $\lambda(N^{-1}) = 0$ as well, and

$$\begin{aligned}
0 &= \int_G \int_G 1_{\{sx \in B_n\}} 1_{\{x \in N^{-1}\}} \, \lambda(dx) \, \mu^n(ds) \\
&= \int_G \int_G 1_{\{x \in B_n\}} 1_{\{s^{-1}x \in N^{-1}\}} \, \lambda(dx) \, \mu^n(ds) \\
&= \int_G \int_{B_n} 1_{\{s \in xN\}} \, \lambda(dx) \, \mu^n(ds) \\
&= \int_{B_n} \theta_x^{-1}\mu^n(N) \, \lambda(dx) \\
&\geqslant \int_{B_n} \mu^n(N) - |\theta_x^{-1}\mu^n(N) - \mu^n(N)| \, \lambda(dx) \\
&\geqslant \int_{B_n} 1 - \frac{1}{2}\left\|\theta_x^{-1}\mu^n - \mu^n\right\|_{\mathrm{TV}} \, \lambda(dx) \\
&\geqslant \frac{1}{2}\lambda(B_n) \\
&> 0
\end{aligned}$$

which is a contradiction. It follows that μ^n must not be singular with respect to λ, and hence μ is spread out.

For the other direction, suppose G is Abelian and that $\nu := \mu^n \geqslant \int_\cdot f \, d\lambda$ as stated. By replacing f with $\min\{f, b\}1_K$ for some $b > 0$ and $K \subseteq G$ compact, one may assume f is bounded and compactly supported. Furthermore, it is claimed that by replacing n with $2n$ one may assume $f > \epsilon$ on some nonempty open set for some $\epsilon > 0$. Indeed,

$$\mu^{2n} = \nu * \nu \geqslant \int_\cdot f * f \, d\lambda.$$

Since f is bounded and compactly supported, the convolution $f * f$ is continuous, and also $\|f * f\|_{L^1} = \|f\|_{L^1}^2 > 0$, so f is not constant 0. Thus the assumption that $f > \epsilon > 0$ on some nonempty open set U and for some $\epsilon > 0$ is justified. In particular, choosing a symmetric neighborhood V of the identity such that $(U - x) \cap U \neq \emptyset$ for each $x \in V$, it holds that

$$\nu \wedge \theta_x^{-1}\nu(G) \geqslant \int_G \min\{f(y), f(x + y)\} \, \lambda(dy) \geqslant \int_{(U-x)\cap U} \epsilon\lambda(dy) > 0$$

for every $x \in V$. It follows that $G_s \supseteq V$. By Corollary 3.1.4, G_s is a subgroup of G, and thus G_s is either clopen or has empty interior. Since G_s contains the nonempty open set V, G_s must be clopen. If G is connected then this implies $G_s = G$. □

The connectedness assumption in Corollary 3.2.2 plays a nontrivial role. For example, consider when G is a countable group. Then any choice of μ is automatically purely atomic because G is countable and spread out

because the Haar measure is a counting measure. The following corollary shows that in that case the conclusion of Corollary 3.2.2 does not hold. This is also part (d) of Theorem 3.0.4.

Corollary 3.2.3. *Suppose G is Abelian and μ is purely atomic with A the set of atoms of μ. Then G_s is the subgroup generated by $A - A$.*

Proof. The atoms of μ^n are $nA := A + \cdots + A$. Then since μ^n is atomic, $\mu^n \wedge \theta_x^{-1}\mu^n \neq \mathbf{0}$ if and only if $nA \cap (nA - x) \neq \emptyset$ if and only if $x \in nA - nA = n(A - A)$. Finally, note that $\bigcup_{n=1}^{\infty} n(A - A)$ is exactly the subgroup generated by $A - A$ since $A - A$ is symmetric. Corollary 3.2.1 then finishes the claim. □

The section ends by showing that, in the Abelian case, any countable subgroup can be a successful exact coupling set, and that the Haar measure is insufficient to measure the size of G_s.

Corollary 3.2.4. *Suppose G is Abelian and H is a countable subgroup of G. Then there is a choice of μ for which $G_s = H$.*

Proof. Any purely atomic μ whose set of atoms is H suffices. If μ is as mentioned, then since the subgroup generated by $H - H$ is H itself, one finds that $G_s = H$ by Corollary 3.2.3. □

Corollary 3.2.5. *Suppose G is locally compact with Haar measure λ, and that G is connected and Abelian as well. If μ is not spread out, then $\lambda(G_s) = 0$.*

Proof. The measurability of G_s is proved in the upcoming Corollary 3.3.2. Here it is taken for granted. If $\lambda(G_s) > 0$, then $G_s = G_s - G_s$ contains a neighborhood of the identity by the Steinhaus Theorem [Str72]. In this case it follows as in the proof of Corollary 3.2.2 that $G_s = G$, which implies that μ is spread out by the same corollary. □

3.3. Properties of the Successful Exact Coupling Set

The primary goal of this section is to treat the measurability issues previously neglected. In the Abelian case, the successful exact coupling set is Borel measurable. To show this, a slight but natural extension of Exercise 6.10.72 in volume II of [Bog07], is required. The following gives the existence of a measurable choice of a family of Radon-Nikodym derivatives. Importantly, the following does not assume absolute continuity and instead produces Radon-Nikodym derivatives of the absolutely continuous parts of measures.

Proposition 3.3.1. *Let $(X, \mathcal{A}, \mu)$ be a finite measure space with $\mathcal{A}$ countably generated, and let $(T, \mathcal{B})$ be a measurable space. Let $(\mu_t)_{t\in T}$ be any family of finite measures on X such that for each $A \in \mathcal{A}$, the function $t \mapsto \mu_t(A)$ is $\mathcal{B}$-measurable. Then there is an $\mathcal{A}\otimes\mathcal{B}$-measurable $f : X\times T \to \mathbb{R}$ such that for every $t \in T$, $x \mapsto f(x,t)$ is a version of the Radon-Nikodym derivative of the absolutely continuous part of μ_t with respect to μ.*

Proof. First consider $X := [0,1]$ and $\mathcal{A} := \mathcal{B}([0,1])$, the Borel sets on $[0,1]$.

Fix a sequence $(\epsilon_n)_{n=0}^{\infty}$ with $\epsilon_n \searrow 0$. For every $t \in T$,

$$\lim_n \frac{\mu_t(B(x,\epsilon_n))}{\mu(B(x,\epsilon_n))} = \frac{d\mu_{t,a}}{d\mu}(x), \qquad \mu\text{-a.e. } x, \tag{3.17}$$

where $\mu_{t,a}$ denotes the absolutely continuous part of μ_t with respect to μ. This follows from, e.g., Theorem 5.8.8. in volume I of [Bog07]. Define

$$f(x,t) := \limsup_n \frac{\mu_t(B(x,\epsilon_n))}{\mu(B(x,\epsilon_n))} \tag{3.18}$$

for $x \in \operatorname{supp} \mu$ and $t \in T$, and $f(x,t) := 0$ otherwise. By (3.17), it suffices to show f is $\mathcal{A} \otimes \mathcal{B}$-measurable. Indeed, consider a fixed n and consider the numerator

$$(x,t) \mapsto \mu_t(B(x,\epsilon_n)) = \int_{[0,1]} 1_{\{|y-x|<\epsilon_n\}}\, \mu_t(dy).$$

Let $g(x,y) := 1_{\{|y-x|<\epsilon_n\}}$ and choose a sequence of measurable simple functions $(s_k)_{k=0}^{\infty}$ of the form

$$s_k(x,y) := \sum_{i=0}^{m_k} \alpha_{i,k} 1_{\{x\in A_{i,k}\}} 1_{\{y\in B_{i,k}\}}, \tag{3.19}$$

with $0 \leqslant s_k \leqslant 1$ and $A_{i,k}, B_{i,k} \in \mathcal{B}([0,1])$ for each k, and $s_k \to g$ as $k \to \infty$. Then

$$\int_{[0,1]} 1_{\{|y-x|<\epsilon_n\}}\, \mu_t(dy) = \lim_k \sum_{i=0}^{m_k} \alpha_{i,k} 1_{\{x\in A_{i,k}\}} \mu_t(B_{i,k}),$$

which shows $(x,t) \mapsto \mu_t(B(x,\epsilon_n))$ is a limit of $\mathcal{A} \otimes \mathcal{B}$-measurable functions, showing its measurability. The argument for the denominator $(x,t) \mapsto \mu(B(x,\epsilon_n))$ is similar and easier. It follows that f is $\mathcal{A} \otimes \mathcal{B}$-measurable.

Next, consider a general X and $\mathcal{A}$. Since $\mathcal{A}$ is countably generated, choose an $\mathcal{A}$-measurable $\phi : X \to [0,1]$ such that $\mathcal{A} = \{\phi^{-1}(B) : B \in \mathcal{B}([0,1])\}$, cf. Theorem 6.5.5 in volume II of [Bog07]. Also set

$$\nu := \mu(\phi \in \cdot), \qquad \nu_t := \mu_t(\phi \in \cdot), \tag{3.20}$$

for each $t \in T$. For each $B \in \mathcal{B}([0,1])$, it holds that $A := \phi^{-1}(B) \in \mathcal{A}$ and $t \mapsto \nu_t(B) = \mu_t(A)$ is $\mathcal{B}$-measurable. By the case where $X = [0,1]$ and $\mathcal{A} = \mathcal{B}([0,1])$, choose $f : [0,1] \times T \to \mathbb{R}$ that is $\mathcal{B}([0,1]) \otimes \mathcal{B}$-measurable and such that for all $t \in T$, $f(\cdot, t)$ is a version of the Radon-Nikodym derivative of the absolutely continuous part of ν_t with respect to ν. Define $f_0 : X \times T \to \mathbb{R}$ by $f_0(x,t) := f(\phi(x), t)$. Then f_0 is $\mathcal{A} \otimes \mathcal{B}$-measurable. Fix $t \in T$ and let $A \in \mathcal{A}$ be given. Choose $B \in \mathcal{B}([0,1])$ with $A = \phi^{-1}(B)$. Then

$$\begin{aligned}
\int_X 1_{\{x \in A\}} f_0(x,t)\, \mu(dx) &= \int_X 1_{\{\phi(x) \in B\}} f(\phi(x), t)\, \mu(dx) \\
&= \int_{[0,1]} 1_{\{y \in B\}} f(y,t)\, \nu(dy) \\
&= \nu_{t,a}(B) \\
&= \mu_{t,a}(A).
\end{aligned}$$

Some care should be taken in the last equality, where it is used that the absolutely continuous part of $\mu_t(\phi \in \cdot)$ with respect to $\mu(\phi \in \cdot)$ is the same as the push-forward with respect to ϕ of the absolutely continuous part of μ_t with respect to μ. Write

$$\nu_t = \nu_{t,a} + \nu_{t,s}, \qquad \mu_t = \mu_{t,a} + \mu_{t,s},$$

with $\nu_{t,a} \ll \nu$ and $\nu_{t,s} \perp \nu$, and $\mu_{t,a} \ll \mu$ and $\mu_{t,s} \perp \mu$. Then also

$$\nu_t = \mu_t(\phi \in \cdot) = \mu_{t,a}(\phi \in \cdot) + \mu_{t,s}(\phi \in \cdot),$$

so it suffices to show by the uniqueness of Lebesgue decompositions that

$$\mu_{t,a}(\phi \in \cdot) \ll \nu \text{ and } \mu_{t,s}(\phi \in \cdot) \perp \nu.$$

Indeed, if $B \in \mathcal{B}([0,1])$ is such that $0 = \nu(B) = \mu(\phi \in B)$, then $\mu_{t,a}(\phi \in \cdot) = 0$ because $\mu_{t,a} \ll \mu$. Thus $\mu_{t,a}(\phi \in \cdot) \ll \nu$. Similarly, choose $A \in \mathcal{A}$ such that $\mu_{t,s}(A^c) = \mu(A) = 0$. Choose $B \in \mathcal{B}([0,1])$ with $A = \phi^{-1}(B)$, then compute $\mu_{t,s}(\phi \in B^c) = \mu_{t,s}(A^c) = 0$ and $\nu(B) = \mu(\phi \in B) = \mu(A) = 0$, so that $\mu_{t,s}(\phi \in \cdot) \perp \nu$. The previous use of $\nu_{t,a}(B) = \mu_{t,a}(A)$ is now justified, showing that $f_0(\cdot, t)$ is a version of the Radon-Nikodym derivative of the absolutely continuous part of μ_t with respect to μ, completing the claim. □

The following, together with Corollary 3.1.4, gives Theorem 3.0.4 part (e).

Corollary 3.3.2. *For a probability measure ν on G, the maps $x \mapsto \left\|\nu - \theta_x^{-1}\nu\right\|_{\mathrm{TV}}$, $x \mapsto \left\|\nu \wedge \theta_x^{-1}\nu\right\|_{\mathrm{TV}}$, and the set $\{x : \nu \wedge \theta_x^{-1}\nu \neq \mathbf{0}\}$ are Borel measurable. In particular, if G is Abelian then $G_s = \cup_{n=1}^{\infty}\{x \in G : \mu^n \wedge \theta_x^{-1}\mu^n \neq \mathbf{0}\}$ is Borel measurable.*

Proof. Apply Proposition 3.3.1 with $X := T := G$ and the family of measures $\nu_t := \theta_t^{-1}\nu$ for $t \in G$. For $A \subseteq G$ open and $t_n \to t \in G$, Fatou's lemma

implies that

$$\begin{aligned}
\nu_t(A) &= \int_G 1_{\{x \in tA\}}\, \nu(dx) \\
&= \int_G 1_{\{t^{-1} \in Ax^{-1}\}}\, \nu(dx) \\
&\leqslant \int_G \liminf_n 1_{\{t_n^{-1} \in Ax^{-1}\}}\, \nu(dx) \\
&\leqslant \liminf_n \int_G 1_{\{t_n^{-1} \in Ax^{-1}\}}\, \nu(dx) \\
&= \liminf_n \nu_{t_n}(A),
\end{aligned}$$

so that $t \mapsto \nu_t(A)$ is semicontinuous and hence measurable. A monotone class argument shows that $t \mapsto \nu_t(A)$ is measurable for all Borel $A \subseteq G$. Thus, Proposition 3.3.1 gives a measurable $f : G \times G \to \mathbb{R}$ such that for every $t \in G$, $x \mapsto f(x,t)$ is a version of the Radon-Nikodym derivative of the absolutely continuous part of $\theta_t^{-1}\nu$ with respect to ν. It follows that

$$M(t) := \int_G \min\{f(x,t), 1\}\, \nu(dx) = \left\|\nu \wedge \theta_t^{-1}\nu\right\|_{\mathrm{TV}} \tag{3.21}$$

is measurable in t. Hence

$$\left\|\nu - \theta_t^{-1}\nu\right\|_{\mathrm{TV}} = 2 - 2\left\|\nu \wedge \theta_t^{-1}\nu\right\|_{\mathrm{TV}} \tag{3.22}$$

is measurable in t, and

$$\{t : \nu \wedge \theta_t^{-1}\nu \neq \mathbf{0}\} = \{t : M(t) > 0\} \tag{3.23}$$

is measurable as well. □

It is not known to the author in the non-Abelian case whether G_s is measurable. Even in the Abelian case though, little is known about other

structural properties of G_s. When is G_s nicer than Borel measurable? The worst case seen so far in Corollary 3.2.4 is that G_s may be any countable subgroup of G, which gives cases where G_s is an F_σ set but not closed (e.g. $\mathbb{Q} \subseteq \mathbb{R}$). Depending on G, this also gives cases where G_s is dense (e.g. $\mathbb{Q} \subseteq \mathbb{R}$), infinite but not dense (e.g. $\mathbb{Z} \subseteq \mathbb{R}$), and finite but not trivial (e.g. $\{-1, 1\} \subseteq \mathbb{R} \setminus \{0\}$). Corollary 3.2.5 indicates that in many cases either $G_s = G$ or $\lambda(G_s) = 0$, so in these cases the Haar measure on G is not useful to measure the size of G_s. Is there a natural measure with which to measure the size of G_s? What is the Hausdorff dimension of G_s, and can it be related, say, to the Hausdorff dimension of the subgroup generated by supp μ? All of these questions remain open and are not investigated further here.

3.4. Possible Exact Coupling

In this section, a weaker notion of exact coupling is studied. Suppose that (S, S^x, T) is an exact coupling of RW(e, μ) and RW(x, μ). If $\mathbf{P}(T < \infty) > 0$, then (S, S^x, T) is called a **possible exact coupling**. The difference between possible exact coupling and successful exact coupling is that a possible exact coupling only requires $T < \infty$ with positive probability, whereas a successful exact coupling would require $T < \infty$ a.s.

Definition 3.4.1. Define the **possible exact coupling set** G_p to be the subset of all $x \in G$ such that there exists a possible exact coupling of RW(e, μ) and RW(x, μ).

Carefully looking over the proofs in Section 3.1 reveals that in many places, the fact that a coupling time T satisfies $T < \infty$ a.s. is used only to guarantee that $\mathbf{P}(T = n) > 0$ for some n, allowing the same proofs work for possible exact couplings as well. In particular, the following variations on Proposition 3.1.5 and Theorem 3.1.6 hold without the need for any kind of assumption about the existence of large sets that commute with x.

Proposition 3.4.2. *Fix $x \in G$ and suppose that $n \geqslant 1$ is such that $\mu^n \geqslant \nu + \theta_x^{-1}\nu$ for a nonzero measure ν. Then $x \in G_p$ and there exists a possible exact coupling of* $\mathrm{RW}(e, \mu)$ *and* $\mathrm{RW}(x, \mu)$ *with a coupling time T satisfying $\mathbf{P}(T = n) = \nu(G)$.*

Proof. In the proof of Proposition 3.1.5, the only place where the assumption that there exists a B with $\mu^n(B) = 1$ such that x commutes with all of B is needed is to show that the constructed coupling time T is a.s. finite and T/n looks like a hitting time of a random walk. When this assumption is not met, the coupling from that proof still works, and the coupling time T still satisfies $\mathbf{P}(T = n) = \nu(G)$, but not necessarily $\mathbf{P}(T < \infty) = 1$, and T/n does not necessarily look like a hitting time of a random walk on $\langle x \rangle$. □

Theorem 3.4.3. *For all $x \in G$, there exists a possible exact coupling of* $\mathrm{RW}(e, \mu)$ *and* $\mathrm{RW}(x, \mu)$ *with a coupling time T satisfying $\mathbf{P}(T = n) > 0$ if and only if $\mu^n \wedge \theta_x^{-1}\mu^n \neq \mathbf{0}$. In particular, $G_p = \{x \in G : \exists n \geqslant 1, \mu^n \wedge \theta_x^{-1}\mu^n \neq \mathbf{0}\}$.*

Proof. The proof is nearly identical to that of Theorem 3.1.6, except one appeals to Proposition 3.4.2 to construct a possible exact coupling instead of Proposition 3.1.5. □

One may now reap some low-hanging fruit. In particular, it is shown that the possible exact coupling set is Borel measurable, that in the Abelian case admitting a possible exact coupling and admitting a successful exact coupling are the same, and that if an n-fold convolution of a measure overlaps with one of its shifts, then all higher-fold convolutions of the measure admit the same property.

Corollary 3.4.4. *G_p is Borel measurable.*

Proof. The set in question, by Theorem 3.4.3, equals $\bigcup_{n=1}^{\infty}\{y : \mu^n \wedge \theta_y^{-1}\mu^n \neq \mathbf{0}\}$, which is Borel measurable by Corollary 3.3.2. □

Corollary 3.4.5. *Suppose G is Abelian. Then $G_p = G_s$.*

Proof. By Theorems 3.4.3 and 3.2.1, both equal $\{x \in G : \exists n \geqslant 1, \mu^n \wedge \theta_x^{-1}\mu^n \neq \mathbf{0}\}$. □

Note that the previous corollary says that if an exact coupling with coupling time T satisfying $\mathbf{P}(T < \infty) > 0$ exists, then an exact coupling with coupling time T' with $\mathbf{P}(T' < \infty) = 1$ exists. It does not show that if $\mathbf{P}(T < \infty) > 0$ then $\mathbf{P}(T < \infty) = 1$.

Corollary 3.4.6. *For a probability measure ν on G, if $\nu^{n_0} \wedge \theta_x^{-1}\nu^{n_0} \neq \mathbf{0}$ for some $n_0 \geqslant 1$, then $\nu^n \wedge \theta_x^{-1}\nu^n \neq \mathbf{0}$ for all $n \geqslant n_0$.*

Proof. Let n_0 as above and let $n \geqslant n_0$ be given. By Theorem 3.4.3, choose a possible exact coupling (S, S^x, T) of RW(e, ν) and RW(x, ν) with $\mathbf{P}(T =$

$n_0) > 0$. Then $T' := T + (n - n_0)$ is also a coupling time for S and S^x with $\mathbf{P}(T' = n) > 0$, so by Theorem 3.4.3 it holds that $\nu^n \wedge \theta_x^{-1}\nu^n \neq \mathbf{0}$. □

In the Abelian case, admitting a possible exact coupling and admitting a successful exact coupling turned out to be the same. Lastly, it is shown that in the non-Abelian case this is not necessarily the case.

Example 3.4.7. Let $G := \mathbb{F}_2$ be the free group on two letters a, b and consider S and S^{ab} simple random walks on G. That is, the step-length distribution μ is supported on four atoms:

$$\mu(\{a\}) = \mu(\{a^{-1}\}) = \mu(\{b\}) = \mu(\{b^{-1}\}) = \frac{1}{4}. \tag{3.24}$$

Suppose S starts at the empty word e, and S^{ab} starts at ab. If S and S^{ab} are taken to be independent, then with positive probability $S_1 = a = S_1^{ab}$, so a possible exact coupling can be easily constructed. Furthermore, note that the length $\operatorname{len} S$ of S is itself a Markov chain on $\mathbb{N}$. In fact, with W denoting a simple random walk on $\mathbb{Z}$ having probability $1/4$ of decreasing and $3/4$ of increasing at each step, and which, for any $x \in \mathbb{Z}$, is started at x under a measure $\mathbf{P}_x$, one has

$$\mathbf{P}(\operatorname{len} S \text{ returns to } 0) = \mathbf{P}_1(W \text{ hits } 0) < 1, \tag{3.25}$$

where the last inequality is a standard fact about asymmetric simple random walks on $\mathbb{Z}$. It follows that 0 is a transient state for the Markov chain $\operatorname{len} S$ and, since the chain is irreducible, all states are transient. Hence a.s. the length of S tends to ∞ and a limiting word is finalized. A similar statement

holds for S^{ab}. Denote the limiting words $\lim S$ and $\lim S^{ab}$. Admitting a successful exact coupling is also equivalent, cf. Theorem 9.4 in Section 9.5 of [Tho00], to

$$\mathbf{P}(S \in B) = \mathbf{P}(S^{ab} \in B), \qquad B \in \mathcal{T}, \tag{3.26}$$

where $\mathcal{T}$ is the σ-algebra of tail measurable events. The set

$$\{s = (s_n)_{n=0}^{\infty} : \lim s \text{ starts with } b\}$$

is tail measurable. With τ the hitting time of e for S^{ab}, by the strong Markov property and the fact that at time τ it holds that S^{ab} starts anew as a copy of S,

$$\begin{aligned}\mathbf{P}(\lim S^{ab} \text{ starts with } b) &= \mathbf{P}(\tau < \infty)\mathbf{P}(\lim S \text{ starts with } b) \\ &= \mathbf{P}_2(W \text{ hits } 0)\mathbf{P}(\lim S \text{ starts with } b) \\ &< \mathbf{P}(\lim S \text{ starts with } b).\end{aligned}$$

Thus there is no successful exact coupling between S and S^{ab}.

3.5. Bibliographical Comments

The work in this chapter heavily relies on the inspiration of previous works. The initial idea for exact coupling comes from Doeblin's 1938 work [Doe38] on irreducible and aperiodic Markov chains on finite state spaces. In 1968, Ornstein [Orn69] provided an exact coupling that is successful for random walks on $\mathbb{Z}$ under aperiodicity assumptions. In 1979,

Berbee [Ber79] constructed an exact coupling that is successful for spread out random walks on $\mathbb{R}$. Interestingly, earlier in 1966 and using non-coupling methods, Stam [Sta66] determined necessary and sufficient conditions for random walks on $\mathbb{R}$ with either spread out or discrete step-lengths to have total variation bounds that are now known to be equivalent to admitting a successful exact coupling. The equivalence is deep in nature and work on it spans several papers. The equivalence is neatly presented in Thorisson's work [Tho00], see Theorem 9.4 in Chapter 4, Section 9.5. Coupling proofs of the spread out case and discrete case on $\mathbb{R}$ using more modern notation are given in [Tho00] and [Arn10] respectively, and the proofs in these two sources were the main starting points for the proof of the general case for any step-length distribution given in this chapter.

Chapter 4

Doeblin Trees

The first incarnation of the Propp and Wilson [PW96] coupling from the past (CFTP) algorithm was designed to build a perfect sample from the stationary distribution π of an irreducible and aperiodic Markov chain on a finite state space S. It uses a Doeblin-type coupling of a family of copies of the Markov chain started in all possible states at all possible times, whereby when two chains meet, they merge. This coupling is represented with a random directed graph on $\mathbb{Z} \times S$ depicting the trajectories of these Markov chains. Below, this random graph will be referred to as the **Doeblin graph** of the chain. It will be shown that the Doeblin graph admits a unimodularizable subgraph, and, when the graph is a tree, this fact is used to show the existence and uniqueness of a bi-infinite path in the graph that is recurrent both forwards and backwards in time, which is in-turn used to determine other properties of the Doeblin graph and of Markov chains in general, particularly in relation to this recurrence property.

Prior to this research, the study of this random graph has been mostly a by-product of research on perfect simulation. In 1992–1993, Borovkov and Foss [BF92, BF94] laid out the framework of stochastically recursive se-

quences (SRS), of which Markov chains are a special case, and they proved the main results on the existence of a stationary version of an SRS to which non-stationary versions converge in a certain sense. The CFTP algorithm itself was introduced by Propp and Wilson in 1996 in [PW96] for obtaining samples from the stationary distribution of a Markov chain. The CFTP algorithm can be seen as a specialization of the general ideas of [BF92] for SRS to the Markov case aiming at perfect simulation. Foss and Tweedie [FT98] then gave a necessary and sufficient condition for the CFTP algorithm to converge a.s. From 1996 to 2000, many papers [PW98a, MG98, PW98b, Ken98, Fil99, HN99, HvLM99, Møl99, KM00, Wil00, Men00] investigated how to improve CFTP implementations or how to apply CFTP or a CFTP-inspired algorithm to obtain a perfect sample from a particular Markov chain's stationary distribution. Of particular importance is Wilson's read-once CFTP algorithm [Wil00], which allows CFTP to be done by only simulating forwards in time. A review of perfect simulation in stochastic geometry up to that point is provided in [Møl01]. Since then, [Ken04, CK07] showed that (possibly impractical) generalizations of the CFTP algorithm can be applied under weaker conditions, and [FK03] gives a CFTP-like algorithm that applies even in the non-Markovian setting.

In this chapter, focus is shifted away from finding an individual sample from the stationary distribution of a Markov chain, and instead properties of the Doeblin graph as a whole are studied. The SRS framework will be used, but, because the Markov case is a fundamental special case,

most sections will spell out what can be said in the Markov case. The main tool of study is the theory of unimodular random (rooted) networks in the sense of Aldous and Lyons [AL07]. Unimodular networks are rooted networks where, heuristically, the root is picked uniformly at random. In order to generalize this concept for infinite networks, instead of picking the root uniformly at random, the network is required to satisfy a mass-transport principle. The primary new object of study is the subgraph of bridges between a fixed recurrent state, which is referred to as the **bridge graph** for this state and is roughly inspired by the population process in [BS18]. The subgraph is defined by looking at processes started at any time from this fixed state. General setup and definitions of the Doeblin graph and the bridge graph are given in Section 4.1 along with how to view the bridge graph as a random network. The main theorem is then proved in Section 4.2.

Section 4.2.1 proves the main theorem, identifying the unimodular structure in the bridge graph. Section 4.2.2 studies properties of the bridge graph that are inherited due to its I/F component structure as a unimodular network. Here I/F refers to the class of a component in the sense of the foil classification theorem in unimodular networks in [BHMK18], which was reviewed in Section 2.4. The most interesting case is when S is infinite and the Doeblin graph is connected. In this case (see Corollary 4.2.8), although there may be infinitely many bi-infinite paths in the Doeblin graph, there exists a unique **bi-recurrent** path, a bi-infinite path that visits every state infinitely often in the past, as well as in the future. This unique path also

has the property that the states in S that the path traverses form a stationary version of the original Markov chain (or SRS), and hence give samples from its stationary distribution. Indeed, the original CFTP algorithm ultimately computes the time zero point on the bi-recurrent path. By embedding Markov chains inside Doeblin graphs, bi-recurrence is also shown to be a decisive property for Markov chains indexed by $\mathbb{Z}$. Theorem 4.2.11 shows that if a Markov chain $(X_t)_{t\in\mathbb{Z}}$ has an irreducible, aperiodic, and positive recurrent transition matrix, then $(X_t)_{t\in\mathbb{Z}}$ is stationary if and only if it is bi-recurrent for any (and hence every) state. The I/F structure of a component leads to further useful qualitative properties discussed in Section 4.2.2. In reversed time, the bridge tree can be seen as a multi-type branching-like process where the types are the elements of S, and for which there is at most one child of each type per generation. The nodes in this branching process are either mortal (i.e., with finitely many descendants) or immortal (resp. infinitely many). The mortal descendants of the nodes on the bi-infinite path form a stationary sequence of finite trees. Mean values in these trees are linked to coupling times by mass-transport relations. Finally, Section 4.2.3 gives results that are relevant to simulating the bridge graph, such as approximating the bridge graph by finite networks, and viewing the process of vertical slices of the bridge graph as a Markov chain in its own right.

4.1. The Doeblin Graph

4.1.1. Definition

In this section, the Doeblin graph is constructed. Fix a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, a countable state space S, and a complete separable metric space Ξ for the remainder of the chapter. The first ingredient needed is a **pathwise transition generator**, a function $h_{\text{gen}} : S \times \Xi \to S$ that will be used for determining transitions between states of S. Such an h_{gen}, combined with a **driving sequence** $(\xi_t)_{t\in\mathbb{N}}$, is used to give a pathwise representation of a stochastic process $(X_t)_{t\in\mathbb{N}}$ satisfying

$$X_{t+1} := h_{\text{gen}}(X_t, \xi_t), \qquad t \geqslant 0. \tag{4.1}$$

Equation (4.1) is the defining property of a stochastically recursive sequence (SRS) in the sense of Borovkov and Foss [BF92]. If the driving sequence is taken to be i.i.d. and independent of X_0, then $(X_t)_{t\in\mathbb{N}}$ is a (discrete time) Markov chain with transition matrix $P = (p_{x,y})_{x,y\in S}$ determined by $p_{x,y} := \mathbf{P}(h_{\text{gen}}(x, \xi_0) = y)$ for each $x, y \in S$. It is a classical result that, when $\Xi := [0, 1]$, all possible transition matrices P can be achieved by choosing h_{gen} and the distribution of ξ_0 accordingly (c.f. Chapter 17 in [Bor13]). Many processes in this chapter will be indexed by $\mathbb{Z}$ or an interval of $\mathbb{Z}$ instead of just $\mathbb{N}$. The pathwise transition generator h_{gen} and a stationary and ergodic *bi-infinite* driving sequence $\xi := (\xi_t)_{t\in\mathbb{Z}}$, are fixed for the remainder of the chapter. The notation for the transition matrix $P = (p_{x,y})_{x,y\in S}$ is also fixed for the remainder of the chapter, even when ξ is not assumed to be i.i.d.

The space $\mathbb{Z} \times S$ should be thought of as time and space coordinates, with $(t, x) \in \mathbb{Z} \times S$ being in state x at time t. The vertices and edges of a graph Γ will be written $V(\Gamma)$ and $E(\Gamma)$, and if $V(\Gamma) \subseteq \mathbb{Z} \times S$, the vertices of Γ sitting at a particular time t or in a particular state x will be denoted, respectively, as

$$V_t(\Gamma) := \left\{(s, y) \in V(\Gamma) : s = t\right\}, \quad V^x(\Gamma) := \left\{(s, y) \in V(\Gamma) : y = x\right\}. \tag{4.2}$$

Note that $V_t(\Gamma)$ and $V^x(\Gamma)$ are subsets $\mathbb{Z} \times S$, i.e. their elements have both a time component and a space component. If instead just states (elements of S) or just times (elements of $\mathbb{Z}$) are desired, then the following are used instead

$$\Gamma_t := \left\{y \in S : (t, y) \in V_t(\Gamma)\right\}, \quad \Gamma^x := \{s \in \mathbb{Z} : (s, x) \in V^x(\Gamma)\}. \tag{4.3}$$

Then the **Doeblin graph** $\mathbf{G} = \mathbf{G}(h_{\text{gen}}, \xi)$ is constructed as follows. It has vertices $V(\mathbf{G}) := \mathbb{Z} \times S$. The edges of $\mathbf{G}$ are determined by the **follow map** $f_+ : V(\mathbf{G}) \to V(\mathbf{G})$, which is a random map giving directions of where each vertex should move to in the next time step. It is defined by

$$f_+(t, x) := (t + 1, h_{\text{gen}}(x, \xi_t)), \qquad (t, x) \in \mathbb{Z} \times S. \tag{4.4}$$

That is, let the edges of $\mathbf{G}$ be drawn from each $(t, x) \in \mathbb{Z} \times S$ to $f_+(t, x)$. By saying a function $f : A \to B$ is a random map, it is meant that $f : A \times \Omega \to B$ is measurable and the second argument will be omitted. Iterates of f_+ are denoted by f_+^n for $n \geqslant 0$. Thinking of each vertex in $\mathbf{G}$ as an individual, one may also interpret the follow map as mapping each vertex to its unique parent vertex.

4.1.2. Modeling

When dealing with paths in **G**, it will often be convenient to ignore the time coordinate and focus only on the space coordinate. If $(X_t)_{t\in I}$ is a stochastic process defined on Ω that takes values in S and is such that $(t, X_t)_{t\in I}$ is a.s. a path in **G** over some fixed time interval $I \subseteq \mathbb{Z}$, then $(X_t)_{t\in I}$ is called the **state path (in G)** corresponding to the path $(t, X_t)_{t\in I}$. That is, there are two ways of looking at every route through **G**: as a path $(t, X_t)_{t\in I} \subseteq V(\mathbf{G})$, or as a state path $(X_t)_{t\in I} \subseteq S$.

Lemma 4.1.1. *Let $I \neq \emptyset$ be an interval in $\mathbb{Z}$. Suppose that $(X_t)_{t\in I}$ is a stochastic process taking values in S a.s. satisfying the recurrence relation $X_{t+1} = h_{\text{gen}}(X_t, \xi_t)$ for each $\inf I \leqslant t < \sup I$, where h_{gen} and $(\xi_t)_{t\in I}$ are the same as are used to define* **G**. *Then $(X_t)_{t\in I}$ is a state path in* **G**.

Proof. One must check that $(t, X_t)_{t\in I}$ is a.s. a path in **G**. Fix $t \in I$. Since $V(\mathbf{G}) = \mathbb{Z} \times S$, (t, X_t) is certainly a vertex of **G**. If $t + 1 \in I$ as well, one must check the edge e from (t, X_t) to $(t + 1, X_{t+1})$ is a.s. an edge in **G**. The edges of **G** are defined to be from each $(t, x) \in \mathbb{Z} \times S$ to $(t + 1, h_{\text{gen}}(x, \xi_t))$, so the relation $X_{t+1} = h_{\text{gen}}(X_t, \xi_t)$ holding a.s. implies the edge e is a.s. an edge of **G**. □

In particular, Lemma 4.1.1 says that any SRS whose driving sequence is defined for all times in $\mathbb{Z}$ can be seen as living inside a Doeblin graph, namely the one generated by its driving sequence and choosing h_{gen} to be the same as in the definition of the SRS.

State paths started at a deterministic vertex will also be used heavily. For the remainder of the chapter, let $F^{(t,x)} := (F_s^{(t,x)})_{s \geqslant t} \subseteq S$ be the **state path in G started at time** t **in state** x, i.e., $F^{(t,x)}$ is a re-indexing of the states traversed by f_+ defined by

$$(s, F_s^{(t,x)}) = f_+^{s-t}(t,x), \qquad (t,x) \in \mathbb{Z} \times S, \quad s \geqslant t. \tag{4.5}$$

One has that $F^{(t,x)}$ is a version of the SRS or Markov chain started in state x with initial condition given at time t. Generally speaking, throughout the chapter, a parenthesized superscript, as in $F^{(t,x)}$, refers to a starting location. For every $x \in S$, the distribution of $(F_{s+t}^{(t,x)})_{s \geqslant 0}$ does not depend on t because ξ is stationary. An example of a Doeblin graph and path of $F^{(t,x)}$ are drawn in Figure 4.1.

It has already been noted (see [Bor13]) that a Markov chain $(X_t)_{t \in \mathbb{N}}$ with any given desired transition matrix can be constructed as an SRS with i.i.d. driving sequence. The following is an analogous result saying that any Markov chain $(X_t)_{t \in \mathbb{Z}}$ may be realized as a state path in a Doeblin graph with i.i.d. driving sequence. Note here that the time index set is all of $\mathbb{Z}$, not just $\mathbb{N}$.

Theorem 4.1.2. *Suppose that $(X_t)_{t \in \mathbb{Z}}$ is a Markov chain with transition matrix P on some probability space, where P is the same as was defined for the Doeblin graph* **G**. *Also suppose the driving sequence ξ is i.i.d. Then there is a probability space $(\Omega', \mathcal{F}', \mathbf{P}')$ and $(X'_t)_{t \in \mathbb{Z}} \sim (X_t)_{t \in \mathbb{Z}}$ on Ω' such that $(X'_t)_{t \in \mathbb{Z}}$ is state path in* $\mathbf{G}'$, *where* $\mathbf{G}'$ *is the Doeblin graph generated by some i.i.d. driving sequence*

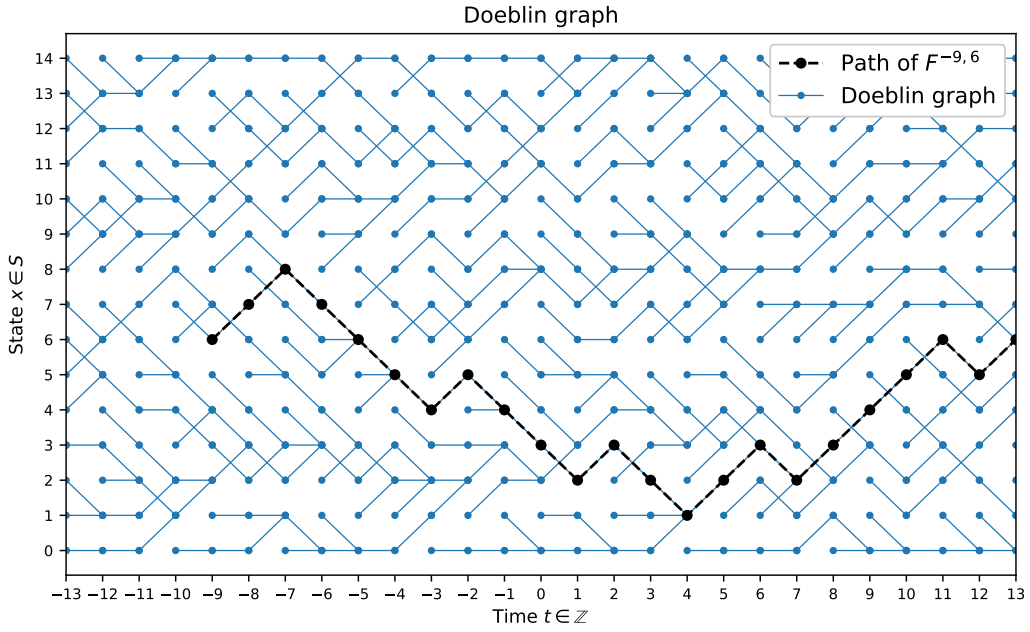


Figure 4.1: An example of a Doeblin graph with the path corresponding to the state path $F^{(t,x)}$ distinguished. All edges are directed from left to right.

$\xi' = (\xi'_t)_{t\in\mathbb{Z}} \sim \xi$ *in* Ω' *with pathwise transition generator* h_{gen}. *Moreover, for each* $t \in \mathbb{Z}$, X'_t *is independent of* $(\xi'_s)_{s\geqslant t}$.

Proof (sketch). Consider a probability space housing independent copies of $(X_t)_{t\in\mathbb{Z}}$ and **G**. Then consider for each $t \in \mathbb{Z}$ the state path in **G** started at X_t. The distributions of these state paths determine a consistent set of finite dimensional distributions for the desired pair of processes $((X'_t)_{t\in\mathbb{Z}}, (\xi'_t)_{t\in\mathbb{Z}})$. By the Kolmogorov extension theorem, the result follows. The full proof of Theorem 4.1.2 is given in Appendix A.

4.1.3. Basic Properties

Plainly, **G** is acyclic as an undirected graph because all outgoing edges point forward one unit in time and each vertex has only one outgoing edge. When **G** is a.s. connected, it is called a **Doeblin Eternal Family Tree** or a **Doeblin EFT** for short. More generally, **G** may have up to countably many components and is referred to as a **Doeblin Eternal Family Forest** or **Doeblin EFF**. The EFT and EFF terminology is inspired by [BHMK18] and the word eternal refers to the fact that every vertex of **G** has a unique outgoing edge. That is, every individual has exactly one parent. An EFF is a more general object than an EFT, i.e. an EFF may also be an EFT.

If the driving sequence ξ is i.i.d., so that the state paths $F^{(t,x)}$ for each $(t,x) \in \mathbb{Z} \times S$ are Markov chains, then say that **G** is **Markovian**. If ξ is such that for each $t \in \mathbb{Z}$, $(f_+(t,x))_{x\in S}$ is an independent family, then **G** is said to have **vertical independence**. If **G** is Markovian and has vertical independence, then say that **G** has **fully independent transitions**.

Some later results are only valid for EFTs, so the following result gives an easy case when **G** can be shown to be connected.

Proposition 4.1.3. *Suppose* **G** *has fully independent transitions, and P is irreducible and positive recurrent with period d. Then a.s.* **G** *has d components. In particular, if P is irreducible, aperiodic, and positive recurrent, then* **G** *is an EFT.*

Proof (sketch). The case of a general d is reduced to $d = 1$ by viewing the chain only every d steps and with state space restricted to one of the d

classes appearing in a cyclic decomposition of the state space. Consider the state paths in $\mathbf{G}$ started at $(0, x)$ and $(0, y)$ for any two x, y. Strictly before hitting the diagonal, the pair of state paths has the same distribution as a product chain, i.e. two independent copies of the chain with one started at x and the other at y. The product chain is irreducible, aperiodic, and positive recurrent, and therefore a.s. hits the diagonal, showing the state paths started at $(0, x)$ and $(0, y)$ eventually merge. The full proof of Proposition 4.1.3 is given in Appendix A.

A ξ-measurable subgraph $\Gamma = \Gamma((\xi_t)_{t\in\mathbb{Z}})$ of $\mathbf{G}$ is called **shift-covariant** if, for all $s \in \mathbb{Z}$, $\Gamma((\xi_{t+s})_{t\in\mathbb{Z}})$ is a.s. the time-translation of Γ by $-s$. Say a state path $(X_t)_{t\in\mathbb{Z}}$ is **shift-covariant** if the corresponding path in $\mathbf{G}$ is shift-covariant. In other words, if the driving sequence ξ is translated by some amount s in time, then shift-covariant objects are also translated in time by the same amount. Let $E \in \mathcal{F}$ be ξ-measurable, say $1_E = g((\xi_t)_{t\in\mathbb{Z}})$. Say that E is **shift-invariant** if $g((\xi_t)_{t\in\mathbb{Z}}) = g((\xi_{t+1})_{t\in\mathbb{Z}})$ a.s. That is, shift-invariant events are those events whose occurrence is unaffected by time translations of the driving sequence ξ. One has that $\mathbf{P}(E) \in \{0, 1\}$ for all shift-invariant events E due to the ergodicity of ξ. All of the following are shift-invariant and hence happen with probability zero or one: $\mathbf{G}$ is locally finite, $\mathbf{G}$ contains no cycles, $\mathbf{G}$ is connected, $\mathbf{G}$ has exactly $n \in \mathbb{N}\cup\{\infty\}$ components, $\mathbf{G}$ contains exactly $n \in \mathbb{N} \cup \{\infty\}$ bi-infinite paths. Generally it will be obvious whether an event is shift-invariant.

When **G** is a Markovian, one needs to be cautious that not all state paths in **G** are Markov chains with transition matrix P.

Example 4.1.4. Let $S := \mathbb{Z}$ and suppose **G** has fully independent transitions with $p_{x,x-1} = p_{x,x} = p_{x,x+1} = \frac{1}{3}$ for all $x \in S$. Choose X_0 to be the smallest element of $\mathbb{Z}$ (in some well-ordering of $\mathbb{Z}$) such that $F_1^{(0,X_0)} = F_2^{(0,X_0)}$. In this case, a.s. $X_1 = X_2$, so $(X_t)_{t\in\mathbb{N}}$ is not even Markovian.

The problem with the path in the previous example is that it looks into the future. Namely, the value of X_0 depends on information at time 1 and time 2. To exclude state paths like those in Example 4.1.4, the notion of properness is introduced. For a nonempty interval I of $\mathbb{Z}$, if for each $t \in I$, X_t is independent of $(\xi_s)_{s\geqslant t}$, then $(X_t)_{t\in I}$ is called a **proper** state path. In the Markovian case, if I has a minimum element t_0, then to show that a state path $(X_t)_{t\in I}$ is proper it is sufficient that X_{t_0} is independent of $(\xi_s)_{s\geqslant t_0}$ because for any $s \in \mathbb{N}$, X_{t_0+s} is measurable with respect to the σ-algebra generated by X_{t_0} and $\xi_{t_0}, \ldots, \xi_{t_0+s-1}$. Unlike general state paths in **G**, proper state paths inherit a Markov transition structure.

Lemma 4.1.5. *Suppose* **G** *is Markovian. If* $(X_t)_{t\in I}$ *is a proper state path in* **G** *over a nonempty interval* $I \subseteq \mathbb{Z}$*, then* $(X_t)_{t\in I}$ *is a Markov chain with transition matrix* P*.*

Proof. Fix $t < \sup I$. Let $E := \{X_t = x_t, \ldots, X_{t-k} = x_{t-k}\}$ be given with $k \in \mathbb{N}$ such that $t - k \geqslant \inf I$, and $x_t, \ldots, x_{t-k} \in S$. Note that whether E occurs is

a function of X_{t-k} and $\xi_{t-k}, \ldots, \xi_{t-1}$, so the fact that X_{t-k} is independent of $(\xi_s)_{s \geqslant t-k}$ and the fact that ξ is i.i.d. imply that E is independent of $(\xi_s)_{s \geqslant t}$. Then for any $x \in S$,

$$\begin{aligned}
\mathbf{E}[1_{\{X_{t+1}=x\}} 1_E] &= \mathbf{E}[1_{\{h(x_t,\xi_t)=x\}} 1_E] \\
&= \mathbf{P}(h(x_t, \xi_t) = x)\mathbf{P}(E) \\
&= p_{x_t,x}\mathbf{P}(E) \\
&= \mathbf{E}[p_{X_t,x} 1_E].
\end{aligned}$$

Since E was an arbitrary cylinder set, it follows that for all $x \in S$,

$$\mathbf{P}(X_{t+1} = x \mid (X_s)_{s \in I, s \leqslant t}) = p_{X_t,x}.$$

Thus $(X_s)_{s \in I}$ is a Markov chain with transition matrix P. □

4.1.4. Connections with CFTP

Consider the following structural result that will be expanded upon in Section 4.2.2. It is a special case of Proposition 4.2.7 and Corollary 4.2.8, which will be proved later.

Proposition 4.1.6. *Suppose* **G** *is Markovian, and that P is irreducible, aperiodic, and positive recurrent. Then a.s. in every component of* **G** *there exists a unique bi-infinite path that visits every state in S infinitely often in the past. All other bi-infinite paths in* **G** *do not visit any state infinitely often in the past. If* **G** *is an EFT, then with β_t denoting the state at time t of the unique bi-infinite path visiting every state infinitely often in the past, one has that $(\beta_t)_{t \in \mathbb{Z}}$ is a stationary Markov*

chain with transition matrix P, so that $\beta_t \sim \pi$ for all $t \in \mathbb{Z}$, where π is the invariant distribution for P.

The main result of the original Propp and Wilson paper can be translated into the language of Doeblin EFFs and summarized as follows. The reader is encouraged to ponder what it says about the structure of $\mathbf{G}$, and in doing so one sees that is has much the same spirit as Proposition 4.1.6.

Proposition 4.1.7 (Perfect Sampling [PW96])**.** *If S is finite and $\mathbf{G}$ is Markovian and an EFT (which necessitates that P is irreducible and aperiodic), then there is an a.s. finite time τ such that all paths in $\mathbf{G}$ started at any time $t \leqslant -\tau$ have merged by time 0, all reaching a common vertex $(0, \beta_0)$. Moreover, $\beta_0 \sim \pi$, where π is the stationary distribution of P, and there is an algorithm A that a.s. terminates in finite time returning β_0.*

Remark 4.1.8. *In fact, the β_0 appearing in Proposition 4.1.7 and the β_0 appearing in Proposition 4.1.6 are the same. That is, the perfect sampling algorithm A is ultimately computing the point in $\mathbf{G}$ on the unique bi-infinite path and returning its state. This can be seen by the fact that, since all paths started at time $-\tau$ reach the common vertex $(0, \beta_0)$, any bi-infinite path in $\mathbf{G}$ must also pass through $(0, \beta_0)$. However, what is notably absent in Proposition 4.1.6 is any mention of an algorithm to compute β_0. Whether such an algorithm exists in general is not studied in the present research.*

4.1.5. Bridge Graphs

The primary tool used in this chapter will be the theory of unimodular networks in the sense of [AL07]. Local finiteness is essential in the theory of unimodular networks, but the Doeblin graph $\mathbf{G}$ may not be locally finite, as the following result shows.

Proposition 4.1.9. *If $\sum_{x\in S} p_{x,y} < \infty$ for all $y \in S$, then $\mathbf{G}$ is a.s. locally finite. If $\mathbf{G}$ has fully independent transitions and for some $y \in S$, $\sum_{x\in S} p_{x,y} = \infty$, then $\mathbf{G}$ is a.s. not locally finite.*

Proof. Both statements follow from the Borel-Cantelli lemmas. That is, for any fixed $(t, y) \in \mathbb{Z} \times S$, if $\sum_{x\in S} p_{x,y} < \infty$, then a.s. one has that only finitely many of the events $\left\{f_+(t-1, x) = (t, y)\right\}_{x\in S}$ occur, showing (t, y) has finite in-degree, and hence finite degree, in $\mathbf{G}$. On the other hand, if $\mathbf{G}$ has fully independent transitions and for some fixed $(t, y) \in \mathbb{Z}\times S$ one has $\sum_{x\in S} p_{x,y} = \infty$, then a.s. infinitely many of the events $\left\{f_+(t-1, x) = (t, y)\right\}_{x\in S}$ occur, so that (t, y) has infinite degree. □

The remedy taken here is to instead concentrate on particular subgraphs of $\mathbf{G}$. In this section, subgraphs are introduced that are locally finite under a positive recurrence assumption and turn out to have nice properties when considered as random networks.

For each $(t, x) \in \mathbb{Z} \times S$, and each $y \in S$, let

$$\tau^{(t,x)}(y) := \inf\left\{s > t : F_s^{(t,x)} = y\right\}, \qquad \sigma^{(t,x)}(y) := \tau^{(t,x)}(y) - t \tag{4.6}$$

be, respectively, the **return time** and **time until return** of $F^{(t,x)}$ to y. The word return is used even when $y \neq x$, in which case it may be that $F^{(t,x)}$ is not part of a state path that has visited y before time t. Note that the distribution of $\sigma^{(t,x)}(y)$ does not depend on t because ξ is stationary. Call a state $x \in S$ **positive recurrent** if $\mathbf{E}[\sigma^{(0,x)}(x)] < \infty$ or **recurrent** if $\sigma^{(0,x)}(x) < \infty$ a.s. In the Markovian case these are the usual definitions. If a state $x \in S$ is recurrent, then indeed for every $t \in \mathbb{Z}$, $F^{(t,x)}$ visits x infinitely often.

For each fixed $x \in S$, consider the subgraph $\mathbf{B}(x)$ of $\mathbf{G}$ of all paths starting from state x at any time. That is, $\mathbf{B}(x)$ is the subgraph of $\mathbf{G}$ with

$$V(\mathbf{B}(x)) := \bigcup_{t\in\mathbb{Z}} \left\{ (s, F_s^{(t,x)}) : s \geqslant t \right\} = \bigcup_{t\in\mathbb{Z}} \left\{ f_+^n(t,x) : n \geqslant 0 \right\}. \tag{4.7}$$

Call $\mathbf{B}(x)$ the **bridge graph for state** x and refer to it as either a **bridge EFF** or **bridge EFT** depending on whether it is a forest or a tree. Note that one of these possibilities happens with probability 1 because the number of components in $\mathbf{B}(x)$ is shift-invariant.

Assumption 4.1.10. *For the remainder of the chapter, assume there exists a positive recurrent state $x^* \in S$, which is fixed, and the notation $\mathbf{B} := \mathbf{B}(x^*)$ refers to the bridge graph for state x^*.*

An example bridge graph appears in Figure 4.2. Equivalently, $\mathbf{B}$ can be described in terms of descendants of vertices, viewing directed edges in $\mathbf{G}$ as pointing from a vertex to its parent. For each $(t, y) \in \mathbb{Z} \times S$, define the

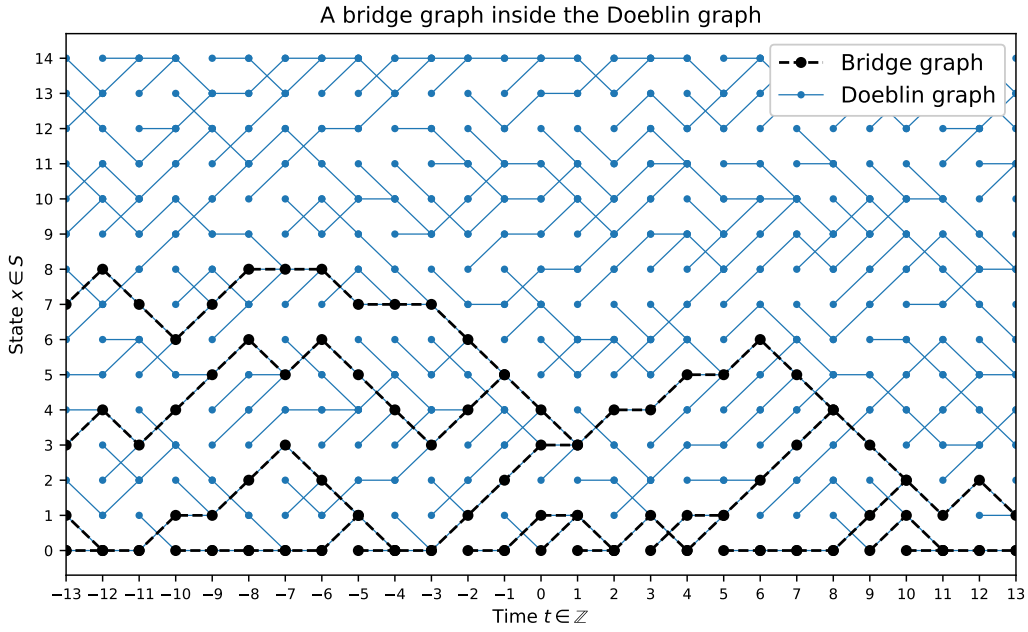


Figure 4.2: An example bridge graph, in this case for state $x^* = 0$, sitting inside the Doeblin graph.

descendants of (t, y) in $\mathbf{G}$ to be

$$D^{(t,y)} := \left\{(s, x) \in \mathbb{Z} \times S : F_t^{(s,x)} = y\right\}. \tag{4.8}$$

Then $\mathbf{B}$ is also the subgraph of $\mathbf{G}$ with

$$V(\mathbf{B}) = \left\{(t, y) \in \mathbb{Z} \times S : \exists s,\, (s, x^*) \in D^{(t,y)}\right\}.$$

That is, $\mathbf{B}$ is the subgraph of $\mathbf{G}$ generated by vertices that have some descendant in state x^*. In particular, recalling the notation (4.3),

$$y \in \mathbf{B}_t \iff \exists s,\, x^* \in D_s^{(t,y)}, \qquad (t, y) \in \mathbb{Z} \times S. \tag{4.9}$$

Lemma 4.1.11 shows that if **G** is a.s. connected, then **B** is too.

Lemma 4.1.11. *If $u, v \in V(\mathbf{B})$ are in the same component of* **G***, then they are in the same component of* **B***. In particular, if* **G** *is an EFT, then* **B** *is an EFT.*

Proof. Consider times $s, t \in \mathbb{Z}$. Suppose (s, x^*) and (t, x^*) are in the same component of **G**. Then $F^{(s,x^*)}$ and $F^{(t,x^*)}$ meet at some point. But, by definition, the paths of $F^{(s,x^*)}$ and $F^{(t,x^*)}$ are included in **B**. Hence (s, x^*) and (t, x^*) are in the same component of **B**. Now if $u, v \in V(\mathbf{B})$ are in the same component of **G**, u is in the same component in **G** as some (s, x^*) and v is in the same component of **G** as some (t, x^*), and (s, x^*) and (t, x^*) are in the same component of **B** by the previous part. Hence u, v are in the same component of **B**. □

The condition that **B** is an EFT is equivalent to strong coupling convergence (defined and studied in [BF92, BF94, FK03]) of $F^{(0,x^*)}$ to a stationary version of the SRS. However, simple conditions for **B** to be an EFT are not known outside of the fully i.i.d. Markovian case, where Proposition 4.1.3 showed that if P is irreducible, aperiodic, and positive recurrent, then **G** is an EFT. Another (not necessarily easy to check) condition for **B** to be an EFT will be given in Corollary 4.2.9.

The main tool used in this chapter is unimodularity of random networks. The first form of unimodularity used is stationarity, i.e., the unimodularity of the deterministic network $\mathbb{Z}$ rooted at 0 and with neighboring integers connected. Unimodularity of $\mathbb{Z}$ gives a helpful way to reorganize

proofs based on stationarity in terms of transporting mass between different times. Recall that a (measurable) group action $\theta : \mathbb{Z} \times \Omega \to \Omega$ of $\mathbb{Z}$ on Ω is called $\mathbf{P}$-invariant if $\mathbf{P}(\theta_t \in \cdot) = \mathbf{P}$ for all $t \in \mathbb{Z}$. The shift operator on $\Xi^{\mathbb{Z}}$ is an example of such an action.

Lemma 4.1.12 (Mass-transport Principle for $\mathbb{Z}$)**.** *Suppose $w : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}_{\geqslant 0}$ is a random map. Also suppose $\theta : \mathbb{Z} \times \Omega \to \Omega$ is a $\mathbf{P}$-invariant $\mathbb{Z}$-action on Ω, and that the two are compatible in the sense that $w(s,t) \circ \theta_r = w(s+r, t+r)$ almost surely for each $s, t, r \in \mathbb{Z}$. Then with $w^+ := \sum_{t \in \mathbb{Z}} w(0,t)$ and $w^- := \sum_{s \in \mathbb{Z}} w(s,0)$, one has*

$$\mathbf{E}[w^+] = \mathbf{E}[w^-]. \tag{4.10}$$

Proof. One calculates

$$\mathbf{E}[w^+] = \sum_{t \in \mathbb{Z}} \mathbf{E}[w(0,t)] = \sum_{t \in \mathbb{Z}} \mathbf{E}[w(0,t) \circ \theta_{-t}] = \sum_{t \in \mathbb{Z}} \mathbf{E}[w(-t,0)] = \mathbf{E}[w^-]$$

as desired. □

The mass-transport principle for $\mathbb{Z}$ immediately gives the following.

Proposition 4.1.13. *For all $t \in \mathbb{Z}$, $\mathbf{E}[\#\mathbf{B}_t] \leqslant \mathbf{E}[\sigma^{(0,x^*)}(x^*)]$. In particular, $\mathbf{B}$ is a.s. locally finite, even if $\mathbf{G}$ itself is not.*

Proof. Without loss of generality, Ω is the canonical space $\Xi^{\mathbb{Z}}$, with the driving sequence $(\xi_t)_{t \in \mathbb{Z}}$ being coordinate maps. Then $\theta : \mathbb{Z} \times \Omega \to \Omega$ defined by $\theta_s((\xi_t)_{t \in \mathbb{Z}}) := (\xi_{s+t})_{t \in \mathbb{Z}}$ is a $\mathbf{P}$-invariant measurable $\mathbb{Z}$-action on

Ω. Choose the mass-transport $w(s,t) := 1_{\{\sigma^{(s,x^*)}(x^*)>t-s>0\}}$. The fact that one has $\sigma^{(s,x^*)}(x^*) \circ \theta_r = \sigma^{(s+r,x^*)}(x^*)$ for all $s,t,r \in \mathbb{Z}$ implies w is compatible with θ. Then $w^+ = \sigma^{(0,x^*)}(x^*) - 1$, and $w^- = \#\left\{s < 0 : \sigma^{(s,x^*)}(x^*) > |s|\right\} \geqslant \#\mathbf{B}_0 - 1$, where this inequality follows from the fact that for every $y \in \mathbf{B}_0 \setminus \{x^*\}$, there is $s < 0$ such that $\sigma^{(s,x^*)}(x^*) > |s|$ and $F_0^{(s,x^*)} = y$. Thus the mass-transport principle for $\mathbb{Z}$ gives $\mathbf{E}[\sigma^{(0,x^*)}(x^*) - 1] \geqslant \mathbf{E}[\#\mathbf{B}_0 - 1]$, from which the result follows. □

The proof style of Proposition 4.1.13 may be repeated in many different ways and the boilerplate setup of the proof can be mostly omitted once one understands the flow of the proof. The shortened version of the proof of Proposition 4.1.13 is given to exemplify how much can be omitted without losing the main idea.

Proof (shortened). Let the mass-transport $w(s,t)$ send mass 1 from s to all times t strictly after s and strictly before $F^{(s,x^*)}$ returns to x^*. Then $w^+ = \sigma^{(0,x^*)}(x^*) - 1$ and $w^- = \#\left\{s < 0 : \sigma^{(s,x^*)}(x^*) > |s|\right\} \geqslant \#\mathbf{B}_0 - 1$, where this inequality follows from the fact that for every $y \in \mathbf{B}_0 \setminus \{x^*\}$, there is $s < 0$ such that $\sigma^{(s,x^*)}(x^*) > |s|$ and $F_0^{(s,x^*)} = y$. The mass-transport principle finishes the claim. □

One now sees the versatility of using even the simplest form of unimodularity. A list of mass-transports and the results they give, all by following the same proof style, appears in Section A.2 in Appendix A. Some of the mass-transports give new results, and others recover well-known results,

such as fact that $\pi(y)/\pi(x^*)$ is the expected number of visits of a Markov chain started at x^* to y before returning to x^*, and $1/\pi(x^*)$ is the expected return time of a Markov chain started at x^* to return to x^*, where π is the invariant distribution for the Markov chain. The next section shows how to embed subgraphs of $\mathbf{G}$ as random networks, so that eventually one may find a unimodular structure inside $\mathbf{G}$.

4.1.6. Embedding Subgraphs of the Doeblin Graph as Random Networks

In order to view a subgraph of $\mathbf{G}$ as a random network, one must ensure the subgraph is nonempty, locally finite, connected, and a root $\boldsymbol{o}$ has been suitably chosen. Since the vertices of $\mathbf{G}$ come from the fixed countable space $\mathbb{Z}\times S$, the following setup will help to verify all the technicalities. Before proceeding, the reader may wish to recall the notation used for random networks that was established in Section 2.4.

Let $V := \mathbb{Z} \times S$ and suppose that

$$\mathbf{\Gamma} : \Omega \to \{0,1\}^V \times \Xi^V \times \{0,1\}^{V\times V} \times \Xi^{V\times V} =: (f_V, \xi_V, f_E, \xi_E) \tag{4.11}$$

is measurable (where the codomain is given its product topology and corresponding Borel σ-algebra). Then $\mathbf{\Gamma}(\omega)$ can be considered for each $\omega \in \Omega$ as a (possibly empty, possibly not locally finite, possibly disconnected) network in the following way. For each $u, v \in V$, interpret:

(i) $f_V(v)$ as the indicator that $v \in V(\mathbf{\Gamma})$,

(ii) $f_E(u, v)$ as the indicator that the edge $\{u, v\} \in E(\mathbf{\Gamma})$,

(iii) $\xi_V(u)$ as the mark of u, and

(iv) $\xi_E(u, v)$ as the mark of the vertex-edge pair $(u, \{u, v\})$.

That is, use items (i) to (iv) to *define* $V(\mathbf{\Gamma}), E(\mathbf{\Gamma})$, and the marks of vertices and edges in $\mathbf{\Gamma}$. Note that f_E must be symmetric because edges are not directed, but ξ_E may not be, since each edge is associated with two marks, one per vertex. If one wants to consider directed edges, one instead uses undirected edges and uses the marks on edges to specify which direction the edge should point. The definition of $\xi_V(u)$ when $u \notin V(\mathbf{\Gamma})$ is irrelevant, and similarly for the definition of $f_E(u, v)$ and $\xi_E(u, v)$ if either of u or v is not in $V(\mathbf{\Gamma})$. All statements about the network defined by $\mathbf{\Gamma}$ are then translated into statements about the maps (f_V, ξ_V, f_E, ξ_E). For instance,

$$\{\mathbf{\Gamma} \text{ is not empty}\} = \left\{ \sum_{v \in V} f_V(v) > 0 \right\}.$$

This is exactly the kind of construction used to define the Doeblin graph $\mathbf{G}$. In the case of $\mathbf{G}$,

(i) $f_V = 1$ on $\mathbb{Z} \times S$,

(ii) $f_E\left((t, x), (t+1, h(x, \xi_t))\right) = f_E\left((t+1, h(x, \xi_t)), (t, x)\right) = 1$ for all $(t, x) \in \mathbb{Z} \times S$ and $f_E = 0$ otherwise,

(iii) $\xi_V(t, x) = \xi_t$ for all $(t, x) \in \mathbb{Z} \times S$, and

(iv) $\xi_E((t,x),(t+1,h(x,\xi_{t,x}))) = 1$ for all $t \in \mathbb{Z}, x \in S$ to indicate the edge is directed forwards in time.

This construction also works for the bridge graph $\mathbf{B}$ as well. When a $\mathbf{\Gamma}$ has been constructed as in this section, one can see $\mathbf{\Gamma}$ as a random network after any measurable choice of root, given that it is nonempty and locally finite.

Lemma 4.1.14. *Suppose $\mathbf{\Gamma} = (f_V, \xi_V, f_E, \xi_E)$ is as above and a.s. $\mathbf{\Gamma}$ is nonempty, locally finite, and connected. Then for any measurable choice of root $\boldsymbol{o} \in V(\mathbf{\Gamma})$, $[\mathbf{\Gamma}, \boldsymbol{o}]$ is a random network.*

Proof (sketch). Write the event that $[\mathbf{\Gamma}, \boldsymbol{o}]$ is within $\epsilon > 0$ of some fixed network $[\Gamma, o]$ as a countable union over rooted isomorphic copies (Γ', o') of (Γ, o) with vertices in V of the event that $\boldsymbol{o} = o'$, the neighborhood of radius $\lceil \frac{1}{\epsilon} \rceil$ around $\boldsymbol{o}$ is exactly Γ', and the marks $\xi_V(u)$ for $u \in V(\Gamma')$ and $\xi_E(v,w)$ for $\{v,w\} \in E(\Gamma')$ are within ϵ of the corresponding vertex and edge marks of (Γ', o'). Each of these conditions individually are written in terms of events using the maps f_V, ξ_V, f_E, ξ_E, showing the desired measurability of $\omega \mapsto [\mathbf{\Gamma}(\omega), \boldsymbol{o}(\omega)]$. The full proof of Lemma 4.1.14 given in Appendix A.

Thus indeed $\mathbf{G}$ may be seen as a random network when rooted and marked, assuming it is locally finite and connected. But the question remains whether this may be done in such a way as to make $\mathbf{G}$ unimodular. The first approach one might take is to investigate whether $\mathbf{G}$, rooted at $(0, X_0)$ for some (random) choice of $X_0 \in S$, is unimodular. Two natural choices,

at least in the standard CFTP setup, are to take X_0 to be the output of the CFTP algorithm, or to take X_0 to be independent of $\mathbf{G}$. For simplicity, the "standard" CFTP setup refers to the case where $\mathbf{G}$ has fully independent transitions, S is finite, and the CFTP algorithm succeeds a.s. The following proposition determines when $\mathbf{G}$ can be unimodular under the previous choices of X_0.

Proposition 4.1.15. *Suppose $\mathbf{G}$ is an EFT, that $\mathbf{G}$ has each $(t, x) \in V(\mathbf{G})$ marked by (x, ξ_t), and that X_0 is a random choice in S. Then the following hold:*

(a) If $[\mathbf{G}, (0, X_0)]$ is unimodular, then S is finite and X_0 is uniformly distributed on S.

(b) If X_0 is independent of $\mathbf{G}$ and uniformly distributed on a finite S, then $[\mathbf{G}, (0, X_0)]$ is unimodular.

(c) If X_0 is the output of the CFTP algorithm in the standard CFTP setup, then $[\mathbf{G}, (0, X_0)]$ is unimodular if and only if S has a single element.

Proof (sketch). The first point follows by constructing for each $x, y \in S$ a mass-transport that, when applied to $\mathbf{G}$, sends mass 1 within vertical slices of $\mathbf{G}$ from the vertex in state x to the vertex in state y. Unimodularity then gives $\mathbf{P}(X_0 = x) = \mathbf{P}(X_0 = y)$. The second point follows from the definition of unimodularity. The third point follows by noting that the output of the CFTP algorithm has at least one child, but unimodularity implies that it must have one on average, so a.s. it has one child. A nonempty tree where

every vertex has one incoming and one outgoing edge is isomorphic to $\mathbb{Z}$, so S can only have one state. The full proof of Proposition 4.1.15 is given in Appendix A.

While choosing X_0 uniformly distributed on S and independent of $\mathbf{G}$ works when S is finite, unimodularity of the whole $\mathbf{G}$ is doomed in the general case, as there is no uniform distribution on an infinite S. This is the reason for introducing the bridge graph $\mathbf{B}$, which is locally finite. However, the bridge graph may still not be connected, so a spine is added to it to make it connected.

Corollary 4.1.16. *Let* $\overline{\mathbf{B}}$ *be* $\mathbf{B}$ ***with spine added****, i.e. with edges from each* (t, x^*) *to* $(t+1, x^*)$ *for all* $t \in \mathbb{Z}$ *added. Then for any measurable marks and any measurable choice of root* $\boldsymbol{o} \in V(\mathbf{B})$*,* $[\overline{\mathbf{B}}, \boldsymbol{o}]$ *is a random network.*

Proof. One has that $(0, x^*) \in V(\overline{\mathbf{B}})$, so $\overline{\mathbf{B}}$ is nonempty. Also $\overline{\mathbf{B}}$ is locally finite by Proposition 4.1.13 and the fact that adding the spine has increased the degree of each vertex by at most two. Finally, since each $v \in V(\overline{\mathbf{B}})$ is connected to some (t, x^*), and the spine in $\overline{\mathbf{B}}$ connects all such vertices, $\overline{\mathbf{B}}$ is connected. Lemma 4.1.14 finishes the claim. □

Everything is in place to see the unimodular structure hidden in $\mathbf{G}$, which is handled in the next section.

4.2. Unimodularizability and its Consequences

4.2.1. Unimodularizability of the Bridge Graph

The following result identifies the unimodular structure inside $\mathbf{G}$. For the rest of the chapter, each $(t, y) \in V(\mathbf{B})$ is marked by (y, ξ_t) whenever considered as a vertex in a rooted network.

Theorem 4.2.1. *Any random network with distribution*

$$\mathbf{P}^{\square}(A) := \frac{1}{\mathbf{E}[\#\mathbf{B}_0]}\mathbf{E}\left[\sum_{w\in V_0(\mathbf{B})} 1_{\{[\overline{\mathbf{B}},w]\in A\}}\right], \qquad A \in \mathcal{B}(\mathcal{G}_*). \tag{4.12}$$

is unimodular. The spine need not be added and $\mathbf{B}$ *may also be used instead of* $\overline{\mathbf{B}}$ *if* $\mathbf{B}$ *is already connected.*

One may interpret the distribution $\mathbf{P}^{\square}$ as a size-biased version of the network obtained by starting with $\mathbf{B}$ and selecting the root uniformly from $\mathbf{B}_0$.

Proof. By Corollary 4.1.16, $\overline{\mathbf{B}}$ with marks as specified and any choice of root is a random network. Therefore, all the quantities in the following calculation are measurable. Let $g : \mathcal{G}_{**} \to \mathbb{R}_{\geqslant 0}$ be given. One has

$$\begin{aligned}
&\int_{\mathcal{G}_*} \sum_{v\in V(\Gamma)} g[\Gamma, o, v]\, \mathbf{P}^{\square}(d[\Gamma, o]) \\
&= \frac{1}{\mathbf{E}[\#\mathbf{B}_0]}\mathbf{E}\sum_{y\in\mathbf{B}_0}\sum_{v\in V(\mathbf{B})} g[\overline{\mathbf{B}}, (0, y), v] \\
&= \frac{1}{\mathbf{E}[\#\mathbf{B}_0]}\sum_{y,y'\in S, t\in\mathbb{Z}} \mathbf{E}\left[1_{\{(0,y),(t,y')\in V(\mathbf{B})\}} g[\overline{\mathbf{B}}, (0, y), (t, y')]\right].
\end{aligned}$$

Stationarity on $\mathbb{Z}$ implies the right hand side is equal to

$$\begin{aligned}
&\frac{1}{\mathbf{E}[\#\mathbf{B}_0]} \sum_{y,y' \in S, t \in \mathbb{Z}} \mathbf{E}\left[1_{\{(-t,y),(0,y') \in V(\mathbf{B})\}} g[\overline{\mathbf{B}}, (-t,y), (0,y')]\right] \\
=&\frac{1}{\mathbf{E}[\#\mathbf{B}_0]} \sum_{y,y' \in S, t \in \mathbb{Z}} \mathbf{E}\left[1_{\{(t,y),(0,y') \in V(\mathbf{B})\}} g[\overline{\mathbf{B}}, (t,y), (0,y')]\right] \\
=&\frac{1}{\mathbf{E}[\#\mathbf{B}_0]} \mathbf{E} \sum_{y' \in \mathbf{B}_0} \sum_{v \in V(\mathbf{B})} g[\overline{\mathbf{B}}, v, (0,y')] \\
=&\int_{\mathcal{G}_*} \sum_{v \in V(\Gamma)} g[\Gamma, v, o]\, \mathbf{P}^{\square}(d[\Gamma, o]).
\end{aligned}$$

Thus $\mathbf{P}^{\square}$ is the distribution of a unimodular network. □

The view of $\mathbf{P}^{\square}$ as a size-biased version of a network is formalized in the following.

Proposition 4.2.2. *Let $\boldsymbol{o}$ be, conditionally on $V_0(\mathbf{B})$, uniformly distributed on $V_0(\mathbf{B})$ and independent of $\mathbf{B}$. Then under the size-biased measure $\hat{\mathbf{P}}(E) := \frac{1}{\mathbf{E}[\#\mathbf{B}_0]}\mathbf{E}[\#\mathbf{B}_0 1_E]$ for each $E \in \mathcal{F}$, the random network $[\overline{\mathbf{B}}, \boldsymbol{o}]$ has the distribution $\mathbf{P}^{\square}$.*

Proof. In what follows, V ranges over the sets for which $\mathbf{P}(V_0(\mathbf{B}) = V) > 0$, of which there are at most countably many because $\mathbf{B}_0$ is a.s. a finite subset of the countable S. For any $A \in \mathcal{B}(\mathcal{G}_*)$ and with $C := \mathbf{E}[\#\mathbf{B}_0]$,

$$\begin{aligned}
&\hat{\mathbf{P}}([\overline{\mathbf{B}}, \boldsymbol{o}] \in A) \\
&= \frac{1}{C}\mathbf{E}[\#\mathbf{B}_0 1_{\{[\overline{\mathbf{B}}, \boldsymbol{o}] \in A\}}] \\
&= \frac{1}{C} \sum_V |V|\, \mathbf{P}(V_0(\mathbf{B}) = V)\mathbf{P}([\overline{\mathbf{B}}, \boldsymbol{o}] \in A \mid V_0(\mathbf{B}) = V) \\
&= \frac{1}{C} \sum_V \sum_{v \in V} |V|\, \mathbf{P}(V_0(\mathbf{B}) = V)\mathbf{P}(\boldsymbol{o} = v, [\overline{\mathbf{B}}, v] \in A \mid V_0(\mathbf{B}) = V)
\end{aligned}$$

which, by the conditional independence of $\boldsymbol{o}$ and $\mathbf{B}$, is

$$\begin{aligned}
&= \frac{1}{C}\sum_{V}\sum_{v\in V}|V|\,\mathbf{P}(V_0(\mathbf{B})=V)\frac{1}{|V|}\mathbf{P}([\overline{\mathbf{B}},v]\in A \mid V_0(\mathbf{B})=V)\\
&= \frac{1}{C}\mathbf{E}\left[\sum_{V}\sum_{v\in V}1_{\{[\overline{\mathbf{B}},v]\in A, V_0(\mathbf{B})=V\}}\right]\\
&= \mathbf{P}^{\square}(A)
\end{aligned}$$

as claimed. □

4.2.2. I/F Component Properties

For any measurable event $A \subseteq \mathcal{G}_*$ in the σ-algebra of **root-invariant** events, i.e., such that if $[\Gamma, o] \in A$ then $[\Gamma, v] \in A$ for all $v \in V(\Gamma)$, one has

$$\mathbf{P}^{\square}(A) = \frac{1}{\mathbf{E}[\#\mathbf{B}_0]}\mathbf{E}\left[\sum_{w\in V_0(\mathbf{B})}1_{\{[\overline{\mathbf{B}},w]\in A\}}\right] = \frac{1}{\mathbf{E}[\#\mathbf{B}_0]}\mathbf{E}\left[\#\mathbf{B}_0 1_{\{[\overline{\mathbf{B}},(0,x^*)]\in A\}}\right].$$

This immediately gives the following.

Lemma 4.2.3. *One has that* $\mathbf{P}^{\square}$ *and* $\mathbf{P}([\overline{\mathbf{B}},(0,x^*)]\in\cdot)$ *have the same root-invariant sets of measure* 0 *or* 1. □

Next, a vertex-shift that is designed to follow the arrows in $\mathbf{B}$ is defined. It plays the same role as f_+ but is defined for all networks. For the rest of the chapter, let Φ denote the **follow vertex-shift** defined on any network Γ for each $u \in V(\Gamma)$ by $\Phi_\Gamma(u) := v$ if either:

(i) there is a unique outgoing edge from u and this edge terminates at v, or

(ii) u is in state x^* and there is a unique outgoing edge from u that does not terminate at a vertex in state x^*, and this edge terminates at v.

If neither of the two conditions above is met for any $v \in V(\Gamma)$, define $\Phi_\Gamma(u) := u$ for concreteness. Here a vertex is considered to be in a state $y \in S$ when the first component of its mark is y (recall that a vertex $(t, y) \in V(\mathbf{B})$ is marked by (y, ξ_t)). The second clause in the definition of Φ is there because of the presence of the spine in $\overline{\mathbf{B}}$, so that if the root is in state x^* the vertex-shift will choose to follow the arrow in $\mathbf{B}$ instead of following the arrow to the next element of the spine, unless the two coincide. By construction, $\Phi_{\overline{\mathbf{B}}}(t, x) = f_+(t, x)$ for all $(t, x) \in V(\overline{\mathbf{B}})$.

The event that all Φ-components of a network are of class I/F is root-invariant, and moreover it has $\mathbf{P}([\overline{\mathbf{B}}, (0, x^*)] \in \cdot)$-probability one because the Φ-graph of $\overline{\mathbf{B}}$ is $\mathbf{B}$ itself, the Φ-components of $\overline{\mathbf{B}}$ are the components of $\mathbf{B}$, and the Φ-foils of $\overline{\mathbf{B}}$ are subsets of the sets $(V_t(\mathbf{B}))_{t\in\mathbb{Z}}$, which are finite. Hence $\mathbf{P}^\square$ is concentrated on the set of networks having only Φ-components of I/F class. It follows that any a.s. root-invariant properties that follow from $\mathbf{P}^\square$ being unimodular and having I/F components automatically apply to $\mathbf{P}([\overline{\mathbf{B}}, (0, x^*)] \in \cdot)$ as well. Such properties will be referred to as **I/F component properties** and are explored in the rest of the section.

Bi-recurrent Paths

This section studies bi-infinite paths in $\mathbf{G}$ and identifies special bi-infinite paths that have a certain recurrence property backwards in time.

Firstly, it is possible to have multiple bi-infinite paths in **G** because **G** is disconnected.

Example 4.2.4. Consider the case where $S := \{1, 2\}$ and h_{gen} and $(\xi_t)_{t\in\mathbb{Z}}$ are chosen so that the transition $(t, 1) \to (t + 1, 2)$ occurs if and only if $(t, 2) \to (t + 1, 1)$ occurs. In this case **G** has two components a.s. Each component is itself a bi-infinite path.

Moreover, even when **G** is connected, it it still possible to have multiple bi-infinite paths in **G**.

Example 4.2.5. Consider the case of $S := \mathbb{N}$ with fully independent transitions. Let the transition matrix P be determined as follows. In state 0, transition to a Geom(1/2) random variable, and from any other $n \neq 0$, deterministically transition from n to $n - 1$. In this case, from every vertex $(s, x) \in \mathbb{Z} \times S$, there is a bi-infinite path $(t, X_t)_{t\in\mathbb{Z}}$ in **G** for which $X_{s-k} = k + x$ for all $k \geqslant 0$. Thus there are infinitely many bi-infinite paths, despite the fact that in this case **G** is an EFT, which follows from Proposition 4.1.3.

In Example 4.2.5, even though **G** is connected, **G** has infinitely many bi-infinite paths. However, amongst the bi-infinite paths, there is one special bi-infinite path. The special path is the unique bi-infinite path that visits every state infinitely often *in the past*. It turns out that this is the correct kind of path to look for in general.

Definition 4.2.6. A bi-infinite sequence $(x_t)_{t\in\mathbb{Z}}$ in S is called **bi-recurrent for state** x if $\{t \in \mathbb{Z} : x_t = x\}$ is unbounded above and below. If $(x_t)_{t\in\mathbb{Z}}$ is bi-recurrent for every $x \in S$, it is simply called **bi-recurrent**. A state path $(X_t)_{t\in I}$ in $\mathbf{G}$ is called **bi-recurrent (for state** x**)** if a.s. its trajectory is bi-recurrent (for state x).

Recall that Φ denotes the follow vertex-shift. The existence of bi-infinite paths in Φ-components of a network is an I/F property, and hence one has the following.

Proposition 4.2.7. *It holds that* $\mathbf{B}$ *has a unique bi-infinite path in each component a.s. The corresponding state paths are bi-recurrent for* x^* *and these are the only state paths in all of* $\mathbf{G}$ *that are bi-recurrent for* x^**. Moreover, for each* $y \in S$*, these state paths either a.s. never visit* y*, or are bi-recurrent for* y*.*

Proof. By Theorem 2.5.1, $\mathbf{P}^{\square}$-a.e. network has a unique bi-infinite path in each Φ-component, where Φ is the follow vertex-shift. But having a unique bi-infinite path in each Φ-component is a root-invariant event, and hence $\mathbf{P}$-a.s. $\overline{\mathbf{B}}$ has a unique bi-infinite path in each Φ-component. Since the Φ-components of $\overline{\mathbf{B}}$ are the components of $\mathbf{B}$, $\mathbf{P}$-a.s. every component of $\mathbf{B}$ contains a unique bi-infinite path.

Let Π be the covariant partition of Φ-components. Define the covariant subset C on a network Γ by letting C_Γ be the subset of vertices of Γ that are either the first or last visit to a given state $y \in S$, if they exist, on the unique bi-infinite path in their Φ-component of Γ, if such a path exists. The

no infinite/finite inclusion lemma, Lemma 2.5.2, implies that $\mathbf{P}^{\square}$ is concentrated on the set of networks Γ with no first or last visit to y on the unique bi-infinite paths in each Φ-component of Γ. This property is root-invariant and hence a.s. the state paths corresponding to the unique bi-infinite paths in each component of $\mathbf{B}$ either do not visit state y or are bi-recurrent for y. Taking a countable union over $y \in S$ shows this property holds simultaneously for all $y \in S$. Since the unique bi-infinite path in each component of $\mathbf{B}$ at least hits x^*, one may at least conclude the paths are bi-recurrent for x^*. Finally, there cannot be any other bi-recurrent state paths for x^* in $\mathbf{G}$ because, by definition, a bi-recurrent state path in $\mathbf{G}$ will lie in $\mathbf{B}$ since it visits x^* at arbitrarily large negative times. $\square$

The next result applies Proposition 4.2.7 to the nicest case, where $\mathbf{G}$ is a tree.

Corollary 4.2.8. *Suppose that $\mathbf{G}$ is an EFT. Then $\mathbf{G}$ contains a unique (up to measure zero modifications) state path $(\beta_t)_{t\in\mathbb{Z}}$ that is bi-recurrent for x^*. Moreover, there is a version of $(\beta_t)_{t\in\mathbb{Z}}$ that is shift-covariant, stationary, and for each $t \in \mathbb{Z}$ one has that β_t is measurable with respect to $\sigma(\xi_s : s < t)$. Additionally, $(\beta_t)_{t\in\mathbb{Z}}$ is bi-recurrent for every $x \in S$ that is positive recurrent.*

Proof. Proposition 4.2.7 shows that a.s. there is a unique bi-infinite path in each component of $\mathbf{B}$, and the corresponding state paths are bi-recurrent for x^*. Since $\mathbf{G}$ a.s. has only one component, $\mathbf{B}$ does too. The second part of Proposition 4.2.7 then implies the bi-recurrent state path for x^* in

B is the only bi-recurrent state path for x^* in **G**. One would like to define $(\beta_t)_{t\in\mathbb{Z}}$ to be the unique bi-recurrent state path for x^* in **G**. However, in that case, $(\beta_t)_{t\in\mathbb{Z}}$ would only be defined a.s. For concreteness, define β_t for each $t \in \mathbb{Z}$ by letting $\beta_t := \lim_{s\to-\infty} F_t^{(s,x^*)}$ on the event that the limit exists, and $\beta_t := x^*$ otherwise. The limit here is in the discrete topology, so that convergence means eventually constant. On the a.s. event E that **B** is connected, $\#\mathbf{B}_t < \infty$ for all $t \in \mathbb{Z}$, and there is a unique bi-infinite path in **B**, one has that $(t, \beta_t)_{t\in\mathbb{Z}}$ coincides with the unique bi-infinite path in **B**. This is because if, for some $t \in \mathbb{Z}$, $\lim_{s\to-\infty} F_t^{(s,x^*)}$ does not exist, then either $\#\mathbf{B}_t = \infty$, or there exist two states $x, y \in S$ such that (t, x) and (t, y) have (necessarily disjoint) locally finite infinite trees of descendants in **B**. The former case is forbidden on E, and, in the latter case, König's lemma would imply the existence of two distinct bi-infinite paths in **B**, which is also forbidden on E. Thus $\lim_{s\to-\infty} F_t^{(s,x^*)}$ exists for all $t \in \mathbb{Z}$ on the event E, and, on this event, the unique bi-infinite path in **B** must therefore be $(t, \beta_t)_{t\in\mathbb{Z}}$. The shift-covariance and hence stationarity of $(\beta_t)_{t\in\mathbb{Z}}$ follows from its definition in terms of $F^{(s,x^*)}$ for each $s \in \mathbb{Z}$. For each $t \in \mathbb{Z}$, measurability of β_t with respect to $\sigma(\xi_s : s < t)$ also follows from its definition, since each $F_t^{(r,x^*)}$ with $r \leqslant t$ is $\sigma(\xi_s : s < t)$-measurable.

Now let $(Y_t)_{t\in\mathbb{Z}}$ be the unique bi-recurrent state path for some other $y \in S$ that is positive recurrent. Since **G** is a.s. connected, $(\beta_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$ eventually merge, a.s. However, stationarity forbids that there is a first time such that $\beta_t = Y_t$, so it must be that $\beta_t = Y_t$ for all $t \in \mathbb{Z}$. Thus $(\beta_t)_{t\in\mathbb{Z}}$ is

bi-recurrent for every $y \in S$ that is positive recurrent. □

Corollary 4.2.8 shows that, like in the standard CFTP setup, there is a β_0 living at time 0 in $\mathbf{G}$ that is a perfect sample from the stationary distribution of the Markov chain or SRS. However, unlike in the standard CFTP setup, it is not known whether there is an algorithm that can find β_0 in finite time.

Another consequence of the existence of bi-recurrent paths in $\mathbf{B}$ is that one can bound the number of components of $\mathbf{B}$.

Corollary 4.2.9. *The a.s. constant number n of components of $\mathbf{B}$ is no larger than* $\min\{k : \mathbf{P}(\#\mathbf{B}_0 = k) > 0\} < \infty$. *In particular, $\mathbf{B}$ has finitely many connected components, even if $\mathbf{G}$ has infinitely many components, and if $\mathbf{P}(\#\mathbf{B}_0 = 1) > 0$, then $\mathbf{B}$ is an EFT.*

Proof. The number of components of $\mathbf{B}$ is shift-invariant and hence a.s. constant. Each component of $\mathbf{B}$ contains a bi-recurrent path by Proposition 4.2.7. Each bi-recurrent path intersects $V_0(\mathbf{B})$ in a different element since they are in different components of $\mathbf{B}$. It follows that $n \leqslant \#\mathbf{B}_0$ a.s. If $\mathbf{P}(\#\mathbf{B}_0 = k) > 0$ for some k, then it follows that $n \leqslant k$. □

The deterministic cycle on n states shows that the bound in Corollary 4.2.9 can be achieved for each n. In general, any bi-infinite stationary process on S (or any countable set) must be bi-recurrent.

Proposition 4.2.10. *Suppose that $(X_t)_{t\in\mathbb{Z}}$ is a stationary process taking values in S. Then a.s. $(X_t)_{t\in\mathbb{Z}}$ is bi-recurrent for every $x \in \{X_t\}_{t\in\mathbb{Z}}$.*

Proof. For each $x \in S$, stationarity forbids that there is a first or last visit of $(X_t)_{t\in\mathbb{Z}}$ to x since such an occurrence would have to be equally likely to happen at all times $t \in \mathbb{Z}$. Thus, a.s. either $x \notin \{X_t\}_{t\in\mathbb{Z}}$ or $\{t \in \mathbb{Z} : X_t = x\}$ must be unbounded both above and below. The countability of S finishes the claim. □

The remainder of the section specializes to the Markovian setting again. In the Markovian setting, bi-recurrence is actually equivalent to stationarity in the irreducible, aperiodic, positive recurrent case.

Theorem 4.2.11. *Suppose that P is irreducible, aperiodic, and positive recurrent, and that $(X_t)_{t\in\mathbb{Z}}$ is a Markov chain with transition matrix P. Then $(X_t)_{t\in\mathbb{Z}}$ is stationary if and only if it is bi-recurrent for any (and hence every) state.*

Proof. By Theorem 4.1.2, it is possible to assume without loss of generality that $(X_t)_{t\in\mathbb{Z}}$ is a state path in the Doeblin graph **G** with fully independent transitions. By Proposition 4.1.3, **G** is an EFT and therefore Corollary 4.2.8 implies that **G** contains a bi-recurrent state path $(\beta_t)_{t\in\mathbb{Z}}$ that is, for all $y \in S$, the a.s. unique bi-recurrent state path for state y in **G**. Moreover, $\beta_t \sim \pi$ for all $t \in \mathbb{Z}$, where π is the stationary distribution for P. If $(X_t)_{t\in\mathbb{Z}}$ is bi-recurrent for some $y \in S$, then, by uniqueness, $X_t = \beta_t$ for all $t \in \mathbb{Z}$, a.s. In particular, $(X_t)_{t\in\mathbb{Z}}$ is stationary. The converse follows from Proposition 4.2.10 and irreducibility. □

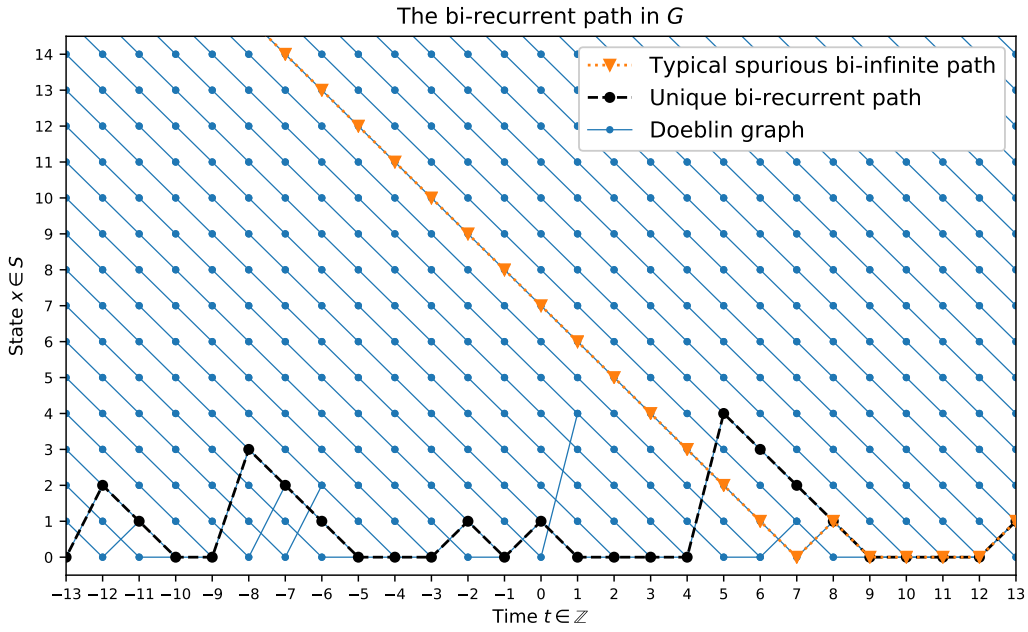


Figure 4.3: The Doeblin graph from Example 4.2.5 with the bi-recurrent path and a spurious path distinguished.

A bi-infinite path in **G** whose state path is not bi-recurrent for any state $x \in S$ will be called **spurious**. Observe the difference between spurious bi-infinite paths and the unique bi-recurrent path in Figure 4.3. Viewed in reverse time, a spurious path must run off to ∞ in the sense that for every finite set $F \subseteq S$, the reversed path eventually leaves F forever. It is possible for **G** to contain spurious bi-infinite paths, as was seen in Example 4.2.5.

Say that P^n **converges uniformly (to π as $n \to \infty$)** if P is irreducible, aperiodic, and positive recurrent with stationary distribution π, and $\sup_{x \in S} \|P^n(x, \cdot) - \pi\|_{\mathrm{TV}} \to 0$ as $n \to \infty$. For example, this is automatic if P is irreducible, aperiodic, and S is finite. Some authors call P uniformly

ergodic, but the term ergodic is not used here to avoid a terminology collision with ergodic theory. Uniform convergence to π is also equivalent (cf. [MT09] Theorem 16.0.2 (v)) to the statement that there is m such that $P^m(x, \cdot) \geqslant \varphi(\cdot)$ for all $x \in S$, for a measure φ which is not the zero measure. It is also equivalent (cf. [FT98] Theorem 4.2) to the fact that the CFTP algorithm succeeds in the case of fully independent transitions, i.e. the backwards vertical coupling time $\inf\left\{t > 0 : F_0^{(-t,x)} = F_0^{(-t,y)}, \forall x, y \in S\right\}$ is a.s. finite.

Together, the following two results say that when a Markov chain that mixes uniformly is started in the infinite past, it has converged to its stationary distribution by any finite time.

Proposition 4.2.12. *Suppose P^n converges uniformly to π as $n \to \infty$ and* **G** *has fully independent transitions. Then* **G** *contains no spurious bi-infinite paths.*

Proof. For every $s < t$ let $C_{s,t}$ be the event that $F_t^{(s,x)} = F_t^{(s,y)}$ for all $x, y \in S$. That is, $C_{s,t}$ is the event that starting at time s, all paths in **G** collapse to a single state by time t. Note that $\mathbf{P}(C_{s,t})$ depends only on $t - s$. Since **G** has fully independent transitions and P^n converges to π uniformly as $n \to \infty$, by e.g. Theorem 5.2 in [FT98], there exists some $k \in \mathbb{N}$ such that $\mathbf{P}(C_{s,t}) > 0$ when $t - s \geqslant k$. Consider $E_n := C_{-k(n+1),-kn}$ for each $n \in \mathbb{N}$. One has $\mathbf{P}(E_n) = \mathbf{P}(E_0) > 0$ for all n and the E_n are independent. It follows that a.s. infinitely many of them occur. On an ω for which infinitely many E_n occur, there is at most one bi-infinite path in **G**, and thus any bi-infinite path in **G** must coincide with the unique bi-recurrent path guaranteed to exist

by Corollary 4.2.8. □

It is a classical result that it is possible to find a bi-infinite stationary version $(X_t)_{t\in\mathbb{Z}}$ of a Markov chain that has a stationary distribution. The following shows that, in the case of uniform convergence to π, this is the only way to extend a Markov chain to have time index set all of $\mathbb{Z}$. That is, if $(X_t)_{t\in\mathbb{Z}}$ is a Markov chain that converges uniformly to its stationary distribution, then it must be that $X_t \sim \pi$ for all $t \in \mathbb{Z}$.

Proposition 4.2.13. *Suppose P^n converges uniformly to π as $n \to \infty$. Then every Markov chain $(X_t)_{t\in\mathbb{Z}}$ with transition matrix P is stationary and bi-recurrent. The subtle assumption here is that the time index set is all of $\mathbb{Z}$.*

Proof. By Theorem 4.1.2, one may assume $(X_t)_{t\in\mathbb{Z}}$ is a state path in $\mathbf{G}$ with fully independent transitions, which is then an EFT by Proposition 4.1.3. Since P^n converges uniformly to π as $n \to \infty$, $\mathbf{G}$ contains no spurious bi-infinite paths by Proposition 4.2.12, and hence $(X_t)_{t\in\mathbb{Z}}$ must be the bi-recurrent state path. Theorem 4.2.11 then implies $(X_t)_{t\in\mathbb{Z}}$ is stationary. □

Proposition 4.2.13 may fail for an irreducible, aperiodic, and positive recurrent P if P does not converge uniformly to its stationary distribution. Indeed, it was already shown, e.g., in Example 4.2.5, that it is possible for $\mathbf{G}$ to admit spurious bi-infinite paths. If $(X_t)_{t\in\mathbb{Z}}$ is a proper state path in $\mathbf{G}$ that corresponds to a spurious bi-infinite path, then $(X_t)_{t\in\mathbb{Z}}$ is a Markov chain with transition matrix P, but it is not stationary since it is not bi-recurrent.

Recall that $\mathbf{B}(x)$ denotes the bridge graph in $\mathbf{G}$ using x as the base point instead of x^*.

Proposition 4.2.14. *Suppose P is irreducible, aperiodic, and positive recurrent, and that $\mathbf{G}$ has fully independent transitions. If*

(i) S is infinite,

(ii) $\mathbf{G}$ is locally finite, and

(iii) $\mathbf{G}$ contains no spurious bi-infinite paths,

then

$$\bigcap_{x \in S} V(\mathbf{B}(x)) = \left\{(t, \beta_t) : t \in \mathbb{Z}\right\}, \tag{4.13}$$

where $(\beta_t)_{t \in \mathbb{Z}}$ is the unique bi-recurrent state path in $\mathbf{G}$. That is, the bi-recurrent path in $\mathbf{G}$ is the only thing common to all of the bridge EFTs. Alternatively, if S is finite and has at least 2 states, then a.s.

$$\bigcap_{x \in S} V(\mathbf{B}(x)) \supsetneq \left\{(t, \beta_t) : t \in \mathbb{Z}\right\}. \tag{4.14}$$

Proof. For each $x \in S$, the bi-recurrent path is in $\mathbf{B}(x)$ because it is bi-recurrent for x. Suppose S is infinite, $\mathbf{G}$ is locally finite, and that $\mathbf{G}$ contains no spurious bi-infinite paths. Consider a vertex $v \in V(\mathbf{G})$ not on the bi-recurrent path. The tree of all descendants of v in $\mathbf{G}$ must be finite, else König's lemma would give a bi-infinite path in $\mathbf{G}$ that is distinct from the unique bi-recurrent path since v is not on the bi-recurrent path. Since $\mathbf{G}$

contains no spurious bi-infinite paths, this is impossible. Since the tree of descendants of v is finite but S is infinite, there is some state $x \in S$ such that v has no descendant in state x. In particular, $v \notin V(\mathbf{B}(x))$, showing that nothing off the bi-recurrent path can be common to all the bridge EFTs.

Next suppose that $2 \leqslant \#S < \infty$. It suffices to give a finite deterministic graph Γ that is a subgraph of $\mathbf{G}$ with positive probability such that when some time-translate of Γ is a subgraph of $\mathbf{G}$, $\bigcap_{x \in S} V(\mathbf{B}(x))$ contains a vertex not on the unique bi-infinite path in $\mathbf{G}$. Firstly, since S is finite, choose a tree T on $\mathbb{Z} \times S$ that occurs with positive probability and is an example witnesses of the a.s. finiteness of the backwards vertical coupling time $\inf\left\{t > 0 : F_0^{(-t,x)} = F_0^{(-t,y)}, \forall x, y \in S\right\}$. Suppose T is rooted at $(0, x_0)$. In particular, $V(T) \subseteq (-\infty, 0] \times S$. By irreducibility of P and the fact that $\#S \geqslant 2$, choose $L = (x_0, x_1, \ldots, x_n)$ a finite path in S using only positive probability transitions from x_0 back to $x_0 = x_n$ that passes through all states of S and has the property that $x_i \neq x_{i+1}$ for any i. Note that

$$L_0 := \{(t, x_t) : t = 0, \ldots, n\}, \qquad L_1 := \{(t+1, x_t) : t = 0, \ldots, n\} \tag{4.15}$$

do not intersect. Moreover, L_0 and T intersect only at the vertex $(0, x_0)$, and L_1 and T do not intersect. Let Γ be the union of T, L_0, and L_1. The edges of T, L_0, and L_1 all occur with positive probability in $\mathbf{G}$, and none of them have the same initial vertex, so that in fact they are comprised of independent edges in $\mathbf{G}$. Since Γ has only a finite number of edges, it follows that $\Gamma \subseteq \mathbf{G}$ occurs with positive probability. Moreover, when $\Gamma \subseteq \mathbf{G}$ occurs,

the vertex $(n, x_{n-1}) \in V(\mathbf{B}(x))$ for all $x \in S$, but it is not on the bi-infinite path. This is because, by construction, $(0, x_0)$ is on the bi-infinite path in **G** and therefore L_0 makes up a segment of the bi-infinite path in **G**. But, L_1 includes a representative for every state, so for every $x \in S$ there is an $s \in \mathbb{Z}$ such that $x \in D_s^{(n,x_{n-1})}$. Finally, $V(L_0) \cap V(L_1) = \emptyset$ so (n, x_{n-1}) is not on the bi-infinite path in **G**. □

Other I/F Component Properties

The existence and uniqueness of a bi-infinite path in each Φ-component of a network is one I/F property that was studied at length in Section 4.2.2, which centered around bi-recurrent paths in **B**. However, there are many other potential things to say about **B** following from its I/F structure. A few of them are discussed in this brief section.

The first is the general structure of a network with only I/F components. Each component of **B** contains a unique bi-infinite path. Points on a bi-infinite path are sometimes referred to as **immortals** due to the fact that they do not disappear after an infinite number of applications of the follow vertex-shift Φ. A component **evaporates** if each point disappears after a finite number (depending on the point) of applications of Φ. Thus, in the case of **B**, none of the components evaporate. **Mortals** are those points in $V(\mathbf{B})$ that do disappear after a finite number of applications of Φ, i.e. those that have only finitely many descendants. Each component of **B** contains a bi-infinite path of immortals, and each immortal has exactly one child who

is immortal. Thus the immortals within a component are ordered like $\mathbb{Z}$ in a shift-covariant way. Hanging off of each immortal is then a (possibly empty) tree of mortals, the descendants of the immortal who are not themselves immortal and whose closest immortal ancestor is the given immortal. With this viewpoint, each component of $\mathbf{B}$ can be seen as a shift-covariant bi-infinite sequence of finite rooted trees, where each immortal is the root of its tree. If there is only one component of $\mathbf{B}$, then it has already been noted that there is a unique bi-infinite path in $\mathbf{B}$ whose state path $(\beta_t)_{t\in\mathbb{Z}}$ is stationary. However, more can be said in this case. If there is only one component of $\mathbf{B}$, then in fact the whole sequence $([Q_t, (t, \beta_t)])_{t\in\mathbb{Z}}$ is stationary, where Q_t is the tree hanging from the immortal (t, β_t). It is important here that the isomorphism class of Q_t is used and each vertex $(t, y) \in V(\mathbf{B})$ is marked with (y, ξ_t), otherwise the sequence would not be stationary due to the strictly increasing time coordinate. This view of $\mathbf{B}$ as a joining of trees gives an alternative way of looking at $\mathbf{B}$ compared to the view of $\mathbf{B}$ as a union of bridges between x^* at different times. Yet another viewpoint is that of $\mathbf{B}$ as a sequence of vertical slices. This idea has already been explored slightly in that the way the root was chosen in the definition of the unimodular measure $\mathbf{P}^\square$ is by choosing a root from one of these vertical slices. The view of $\mathbf{B}$ as a sequence of vertical slices is explored more in Section 4.2.3.

Additionally, the list of mass-transports given in Appendix A gives some integrability results relating these three viewpoints. In particular, in each way of viewing $\mathbf{B}$ there is a natural way to split $\mathbf{B}$ into pieces. In the view

of **B** as a joining of a sequence of trees of mortals hanging off an immortal, the vertices are partitioned by which tree they are in. In the view of **B** as a sequence of vertical slices, the vertices are partitioned by which slice they are in. In the view of **B** as paths started from state x^*, vertices are partitioned by the time they first return to x^*. In fact, the mass-transport arguments given in Appendix A show that the mean number of vertices in a partition element is the same for all three viewpoints. See the list of mass-transports in Appendix A for a more detailed description of these results and other finer-grained results.

4.2.3. Applications to Simulating the Bridge Graph

Local Weak Convergence to the Bridge Graph

It was shown in Proposition 4.2.2 that the measure $\mathbf{P}^{\square}$ may be thought of as an appropriately size-biased version of a network with the root picked uniformly at random from individuals at time 0. A common reason for size-biasing to show up is when picking uniformly at random across a population and asking the size of the group an individual is in. Picking uniformly at random is what unimodularity models, so one might expect that a unimodular network can be approximated by picking the root uniformly at random from a very large but finite sub-network. At present, whether all unimodular networks can be approximated in this way is an open problem [AL07]. In the case of the unimodular bridge EFF, it will be shown directly that indeed it can be approximated by finite sub-networks with a root picked uniformly

at random.

In this section, different ways of approximating the unimodular version of $\mathbf{B}$ by finite subgraphs are considered. Recall that $\overline{\mathbf{B}}$ denotes $\mathbf{B}$ with spine added, i.e. $\mathbf{B}$ with edges connecting each (t, x^*) to $(t+1, x^*)$. For a finite interval $I \subseteq \mathbb{Z}$ define $V_I(\mathbf{B}) := \bigcup_{t \in I} V_t(\mathbf{B})$ and let $\overline{\mathbf{B}} \cap I$ denote the subgraph of $\overline{\mathbf{B}}$ induced by $V_I(\mathbf{B})$. Also define $V'_I(\mathbf{B}) := \bigcup_{t \in I} \left\{ (s, F_s^{(t,x^*)}) : t \leqslant s \leqslant \sup I \right\}$ to be the vertices of $\mathbf{B}$ obtained by simulating paths starting from x^* within the time window I, and let $\overline{\mathbf{B}} \sqcap I$ denote the graph it induces in $\overline{\mathbf{B}}$. Two ways of approximating $\mathbf{B}$ are then as follows:

(i) Restrict to $[-n, 0]$ and pick a uniform root in $V_{[-n,0]}(\mathbf{B})$.

(ii) Simulate paths starting from x^* in the window $[-n, n]$, which gives the vertices of $\overline{\mathbf{B}} \sqcap [-n, n] \subseteq \overline{\mathbf{B}}$, then pick a uniform root in $V'_{[0,n]}(\mathbf{B})$.

After choosing a large viewing window I, a vertex picked at random will not likely be near the edge of this window, so the effects of throwing away all but this finite window can be controlled. However, the first method involves perfect knowledge of some finite window of $\mathbf{B}$. Practically speaking, when S is infinite, one does not have a way to be sure that one has computed all of $\mathbf{B}$ in a finite window, as the only tool available is to simulate sample paths starting from different locations. This is the motivation for the second method of picking a root. For, even if the edge effects caused by only viewing simulations of paths in $\mathbf{B}$ from $-n$ to n cannot be controlled, the edge effects from 0 to n can be controlled using the information from simulating from $-n$

to n. It will be shown shortly that both of these methods enjoy convergence in the local weak sense to the measure $\mathbf{P}^{\square}$.

Lemma 4.2.15. *For any strictly increasing sequence of finite intervals $(I_n)_{n\in\mathbb{N}}$ in $\mathbb{Z}$, and any function $g \in L^1(\mathbf{P}^{\square})$, one has*

$$\frac{1}{\#I_n \mathbf{E}[\#\mathbf{B}_0]} \sum_{v\in V_{I_n}(\mathbf{B})} g[\overline{\mathbf{B}}, v] \to \mathbf{E}^{\square}[g] \tag{4.16}$$

and

$$\frac{1}{\#V_{I_n}(\mathbf{B})} \sum_{v\in V_{I_n}(\mathbf{B})} g[\overline{\mathbf{B}}, v] \to \mathbf{E}^{\square}[g], \tag{4.17}$$

where both convergences happen $\mathbf{P}$-a.s. as $n \to \infty$. In particular

$$\frac{\#V_{I_n}(\mathbf{B})}{\mathbf{E}[\#V_{I_n}(\mathbf{B})]} = \frac{\#V_{I_n}(\mathbf{B})}{\#I_n \mathbf{E}[\#\mathbf{B}_0]} \to 1. \tag{4.18}$$

Proof. Assume without loss that $\Omega = \Xi^{\mathbb{Z}}$ is the canonical space and $(\theta_t)_{t\in\mathbb{Z}}$ is the family of shift operators defined by $\theta_t((\xi_s)_{s\in\mathbb{Z}}) = (\xi_{t+s})_{s\in\mathbb{Z}}$. Both statements follow from rewriting

$$\sum_{v\in V_{I_n}(\mathbf{B})} g[\overline{\mathbf{B}}, v] = \sum_{t\in I_n} \left(\sum_{x\in\mathbf{B}_t} g[\overline{\mathbf{B}}, (t,x)] \right) = \sum_{t\in I_n} g_0 \circ \theta_t,$$

where $g_0 := \sum_{x\in\mathbf{B}_0} g[\overline{\mathbf{B}}, (0,x)]$. The pointwise ergodic theorem for amenable groups (cf. [Lin01]) then proves the claim. □

Proposition 4.2.16. *Fix any strictly increasing sequence of finite intervals $(I_n)_{n\in\mathbb{N}}$ in $\mathbb{Z}$, and for each $n \in \mathbb{N}$, let $\boldsymbol{o}_n$ be, conditionally on $V_{I_n}(\mathbf{B})$, uniformly distributed on $V_{I_n}(\mathbf{B})$ and independent of $\overline{\mathbf{B}} \cap I_n$ (including its marks). Then for all bounded*

measurable $g : \mathcal{G}_* \to \mathbb{R}_{\geqslant 0}$ depending only on vertices at some bounded distance to the root, one has

$$\frac{1}{\#V_{I_n}(\mathbf{B})} \sum_{v \in V_{I_n}(\mathbf{B})} g[\overline{\mathbf{B}} \cap I_n, v] \to \mathbf{E}^{\square}[g], \qquad \mathbf{P}\text{-}a.s. \tag{4.19}$$

as $n \to \infty$. In particular,

$$\mathbf{P}([\overline{\mathbf{B}} \cap I_n, \boldsymbol{o}_n] \in \cdot) \to \mathbf{P}^{\square}, \qquad n \to \infty \tag{4.20}$$

in the sense of local weak convergence.

Proof. Fix $N \in \mathbb{N}$ and let $g : \mathcal{G}_* \to \mathbb{R}_{\geqslant 0}$ measurable, bounded, and such that g depends only on vertices at graph distance at most N from the root. One has

$$\begin{aligned}
&\mathbf{E}[g[\overline{\mathbf{B}} \cap I_n, \boldsymbol{o}_n]] \\
&= \mathbf{E}[\mathbf{E}[g[\overline{\mathbf{B}} \cap I_n, \boldsymbol{o}_n] \mid V_{I_n}(\mathbf{B})]] \\
&= \mathbf{E}\left[\frac{1}{\#V_{I_n}(\mathbf{B})} \sum_{v \in V_{I_n}(\mathbf{B})} g[\overline{\mathbf{B}} \cap I_n, v]\right] \\
&= \mathbf{E}\left[\left(\frac{\mathbf{E}[\#V_{I_n}(\mathbf{B})]}{\#V_{I_n}(\mathbf{B})}\right)\left(\frac{1}{\mathbf{E}[\#V_{I_n}(\mathbf{B})]} \sum_{v \in V_{I_n}(\mathbf{B})} g[\overline{\mathbf{B}} \cap I_n, v]\right)\right].
\end{aligned}$$

Call the two parenthesized expressions in the previous expectation a_n and b_n respectively, then it will be shown that $a_n b_n \to \mathbf{E}^{\square}[g]$ a.s., from which it also follows that $\mathbf{E}[a_n b_n] \to \mathbf{E}^{\square}[g]$ by dominated convergence. This will prove the claims. By stationarity and linearity of expectation, for each $n \in \mathbb{N}$,

$$\mathbf{E}^{\square}[g] = \frac{1}{\mathbf{E}[\#\mathbf{B}_0]} \mathbf{E}\left[\sum_{v \in \mathbf{B}_0} g[\overline{\mathbf{B}}, v]\right] = \mathbf{E}\left[\frac{1}{\mathbf{E}[\#V_{I_n}(\mathbf{B})]} \sum_{v \in V_{I_n}(\mathbf{B})} g[\overline{\mathbf{B}}, v]\right].$$

Call the inside of the last expectation c_n. Letting $[\Gamma, o]_N$ denote the neighborhood of size N around o in a network Γ, for all $n > N$

$$\begin{aligned}
&|b_n - c_n| \\
&\leqslant \frac{1}{\#I_n \mathbf{E}[\#\mathbf{B}_0]} \sum_{v \in V_{I_n}(\mathbf{B})} \left| g[\overline{\mathbf{B}} \cap I_n, v] - g[\overline{\mathbf{B}}, v] \right| \\
&\leqslant \frac{2\|g\|_\infty}{\#I_n \mathbf{E}[\#\mathbf{B}_0]} \# \left\{ v \in V_{I_n}(\mathbf{B}) : [\overline{\mathbf{B}} \cap I_n, v]_N \neq [\overline{\mathbf{B}}, v]_N \right\} \\
&\leqslant \frac{2\|g\|_\infty}{\#I_n \mathbf{E}[\#\mathbf{B}_0]} \left(\sum_{k=\min I_n}^{\min I_n + N} \#\mathbf{B}_k + \sum_{k=\max I_n - N}^{\max I_n} \#\mathbf{B}_k \right) \\
&\leqslant \frac{2\|g\|_\infty}{\#I_n \mathbf{E}[\#\mathbf{B}_0]} \left(\sum_{k \in I_n} \#\mathbf{B}_k - \sum_{k=\min I_n + N}^{\max I_n - N} \#\mathbf{B}_k \right) \\
&\to 2\|g\|_\infty - 2\|g\|_\infty \\
&= 0
\end{aligned}$$

as $n \to \infty$, **P**-a.s., by Lemma 4.2.15. But also $c_n \to \mathbf{E}^{\square}[g]$ and $a_n \to 1$, **P**-a.s., also by Lemma 4.2.15. Hence $a_n b_n \to \mathbf{E}^{\square}[g]$, **P**-a.s., as claimed. □

Proposition 4.2.17. *Fix any increasing sequence of finite intervals $(I_n)_{n\in\mathbb{N}} = ([-a_n, b_n])_{n\in\mathbb{N}}$ in $\mathbb{Z}$ containing 0 with $a_n \to \infty$ and b_n strictly increasing. For each $n \in \mathbb{N}$, let $\mathbf{o}'_n$ be, conditionally on $V'_{I_n}(\mathbf{B})$, uniformly distributed on $V'_{[0,b_n]}(\mathbf{B})$ and independent of $\overline{\mathbf{B}} \sqcap I_n$ (including its marks). Then for all bounded measurable $g : \mathcal{G}_* \to \mathbb{R}_{\geqslant 0}$ depending only on vertices at some bounded distance to the root, one has*

$$\frac{1}{\#(V'_{I_n}(\mathbf{B}) \cap [0, b_n])} \sum_{v \in V'_{I_n}(\mathbf{B}) \cap [0, b_n]} g[\overline{\mathbf{B}} \sqcap I_n, v] \to \mathbf{E}^{\square}[g], \qquad \mathbf{P}\text{-}a.s. \tag{4.21}$$

In particular,

$$\mathbf{P}([\overline{\mathbf{B}} \sqcap I_n, \boldsymbol{o}'_n] \in \cdot) \to \mathbf{P}^{\square}, \qquad n \to \infty \tag{4.22}$$

in the sense of local weak convergence.

Proof. Fix $N \in \mathbb{N}$ and let $g : \mathcal{G}_* \to \mathbb{R}_{\geqslant 0}$ measurable, bounded, and such that g depends only on vertices at graph distance at most N from the root. The finiteness of $\mathbf{B}_0$ implies that one has that $[\overline{\mathbf{B}} \sqcap I_n, v]_N = [\overline{\mathbf{B}} \cap I_n, v]_N = [\overline{\mathbf{B}}, v]_N$ eventually as $n \to \infty$ for all $v \in V_0(\mathbf{B})$, and hence for all $v \in V_{I_n}(\mathbf{B}) \cap [0, b_n - N]$ eventually as $n \to \infty$ as well. For the same reason $V'_{I_n}(\mathbf{B}) \cap [0, b_n] = V_{[0,b_n]}(\mathbf{B})$ eventually as $n \to \infty$ as well. It follows that eventually

$$\begin{aligned}
&\frac{1}{\#(V'_{I_n}(\mathbf{B}) \cap [0, b_n])} \sum_{v \in V'_{I_n}(\mathbf{B}) \cap [0,b_n]} g[\overline{\mathbf{B}} \sqcap I_n, v] \\
&= \frac{1}{\#V_{[0,b_n]}(\mathbf{B})} \sum_{v \in V_{[0,b_n]}(\mathbf{B})} g[\overline{\mathbf{B}} \cap I_n, v] \\
&+ \frac{1}{\#V_{[0,b_n]}(\mathbf{B})} \sum_{v \in V_{[b_n - N + 1, b_n]}(\mathbf{B})} (g[\overline{\mathbf{B}} \sqcap I_n, v] - g[\overline{\mathbf{B}} \cap I_n, v]).
\end{aligned}$$

Of the last two terms, $\frac{1}{\#V_{[0,b_n]}(\mathbf{B})} \sum_{v \in V_{[0,b_n]}(\mathbf{B})} g[\overline{\mathbf{B}} \cap I_n, v] \to \mathbf{E}^{\square}[g]$ by Proposi-

tion 4.2.16, so it suffices to show that the last term goes to 0. Indeed,

$$\begin{aligned}
&\left| \frac{1}{\#V_{[0,b_n]}(\mathbf{B})} \sum_{v \in V_{[b_n-N+1,b_n]}(\mathbf{B})} (g[\overline{\mathbf{B}} \sqcap I_n, v] - g[\overline{\mathbf{B}} \cap I_n, v]) \right| \\
&\leqslant \frac{2\|g\|_\infty \#V_{[b_n-N+1,b_n]}(\mathbf{B})}{\#V_{[0,b_n]}(\mathbf{B})} \\
&= \frac{2\|g\|_\infty (\#V_{[0,b_n]}(\mathbf{B}) - \#V_{[0,b_n-N]}(\mathbf{B}))}{\#V_{[0,b_n]}(\mathbf{B})} \\
&\to 2\|g\|_\infty (1-1) \\
&= 0
\end{aligned}$$

as desired. □

Renewal Structure of the Bridge Graph

In this section, the driving sequence ξ is assumed to be i.i.d., i.e. $\mathbf{G}$ is Markovian. One may ask whether the bridge graph $\mathbf{B}$ admits any kind of renewal structure. Is it possible that $\mathbf{B}_t$ contains only one state? This is not necessarily possible. Indeed, if $p_{x,x} = 0$, then $\mathbf{B}_t$ contains at least two states for every $t \in \mathbb{Z}$. It is true, though, that $\mathbf{B}_t$ is infinitely often equal to any set that it has positive probability of being equal to. Let $S_{\mathbf{B}}$ denote the **possible configurations** of $\mathbf{B}_0$, i.e. $S_{\mathbf{B}} := \{E \subseteq S : \mathbf{P}(\mathbf{B}_0 = E) > 0\}$. By Proposition 4.1.13, $S_{\mathbf{B}}$ consists only of finite subsets of S and is therefore countable.

Lemma 4.2.18. *For any subset $E \in S_{\mathbf{B}}$, the set of t for which $\mathbf{B}_t = E$ forms a simple stationary point process Ψ_E on $\mathbb{Z}$ with $\mathbf{P}(\Psi_E(\mathbb{Z}) = \infty) = 1$ and intensity $\lambda_E = \mathbf{P}(\mathbf{B}_0 = E)$. In particular, $(\mathbf{B}_t)_{t\in\mathbb{Z}}$ is bi-recurrent for each $E \in S_{\mathbf{B}}$.*

Proof. For $E \in S_{\mathbf{B}}$, the event that there is a t such that $\mathbf{B}_t = E$ is shift-invariant and has positive probability. Therefore it happens almost surely. The set of such t is shift-covariant and therefore determines a simple stationary point process Ψ_E. The previous line implies that Ψ_E contains at least one point, and therefore infinitely many a.s. One calculates $\lambda_E = \mathbf{E}[\Psi_E(\{0\})] = \mathbf{E}[1_{\{\mathbf{B}_0 = E\}}]$, completing the proof. □

Moreover, ruling out obvious hurdles to $\mathbf{B}_t$ being a singleton is sufficient.

Lemma 4.2.19. *Suppose* $\mathbf{G}$ *is an EFT and has fully independent transitions. Assume that* $p_{x^*,x^*} > 0$. *Then* $\{x^*\} \in S_{\mathbf{B}}$.

Proof. By Proposition 4.1.13, $\#\mathbf{B}_t$ is a.s. finite for each $t \in \mathbb{Z}$, and thus it is possible to choose $x_1, \ldots, x_n \in S$ such that $\mathbf{P}(\mathbf{B}_0 = \{x_1, \ldots, x_n\}) > 0$. Since $\mathbf{G}$ is an EFT, choose a tree $T \subseteq \mathbb{Z} \times S$ with leaves $(0, x_1), \ldots, (0, x_n)$ and root (t, x^*) for some $t > 0$ such that $\mathbf{P}(T \subseteq \mathbf{G}) > 0$. With $[t] := \{0, \ldots, t\}$, let $I := \{s \in [t] : x^* \notin T_s\}$. Then

$$\begin{aligned}
\mathbf{P}(\mathbf{B}_t = \{x^*\}) &\geqslant \mathbf{P}(\mathbf{B}_0 = \{x_1, \ldots, x_n\}, T \subseteq \mathbf{G}, F_{s+1}^{(s,x^*)} = x^*, \forall s \in I) \\
&\geqslant \mathbf{P}(\mathbf{B}_0 = \{x_1, \ldots, x_n\})\mathbf{P}(T \subseteq \mathbf{G})\mathbf{P}(F_{s+1}^{(s,x^*)} = x^*, \forall s \in I) \\
&= \mathbf{P}(\mathbf{B}_0 = \{x_1, \ldots, x_n\})\mathbf{P}(T \subseteq \mathbf{G})(p_{x,x})^{\#I} \\
&> 0.
\end{aligned}$$

To justify the use of independence in the previous calculation, note that $\mathbf{B}_0$ is $(\xi_s)_{s<0}$-measurable, whereas the events $\{T \subseteq \mathbf{G}\}$ and $\left\{F_{s+1}^{(s,x^*)} = x^*, \forall s \in I\right\}$

are $(\xi_s)_{s\geqslant 0}$-measurable, so the first is independent of the second two. Then the second is independent of the third because, by construction, they involve disjoint sets of edges in $\mathbf{G}$. □

Now it is possible to see the renewal structure in $\mathbf{B}$. Namely, $(\mathbf{B}_t)_{t\in\mathbb{Z}}$ is itself an irreducible, aperiodic, and positive recurrent Markov chain under certain conditions.

Proposition 4.2.20. *One has that $(\mathbf{B}_t)_{t\in\mathbb{Z}}$ is a Markov chain on $S_{\mathbf{B}}$. Additionally, $(\mathbf{B}_t)_{t\in\mathbb{Z}}$ is stationary and bi-recurrent for every $E \in S_{\mathbf{B}}$. Its transition matrix $P_{\mathbf{B}}$ is irreducible and positive recurrent. If $\mathbf{G}$ is an EFT with fully independent transitions and $p_{x^*,x^*} > 0$, then $\{x^*\} \in S_{\mathbf{B}}$ and $P_{\mathbf{B}}(\{x^*\}, \{x^*\}) > 0$ so $P_{\mathbf{B}}$ is aperiodic as well.*

Proof. By Proposition 4.1.13, $\#\mathbf{B}_t$ is a.s. finite for each $t \in \mathbb{Z}$. Moreover, $\mathbf{B}_{t+1} = \{x^*\} \cup \left\{F_{t+1}^{(t,y)} : y \in \mathbf{B}_t\right\}$, so indeed $(\mathbf{B}_t)_{t\in\mathbb{Z}}$ is a Markov chain on the finite subsets of S since, for each $t \in \mathbb{Z}$, $\mathbf{B}_{t+1}$ is a function of $\mathbf{B}_t$ and ξ_t. Here the running assumption that ξ is i.i.d. is used. By Lemma 4.2.18, $(\mathbf{B}_t)_{t\in\mathbb{Z}}$ is bi-recurrent for every state $E \subseteq S$ such that $\mathbf{P}(\mathbf{B}_0 = E) > 0$. In particular, the chain must be irreducible on $S_{\mathbf{B}}$, else a return to some state E_1 could not occur after a return to another state E_2 for some E_1, E_2 that do not communicate. Since $(\mathbf{B}_t)_{t\in\mathbb{Z}}$ is shift-covariant it is stationary. The existence of a positive stationary distribution (the law of $\mathbf{B}_0$) for the irreducible $P_{\mathbf{B}}$ implies $P_{\mathbf{B}}$ is positive recurrent. If $\mathbf{G}$ is an EFT with fully independent transitions, then Lemma 4.2.19 shows that $\{x^*\} \in S_{\mathbf{B}}$. Then $p_{x^*,x^*} > 0$ implies $P_{\mathbf{B}}(\{x^*\}, \{x^*\}) > 0$ as well, so $P_{\mathbf{B}}$ is also aperiodic in that case. □

It is possible that the $S_{\mathbf{B}}$ is strictly smaller than the set of all finite subsets of S containing x^*.

Example 4.2.21. Consider $S := \{0, 1, 2\}$ and $x^* := 0$ with $p_{0,0} = p_{0,1} = p_{0,2} = \frac{1}{3}$ and $p_{1,0} = p_{2,0} = 1$. That is, from 0 make a uniform choice of where to jump, and from 1 and 2 deterministically return to 0. Fix $t \in \mathbb{Z}$. In this case, if $1 \in \mathbf{B}_t$, it must be that $F_t^{(t-1,0)} = 1$. Similarly, if $2 \in \mathbf{B}_t$, it must be that $F_t^{(t-1,0)} = 2$. Thus it cannot be that both $1, 2 \in \mathbf{B}_t$, and hence $\{0, 1, 2\} \notin S_{\mathbf{B}}$.

However, if *every* state has a chance to be lazy, then $S_{\mathbf{B}}$ does turn out to be the set of all finite subsets of S containing x.

Proposition 4.2.22. *Suppose* $\mathbf{G}$ *has fully independent transitions,* P *is irreducible, and* $p_{y,y} > 0$ *for all* $y \in S$. *Then* $(\mathbf{B}_t)_{t\in\mathbb{Z}}$ *is an irreducible, aperiodic, positive recurrent, and stationary Markov chain on the set of all finite subsets of* S *containing* x^*.

Proof. The assumptions imply that, in fact, P is irreducible, aperiodic, and positive recurrent (since x^* is always assumed positive recurrent), so Proposition 4.2.20 implies that the only item left to show is that $S_{\mathbf{B}}$ contains all finite subsets of S containing x^*. Let a finite set E containing x^* be given. Call $(y_1, \ldots, y_n)$ with each $y_i \in S$ a **possible path** if $\prod_{i=1}^{n-1} p_{y_i,y_{i+1}} > 0$. For the rest of the proof, all paths considered are possible paths. One would like to simply draw a path from x^* to each $y \in S$ where after a path reaches its destination it becomes constant while it waits for the other paths to finish. This

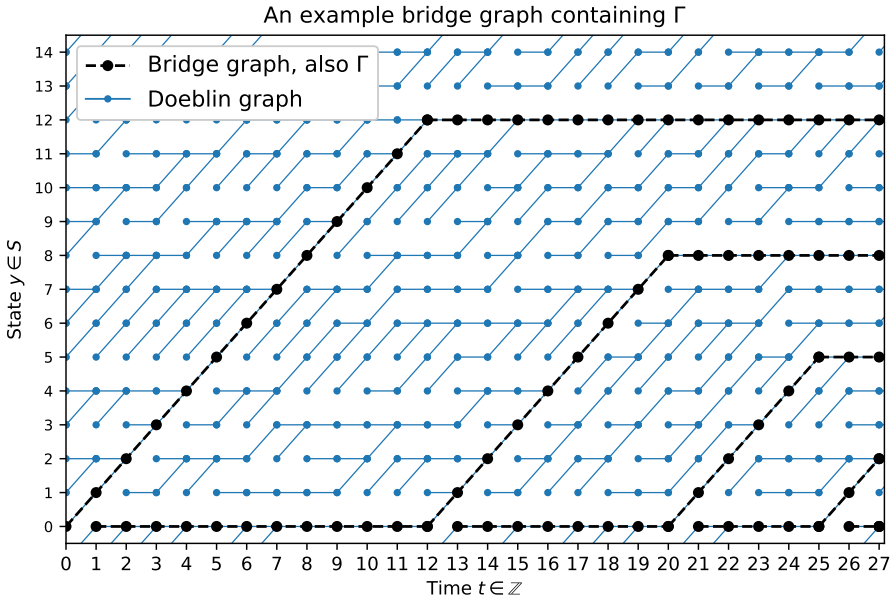


Figure 4.4: The graph Γ from the proof of Proposition 4.2.22 when $S = \mathbb{Z}/15\mathbb{Z}$, $x^* = 0$, and $E = \{0, 2, 5, 8, 12\}$, where the Markov dynamics are the lazy version of the deterministic cycle $x \mapsto x + 1$ on S. The graph Γ is constructed so that, as shown in the figure, if $\mathbf{B}_0 = \{0\}$ and $\Gamma \subseteq \mathbf{G}$, then $\mathbf{B}_{27} = \{0, 2, 5, 8, 12\}$.

approach is slightly flawed because it may be that, for instance, every path from x^* to z passes through y. In this case, one must draw the path from x^* to y before the path from x^* to z, otherwise the resulting graph would have a vertex with multiple outgoing edges, which is an impossibility in $\mathbf{G}$. However, the approach will work as long as it is possible to draw the paths in an order such that no interference occurs.

Define a partial order $\prec_0$ on E by saying $y \prec_0 z$ if all paths from x^* to z pass through y with the convention that the trivial path (x^*) does not

pass through x^* (to prohibit $x^* \prec_0 x^*$). Since E is finite, there is a $\prec_0$-maximal element $x_0 \in E$. That is, for all $y \in E$ there is a path from x^* to y that does not hit x_0. Choose a path L_0 from x^* to x_0. With $\prec_n$, $x_0, \ldots, x_n$, and $L_0, \ldots, L_n$ defined, as long as $E \setminus \{x_0, \ldots, x_n\} \neq \emptyset$, recursively define $\prec_{n+1}$, x_{n+1}, and L_{n+1} as follows. By construction, for all $y \in E \setminus \{x_0, \ldots, x_n\}$, there is a path from x^* to y that avoids $x_0, \ldots, x_n$. Define $\prec_{n+1}$ on $E \setminus \{x_0, \ldots, x_n\}$ by saying $y \prec z$ if all paths from x^* to z avoiding $x_0, \ldots, x_n$ pass through y. Then it is possible to choose a $\prec_{n+1}$-maximal element x_{n+1}, i.e. for all $y \in E \setminus \{x_0, \ldots, x_{n+1}\}$, there is a path from x^* to y that does not pass through any of $x_0, \ldots, x_{n+1}$. Also choose L_{n+1} a path from x^* to x_{n+1} avoiding $x_0, \ldots, x_n$. Necessarily the recursion terminates when $n = \#E - 1$. It is now possible to construct a graph $\Gamma \subseteq \mathbb{Z} \times S$ with $\mathbf{P}(\mathbf{B}_0 = \{x^*\}, \Gamma \subseteq \mathbf{G}) > 0$ and when $\Gamma \subseteq \mathbf{G}$ and $\mathbf{B}_0 = \{x^*\}$, one has $\mathbf{B}_t = E$ for some t. Let t_i be the sum of the lengths of the paths $L_0, \ldots, L_{i-1}$ for each $0 \leqslant i \leqslant \#E$, with $t_0 := 0$. Let Γ be the graph that for each i has:

(i) a path from (t_i, x) to (t_{i+1}, x_i) with state path L_i from time t_i to t_{i+1},

(ii) a path started at (t_{i+1}, x_i) that stays constant at x_i until time $t_{\#E}$, and

(iii) a (possibly trivial) path started at $(t_i + 1, x)$ that stays constant at x^* until time t_{i+1}.

Note that, by construction, Γ is a finite graph that is the union of edges that occur with positive probability. Moreover, the connected components of Γ

are formed from points items (i) and (ii) for some i and item (iii) from $i-1$. Whenever $\Gamma \subseteq \mathbf{G}$ and $\mathbf{B}_0 = \{x^*\}$, one has $\mathbf{B}_{t_{\#E}} = E$. This occurs with positive probability since $p_{y,y} > 0$ for all $y \in S$. □

Proposition 4.2.23. *Suppose* $\mathbf{G}$ *has fully independent transitions. Extend the definition of* $P_{\mathbf{B}}$ *to*

$$P_{\mathbf{B}}(E, E') := \mathbf{P}\left(\{x^*\} \cup \left\{F_1^{(0,y)} : y \in E\right\} = E'\right), \tag{4.23}$$

for all finite $E, E' \subseteq S$ *containing* x^**. Then* $P_{\mathbf{B}}$ *satisfies the following recurrence: for* $E = \{x^*, x_1, \ldots, x_n\}$ *and* $E' = \left\{x^*, y_1, \ldots, y_m\right\}$*,*

$$\begin{aligned} P_{\mathbf{B}}(E, E') = &\left(p_{x_n,x^*} + \sum_{i=1}^{m} p_{x_n,y_i}\right) P_{\mathbf{B}}(E \setminus \{x_n\}, E') \\ &+ \sum_{i=1}^{m} p_{x_n,y_i} P_{\mathbf{B}}(E \setminus \{x_n\}, E' \setminus \left\{y_i\right\}), \end{aligned} \tag{4.24}$$

with recursive depth at most n *and base cases*

$$\begin{cases} P_{\mathbf{B}}(E, E') = 0, & \#E' > \#E + 1 \\ P_{\mathbf{B}}(\{x^*\}, \left\{x^*, y\right\}) = p_{x^*,y}, & y \in S \\ P_{\mathbf{B}}(E, \{x^*\}) = \prod_{y \in E} p_{y,x^*}. \end{cases} \tag{4.25}$$

Proof. First one justifies the extension of the definition of $P_{\mathbf{B}}$ by noting that for $E, E' \in S_{\mathbf{B}}$ one has

$$\begin{aligned} P_{\mathbf{B}}(E, E') &= \mathbf{P}(\mathbf{B}_1 = E' \mid \mathbf{B}_0 = E) \\ &= \mathbf{P}(\{x^*\} \cup \left\{F_1^{(0,y)} : y \in \mathbf{B}_0\right\} = E' \mid \mathbf{B}_0 = E) \\ &= \mathbf{P}(\{x^*\} \cup \left\{F_1^{(0,y)} : y \in E\right\} = E' \mid \mathbf{B}_0 = E) \\ &= \mathbf{P}(\{x^*\} \cup \left\{F_1^{(0,y)} : y \in E\right\} = E'), \end{aligned}$$

where the last equality follows from the fact that $\mathbf{B}_0$ is measurable with respect to $(\xi_t)_{t<0}$, whereas $F_1^{(0,y)}$ is ξ_0-measurable for each $y \in S$. The base cases for $P_{\mathbf{B}}$ are immediate from the definition of $P_{\mathbf{B}}$ and the independence structure. To see the recurrence, suppose $E = \{x^*, x_1, \ldots, x_n\}$ and $E' = \left\{x^*, y_1, \ldots, y_m\right\}$ as above. Split $P_{\mathbf{B}}(E, E')$ depending on the value of $F_1^{(0,x_n)} = x^*$ or $F_1^{(0,x_n)} = y_i$, and on whether $\{x^*\} \cup \left\{F_1^{(0,y)} : y \in E \setminus \{x_n\}\right\} = E'$ still or $\{x^*\} \cup \left\{F_1^{(0,y)} : y \in E \setminus \{x_n\}\right\} = E' \setminus \left\{y_i\right\}$,

$$\begin{aligned}
P_{\mathbf{B}}(E, E') = \mathbf{P}&\left(F_1^{(0,x_n)} = x^*, \{x^*\} \cup \left\{F_1^{(0,y)} : y \in E \setminus \{x_n\}\right\} = E'\right) \\
&+ \sum_{i=1}^{m} \mathbf{P}\left(F_1^{(0,x_n)} = y_i, \{x^*\} \cup \left\{F_1^{(0,y)} : y \in E \setminus \{x_n\}\right\} = E'\right) \\
&+ \sum_{i=1}^{m} \mathbf{P}\left(F_1^{(0,x_n)} = y_i, \{x^*\} \cup \left\{F_1^{(0,y)} : y \in E \setminus \{x_n\}\right\} = E' \setminus \left\{y_i\right\}\right)
\end{aligned}$$

which, since $\mathbf{G}$ has fully independent transitions, equals

$$\begin{aligned}
&p_{x_n,x^*}\mathbf{P}\left(\{x^*\} \cup \left\{F_1^{(0,y)} : y \in E \setminus \{x_n\}\right\} = E'\right) \\
&+ \sum_{i=1}^{m} p_{x_n,y_i}\mathbf{P}\left(\{x^*\} \cup \left\{F_1^{(0,y)} : y \in E \setminus \{x_n\}\right\} = E'\right) \\
&+ \sum_{i=1}^{m} p_{x_n,y_i}\mathbf{P}\left(\{x^*\} \cup \left\{F_1^{(0,y)} : y \in E \setminus \{x_n\}\right\} = E' \setminus \left\{y_i\right\}\right)
\end{aligned}$$

which simplifies to

$$\left(p_{x_n,x^*} + \sum_{i=1}^{m} p_{x_n,y_i}\right) P_{\mathbf{B}}(E \setminus \{x_n\}, E') + \sum_{i=1}^{m} p_{x_n,y_i} P_{\mathbf{B}}(E \setminus \{x_n\}, E' \setminus \left\{y_i\right\}),$$

showing the recurrence holds.

Finally, the recursive depth needed to fully compute $P_{\mathbf{B}}(E, E')$ is at most n because each application of the recurrence removes an element from E. □

Example 4.2.24. By implementing the recurrence of Proposition 4.2.23 in, e.g. Python, one may compute $P_{\mathbf{B}}$ explicitly. Then, given values for the $p_{x,y}$, one may compute the stationary distribution $\pi_{\mathbf{B}}$ of $P_{\mathbf{B}}$. For example, with $S := \{0, 1, 2\}$ and $x^* := 0$, and $p_{x,y} = \frac{1}{3}$ for all $x, y \in S$, one has

$$\begin{aligned}\pi_{\mathbf{B}} &= \begin{bmatrix}\pi_{\mathbf{B}}(\{0\}) & \pi_{\mathbf{B}}(\{0,1\}) & \pi_{\mathbf{B}}(\{0,2\}) & \pi_{\mathbf{B}}(\{0,1,2\})\end{bmatrix} \\ &= \begin{bmatrix}\frac{17}{143} & \frac{45}{143} & \frac{45}{143} & \frac{36}{143}\end{bmatrix}.\end{aligned}$$

It is an open question whether, in the fully independent transitions case, there is a general closed form expression for $P_{\mathbf{B}}$ in terms of P or for the stationary distribution $\pi_{\mathbf{B}}$ of $P_{\mathbf{B}}$ in terms of P and π.

4.3. Bibliographical Comments

While this work may be the first time the Doeblin graph **G** has been explicitly defined and studied in its own right, it is without doubt that most, if not all, who have worked on CFTP-related research have had this picture in mind. Rather, the novelty here lies in the consideration of the bridge graph **B**. While, to the best of the author's knowledge, the bridge graph **B** has not previously been defined or studied, it is not without ties to other objects that have been previously studied.

The first occurrence of some form of the bridge graph appears in [BF92], where Borovkov and Foss consider a family of stochastically recursive sequences started at times $0, -1, -2, \ldots$, all with the same initial condition, and they proved the existence (under suitable conditions) of a stationary version of the SRS. They defined three notions of coupling convergence and studied when coupling convergence to the stationary SRS occurs. Their notion of strong coupling convergence to the stationary SRS is akin to the condition that $\mathbf{B}$ is an EFT. That is, it is the condition that all paths in $\mathbf{B}$ eventually merge. It is conceivable that, in the EFT case, one could derive the existence of the bi-infinite path in $\mathbf{B}$ from the work in [BF92], though it is not clear whether Borovkov and Foss had this in mind, and they did not make any mention of the key bi-recurrence property used in the current chapter to distinguish this bi-infinite path from the potential others in $\mathbf{G}$.

Another occurrence of a similar object to the bridge graph may be found in [BS18] in the very special case of integer-valued renewal processes. The dynamics there are slightly different, where instead of specifying a whole process started from each time, one marks each time with the time of death of an individual who is born at that time. This is akin to marking each $t \in \mathbb{Z}$ by the return time $\tau^{(t,x^*)}(x^*)$ of $F^{(t,x^*)}$ to x^*, though in [BS18] these times of death are assumed to be i.i.d., whereas in the present work they have intricate dependence due to the Doeblin-type coupling. The population process defined in [BS18] is then similar in nature to the sequence of

cardinalities of $(\mathbf{B}_t)_{t\in\mathbb{Z}}$ as considered in Section 4.2.3. It is proved in [BS18] that, under natural conditions, the population process is a stationary regenerative process with independent cycles. In the present work, the process $(\mathbf{B}_t)_{t\in\mathbb{Z}}$ was shown in Proposition 4.2.20 to be an irreducible, aperiodic, and positive recurrent Markov chain under suitable conditions, which therefore also admits an i.i.d. cycle decomposition. The analysis of this special case and, in particular, the identification of the I/F structure of the components has been kept in mind throughout the development of the theory of Doeblin EFFs.

Chapter 5

Point-shifts of Point Processes on Topological Groups

Dynamics on stationary point processes that map each point of the process to another in a translation invariant way, or point-shifts as they are called in the literature, have been studied as far back as Mecke's seminal work [Mec67]. A recent advancement in the study of such dynamics was the classification of a point-shift based on the cardinalities of the foils and connected components of the random graph generated by the point-shift [BHM18]. This graph is drawn on the vertices of the point process with an edge from each point to its image under the point-shift. Two points are considered to be in the same foil if they eventually merge under the same number of repeated applications of the point-shift, and they are in the same component if their forward orbits under the point-shift intersect, i.e., if they are in the same connected component of the graph. The classification states that a.s. each connected component belongs to one of the following three classes:

(i) **Class F/F**: The component and its foils are all finite, and there is a unique cycle in the component.

(ii) **Class I/F**: The component is infinite, but all foils are finite, there are no cycles, and there is a unique bi-infinite path in the component.

(iii) **Class I/I**: The component is infinite, all its foils are infinite, and there are no cycles or bi-infinite paths in the component.

The primary result of this chapter is the extension of the classification theorem verbatim to unimodular groups. In other words, the answers (infinitely many/finitely many) to the questions: "How many points are in a given graph component?" and "Given a point, how many other points will merge with it under repeated application of the point-shift?" fundamentally determine many properties of the components of the graph generated by a point-shift.

The type (F/F, I/F, or I/I) of a component determines whether it is acyclic, the number of bi-infinite paths it contains, and whether any points remain after an infinite number of applications of the point-shift. This cardinality classification was generalized in [BHMK18] to so-called (covariant) vertex-shifts on unimodular random networks (see again Theorem 2.5.1 in the preliminaries). Here a network refers to a graph with extra information called marks associated to its vertices and edges, and a vertex-shift maps each point of the network to another in an isomorphism invariant way. It was shown in an early preprint (arXiv v1) of [BHMK18] that, under its Palm probability, which encodes the view of the world from a "typical" point's perspective, a stationary point process on $\mathbb{R}^d$ can be seen as an embedding of

a unimodular random network in a natural sense, where $0 \in \mathbb{R}^d$ is considered to be the root of the point process. The most current version of [BHMK18] no longer includes this section, but an independent proof is provided in this chapter. Thus the cardinality classification for point-shifts on $\mathbb{R}^d$ is indeed a special case of the cardinality classification for vertex-shifts.

This chapter is concerned with generalizing the cardinality classification in a different direction, namely to point-shifts of G-stationary point processes on a unimodular group G. For a group G, G-stationarity means distributional invariance with respect to the action of G on a space, and unimodularity refers to the existence of a bi-invariant Haar measure, though this form of unimodularity does imply a mass-transport principle for point processes. Point process theory has been developed on any locally compact second-countable Hausdorff (LCSH) space [DVJ08], and G-stationarity has been studied thoroughly when G is a LCSH group acting on a homogeneous space S, cf. [Las10b]. The theory of G-stationarity on homogeneous spaces is very general, but in many cases a G-stationary point process on a homogeneous space S may be pushed forward in a natural way to a G-stationary point process on G itself, and this is the setting that is assumed here. Also note that the embedding techniques used to view stationary point processes on $\mathbb{R}^d$ as unimodular random networks used properties of Delaunay triangulations that do not generalize to arbitrary LCSH groups, and therefore the present work is not subsumed by similar work on unimodular networks. Indeed, it will even be shown that the cardinality classification fails to hold

for a point-shift on an explicit LCSH group that is not unimodular.

The structure of the chapter is as follows. Section 5.1 defines the framework used for dealing with G-stationary point processes and gives some basic results about Palm probabilities and point-shifts. In Section 5.2.1, the structure of the components of the random graph generated by a point-shift is studied when G is unimodular. Theorem 5.2.7 shows that the cardinality classification when $G = \mathbb{R}^d$ extends verbatim to G when G is unimodular, though proofs in [BHM18] relying on the order of $\mathbb{R}$ are replaced by direct mass-transport arguments. Section 5.2.2 shows that unimodularity of G is crucial to the cardinality classification theorem. In this section, G is chosen to be an explicit non-unimodular group, the $ax + b$ group, and a point-shift F is given for which the cardinality classification fails in many respects. The next portion of the chapter explores the boundary between the unimodular and non-unimodular cases and spends a good amount of effort identifying those point-shifts that act like the underlying space is unimodular even when it is not. Such point-shifts will be called **isomodular**.

In Section 5.3, at points it is more natural to use a point-map than a point-shift. Point-maps are equivalent to point-shifts in that they map points of a point process to other points of the point process, except a point-map specifies only where the identity element under the Palm probability measure is mapped. Some simple relationships involving point-maps are given in Section 5.3.1. It is a classical result on $\mathbb{R}^d$, cf. [HL05], that a point-shift preserves the Palm probability measure of a point process if and only if

the point-shift is almost surely bijective on the support of the process. This result is commonly referred to as Mecke's invariance theorem or, in some cases, Mecke's point-stationarity theorem. In Section 5.3.2 it is shown that Mecke's invariance theorem holds if G is unimodular. Moreover, when G is not necessarily unimodular, the class of bijective point-shifts that preserve Palm probabilities is identified as the bijective isomodular point-shifts, the point-shifts that preserve the modular function of the group. Section 5.3.3 continues with the study of isomodularity and investigates for bijective point-shifts the distributional relationship between the reciprocal of the corresponding point-map and the reverse point-map, which corresponds to running the point-shift backwards in time. Section 5.3.4 studies different ways in which functions separate points of a point process. For example, given a function $h : G \to S$ for some set S and a point process Ψ, when are the values of $h(X)$ distinct for all $X \in \Psi$?

Finally, Section 5.4 attempts make explicit the connections between unimodular networks and G-stationary point processes on unimodular groups by defining a way to view a point process as an embedding of a unimodular network. Two natural questions then arise:

(i) Given a G-stationary point process, when can the Palm version of it be seen as an embedding of a unimodular network?

(ii) Given a unimodular network, when is it possible to find a G-stationary point process such that the Palm version of the point process is an

embedding of the given unimodular network?

Some progress in the answering the first question is made in Section 5.4, where the problem is reduced to an invariant geometry problem on the underlying space, which is conjectured to always be solvable when G is unimodular. The results of previous sections also indicate that when G is not unimodular, one should not expect either question to be answered affirmatively.

5.1. Point-shift Basics and Notation

This chapter will make use of simple point processes on a general LCSH group G, which is fixed for the remainder of the chapter. Denote the Borel sets of G by $\mathcal{B}(G)$.

The following framework for dealing with G-stationary point processes was developed in [Las10b, Las10a]. A **stationary framework** $(\Omega, \mathcal{F}, \theta, \mathbf{P})$ on G is a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ equipped with a measurable and $\mathbf{P}$-invariant left G-action $\theta : G \times \Omega \to \Omega$, called a **flow**, which will be identified with the family of mappings $(\theta_x)_{x \in G}$ defined by $\theta_x \omega := \theta(x, \omega)$ for $x \in G, \omega \in \Omega$. A point process Ψ is **flow-adapted** if

$$\Psi(\theta_x \omega, B) = \Psi(\omega, x^{-1}B), \qquad x \in G, \omega \in \Omega, B \in \mathcal{B}(G). \tag{5.1}$$

Another way of expressing (5.1) is

$$\Psi \circ \theta_x = T_x \Psi, \qquad x \in G,$$

where for $\mu \in \mathbf{M}$ and $x \in G$ the translated measure $T_x\mu$ is defined by $T_x\mu(B) := \mu(x^{-1}B)$ for all $B \in \mathcal{B}(G)$. Under these assumptions, any flow-adapted Ψ is G-stationary in the usual sense that Ψ and $T_x\Psi$ have the same distribution for all $x \in G$. For the remainder of the chapter, fix a stationary framework $(\Omega, \mathcal{F}, \theta, \mathbf{P})$ on G. All point processes introduced in this chapter are assumed to be flow-adapted.

A **point-shift** on a point process Ψ is a measurable map $F : \Omega \times G \to G$ on the support of Ψ, i.e. for $\mathbf{P}$-a.e. $\omega \in \Omega$,

$$F(\omega, X) \in \Psi(\omega), \qquad X \in \Psi(\omega). \tag{5.2}$$

From now on, all point-shifts considered in this chapter are assumed to be **flow-adapted** in the sense that

$$F(\theta_y\omega, yx) = yF(\omega, x), \qquad x, y \in G, \omega \in \Omega. \tag{5.3}$$

If unspecified, $F(\omega, x) := x$ for $x \notin \Psi(\omega)$. Dependence on ω is usually dropped and $F(X)$ is written instead of $F(\omega, X)$. Say that F has a functional property, e.g. bijectivity, injectivity, surjectivity, if for $\mathbf{P}$-a.e. $\omega \in \Omega$, $F(\omega, \cdot)$ has the property on the support of $\Psi(\omega)$.

Fix, for the remainder of the chapter, a left-invariant Haar measure λ on G. Also, for the remainder of the section, suppose Ψ is a flow-adapted point processes with finite and nonzero intensity, that is $\Lambda(\cdot) := \mathbf{E}[\Psi(\cdot)]$ is locally finite and not the zero measure. The **Palm probability measure** of Ψ, denoted $\mathbf{P}^\Psi$, is defined by

$$\mathbf{P}^\Psi(A) := \frac{1}{\gamma}\mathbf{E}\int_G 1_{\{\theta_x^{-1} \in A\}} w(x)\, \Psi(dx), \qquad A \in \mathcal{F}, \tag{5.4}$$

where $\gamma = \frac{\mathbf{E}[\Psi(B)]}{\lambda(B)}$ for any $B \in \mathcal{B}(G)$ with $\lambda(B) \in (0, \infty)$, and $w : G \to \mathbb{R}_{\geqslant 0}$ is any nonnegative measurable function with $\int_G w \, d\lambda = 1$. Note that $\gamma \in (0, \infty)$ is uniquely determined and $\mathbf{P}^\Psi$ is independent of the choice of w. Expectation with respect to $\mathbf{P}^\Psi$ is denoted $\mathbf{E}^\Psi$. The Palm probability measure $\mathbf{P}^\Psi$ makes rigorous what is meant by the view of the world from a typical point's perspective.

It is possible to convert between $\mathbf{P}$-a.s. and $\mathbf{P}^\Psi$-a.s. events in the following manner. Intuitively, that which happens almost surely from the typical point's perspective happens almost surely from every point's perspective simultaneously, and vice-versa.

Proposition 5.1.1. *Let $A \in \mathcal{F}$. Then the following are equivalent:*

(a) $\mathbf{P}^\Psi(A) = 1$.

(b) $\mathbf{P}(\Psi(x \in G : \theta_x^{-1} \notin A) = 0) = 1$.

(c) $\mathbf{P}^\Psi(\Psi(x \in G : \theta_x^{-1} \notin A) = 0) = 1$.

A proof of Proposition 5.1.1 could not be found in the literature, so one is given in Appendix B.

For the rest of the chapter, $\Delta : G \to (0, \infty)$ is the modular function of G. Then the mass-transport principle for point processes takes the following form.

Theorem 5.1.2 (Mass-transport Theorem)**.** *[Las10a] For all* ***diagonally invariant*** τ*, i.e. measurable* $\tau : \Omega \times G \times G \to \mathbb{R}_{\geqslant 0}$ *invariant in the sense that*

$$\tau(\theta_z \omega, zx, zy) = \tau(\omega, x, y) =: \tau(x, y), \qquad \omega \in \Omega, \quad x, y, z \in G, \tag{5.5}$$

it holds that

$$\mathbf{E}^{\Psi} \int_G \tau(e, y)\, \Psi(dy) = \mathbf{E}^{\Psi} \int_G \tau(x, e) \Delta(x^{-1})\, \Psi(dx). \tag{5.6}$$

Interpret $\tau(\omega, x, y)$ as the amount of mass sent from x to y on the outcome ω. Under $\mathbf{E}^{\Psi}$, e is a point of Ψ. Thus the left side of (5.6) is an average of mass sent out of $e \in \Psi$ to all points of Ψ. On the other hand, the right side of (5.6) is a weighted average of mass received by $e \in \Psi$ from all points of Ψ. If $\Delta(x) = 1$ for all $x \in G$, i.e. if G is unimodular, then the mass-transport formula is the one expected from the case of translations on $\mathbb{R}^d$, which says that the average mass a typical point of Ψ receives equals the average mass a typical point of Ψ sends. The mass-transport theorem takes an even simpler form for a point-shift F, i.e., when $\tau(x, y) := 1_{\{F(x)=y\}}$ when G is unimodular.

Proposition 5.1.3. *Suppose that G is unimodular. Then for every point-shift F on Ψ, one has that* $\mathbf{E}^{\Psi}[\#\{X \in \Psi : F(X) = e\}] = 1$. □

In particular, this gives the very useful fact that injectivity and surjectivity are equivalent for point-shifts on unimodular G.

Proposition 5.1.4. *Suppose that G is unimodular and let F be a point-shift on Ψ. Then F is injective if and only if it is surjective.*

Proof. If F is injective, then $\#\{X \in \Psi : F(X) = e\} \leqslant 1$ $\mathbf{P}^\Psi$-a.s. by Proposition 5.1.1. But $\mathbf{E}^\Psi[\#\{X \in \Psi : F(X) = e\}] = 1$ so in fact $\#\{X \in \Psi : F(X) = e\} = 1$, $\mathbf{P}^\Psi$-a.s. This implies F is bijective by Proposition 5.1.1 again. The proof when F is surjective is the same with $\leqslant 1$ replaced by $\geqslant 1$. □

5.2. Point-shift Foliations

5.2.1. The Cardinality Classification of Components

In this section, the cardinality classification components of point-shifts in [BHM18] is extended to the general stationary framework for unimodular G. The classification theorem is Theorem 5.2.7, and the fundamental result used in its proof, which says it is impossible to pick out finite subsets of infinite sets in a flow-adapted manner, is Proposition 5.2.3.

Throughout this section, G is assumed to be unimodular. Fix for the rest of the section a flow-adapted simple point process Ψ on G with intensity $\gamma \in (0, \infty)$, and a point-shift F on Ψ. The wording of proofs is substantially cut down by thinking of $F(X)$ as the **father** of X. For example, the **children** of X are the $Y \in \Psi$ such that $F(Y) = X$. Next appear the necessary ingredients needed for the classification theorem.

The iterates F^n are defined by repeatedly applying the point-shift F. That is, $F^0(X) := X$ and $F^{n+1}(X) := F(F^n(X))$ for all $X \in \Psi$. Elements $Y \in \Psi$ that are in the image $F^n(\Psi)$ for all $n \in \mathbb{N}$ are called **primeval**, and $F^\infty(\Psi)$ will denote the set of all primeval elements of Ψ. Here $F^n(\Psi)$ is considered as a set, i.e. multiplicities are ignored, for all $n \leqslant \infty$. Moreover, $F^n(\Psi)$ is a

flow-adapted simple point process for any $n \leqslant \infty$.

Random graphs will be used throughout this section. **However, random graphs in this section are not random networks in the sense of Section 2.4.** Connections with those types of networks will be studied in Section 5.4. In this section, a random (directed) graph $\mathbf{\Gamma}$ on G is specified with a random variable N taking values in $\mathbb{N} \cup \{\infty\}$ and random elements $(x_i)_{i\in\mathbb{N}}$ in G with $V(\mathbf{\Gamma}) := \{x_i : i < N\}$, and measurable indicators $(\xi_{ij})_{i,j\in\mathbb{N}}$ with $E(\mathbf{\Gamma}) := \{(x_i, x_j) : i, j < N, \xi_{ij} = 1\}$. A random subset C of vertices of $\mathbf{\Gamma}$ is a map on Ω taking values in the subsets of $V(\mathbf{\Gamma})$ such that $1_{\{x_i\in C\}}$ is measurable for each i. Similarly, a random (countable) collection $\mathcal{C} = \{C_i\}_{i<N^{\mathcal{C}}}$ of subsets of vertices of $\mathbf{\Gamma}$ is identified with a random variable $N^{\mathcal{C}}$ taking values in $\mathbb{N} \cup \{\infty\}$ and random subsets $(C_i)_{i\in\mathbb{N}}$ with the elements of $\mathcal{C}$ being defined as $\{C_i : i < N^{\mathcal{C}}\}$. In all cases of interest for the present study, the specific numbering of vertices in a random graph or elements of a random collection are of no interest and will not be given upon defining the graph or collection.

The adjective flow-adapted has already been defined for point processes and point-shifts. The same adjective will also be used for random graphs and for random collections. A random graph $\mathbf{\Gamma}$ on G is **flow-adapted** if for all $\omega \in \Omega$ and all $X, Y, z \in G$, one has $X \in V(\mathbf{\Gamma}(\omega))$ if and only if $zX \in V(\mathbf{\Gamma}(\theta_z\omega))$ and $(X, Y) \in E(\mathbf{\Gamma}(\omega))$ if and only if $(zX, zY) \in E(\mathbf{\Gamma}(\theta_z\omega))$. A random collection $\mathcal{C} = \{C_i\}_{i<N}$ is **flow-adapted** if for all $\omega \in \Omega, z \in G$, one has $N(\theta_z\omega) = N(\omega)$ and there is a permutation $\pi(\omega)$ of $\mathbb{N} \cap [0, N(\omega))$ such that $C_i(\theta_z\omega) = \{zx : x \in C_{\pi(\omega)}(\omega)\}$ for each $i < N(\omega)$. That is, $\mathcal{C}(\theta_z\omega)$

contains the same elements as $C(\omega)$, shifted by z, and possibly enumerated in a different order. Now the random graph generated by the point-shift F is defined.

Definition 5.2.1. Define $\mathbf{\Gamma}^F$, the **random graph generated by** F to have vertices at the points of Ψ and directed edges from each $X \in \Psi$ to $F(X)$.

Two natural equivalence relations on the vertices of $\mathbf{\Gamma}^F$ are defined by connected components and foils.

The set of undirected connected components of $\mathbf{\Gamma}^F$ is denoted by C^F and the component of $X \in \Psi$ is denoted $C^F(X)$. Then $X, Y \in \Psi$ are in the same component if and only if there are $n, m \in \mathbb{N}$ such that $F^m(X) = F^n(Y)$. That is, $C^F(X)$ is the set of all **relatives** of X. The graph $\mathbf{\Gamma}^F$ is flow-adapted, and hence so is C^F.

The **foliation** $\mathcal{L}^F$ is defined to be the set of **foils** $L^F(X)$ of F for $X \in \Psi$, which are equivalence classes under the equivalence relation where $X, Y \in \Psi$ are equivalent if and only if there is $n \in \mathbb{N}$ such that $F^n(X) = F^n(Y)$. That is, $L^F(X)$ is the relatives of X from the same **generation** as X. The foliation $\mathcal{L}^F$ is flow-adapted, and $\mathcal{L}^F$ is a subdivision of C^F. For a foil L, also denote $L_+ := L^F(F(X))$ for any $X \in L$. Note that if $X, X' \in L$ then $L^F(F(X)) = L^F(F(X'))$ so L_+ is well-defined. If there is $Y \in \Psi$ such that $F(Y) \in L$, then set $L_- := L^F(Y)$. Then L_- is well-defined because if Y, Y' are both such that $F(Y), F(Y') \in L$, then $L(Y) = L(Y')$. It holds that $(L_+)_- = L$ and when L_- exists $(L_-)_+ = L$.

It will be important later to know that the graph $\mathbf{\Gamma}^F$ is locally finite. The following result, generalizing one in [BHM18], guarantees this. It crucially relies on the unimodularity of G.

Proposition 5.2.2. *Let $D_n(X)$ denote the n-th order descendants of X, i.e. $D_n(X) := \{Y \in \Psi : F^n(Y) = X\}$. Also let $D(X) := \bigcup_{n=1}^{\infty} D_n(X)$ be all descendants of X. Then with $d_n(X) := \#D_n(X)$, $d(X) := \#D(X)$, one has for every $n \geqslant 0$ that $\mathbf{E}^{\Psi}[d_n(e)] = 1$. In particular, $d_n(e)$ is $\mathbf{P}^{\Psi}$-a.s. finite, or equivalently $\mathbf{P}$-a.s. every $X \in \Psi$ has $d_n(X)$ finite. If, in addition, $\mathbf{\Gamma}^F$ is $\mathbf{P}^{\Psi}$-a.s. acyclic, then $\mathbf{E}^{\Psi}[d(e)] = \infty$.*

Proof. F^n is a point-shift in its own right, so Proposition 5.1.3 implies that $\mathbf{E}^{\Psi}[d_n(e)] = \mathbf{E}^{\Psi}[\#D_n(e)] = 1$ since G is unimodular. Thus $d_n(e) < \infty$, $\mathbf{P}^{\Psi}$-a.s., and hence $\mathbf{P}$-a.s. $d_n(X) < \infty$ for all $X \in \Psi$ by Proposition 5.1.1. Moreover, when $\mathbf{\Gamma}^F$ is acyclic, the D_n partition D and hence $\mathbf{E}^{\Psi}[d(e)] = \sum_{n=1}^{\infty} \mathbf{E}^{\Psi}[d_n(e)] = \infty$. □

The primary tool needed to prove the classification theorem follows. It says that it is not possible to extract finite subsets of infinite subsets of Ψ in a flow-adapted way. The proof is modified from the argument proving a similar result for unimodular networks given by Lemma 2.11 in [BHMK18].

Proposition 5.2.3. *Let $\mathfrak{N} = \{\mathfrak{N}_i\}_{i<N}$ be a flow-adapted collection of infinite measurable subsets of Ψ and let k be the number of i such that $e \in \mathfrak{N}_i$. Suppose that $\mathbf{E}^{\Psi}[k] < \infty$. If Ψ' is a measurable flow-adapted subset of Ψ for which $\mathbf{P}$-a.s. $\#(\Psi' \cap \mathfrak{N}_i) < \infty$ for each i, then $\mathbf{P}$-a.s. $\Psi' \cap \mathfrak{N}_i = \emptyset$ for all i. In particular, if $\Psi' \subseteq \bigcup \mathfrak{N}$, then $\mathbf{P}$-a.s. $\Psi' = \emptyset$.*

Proof. Define

$$\tau(\omega, x, y) := \sum_{i<N(\omega)} 1_{\{x,y\in\mathfrak{N}_i(\omega),\, y\in\Psi'(\omega)\}} \frac{1}{\#(\Psi'(\omega)\cap\mathfrak{N}_i(\omega))}.$$

The assumptions about flow-adaptedness of $\mathfrak{N}$, Ψ', and Ψ, imply that τ is diagonally invariant. Then one has $\int_G \tau(e,y)\,\Psi(dy) = k$ by construction since e is in k of the $\mathfrak{N}_i$. Also $\int_G \tau(x,e)\,\Psi(dx) = \infty$ if $e \in \Psi' \cap \mathfrak{N}_i$ for some i because the $\mathfrak{N}_i$ are infinite. But the mass-transport theorem implies

$$\mathbf{E}^{\Psi}\int_G \tau(x,e)\,\Psi(dx) = \mathbf{E}^{\Psi}\int_G \tau(e,y)\,\Psi(dy) = \mathbf{E}^{\Psi}[k] < \infty,$$

and thus it must be that $\mathbf{P}^{\Psi}$-a.s. $e \notin \Psi' \cap \mathfrak{N}_i$ for any i. Equivalently, $\mathbf{P}$-a.s. for all $X \in \Psi$ it holds that $X \notin \Psi' \cap \mathfrak{N}_i$ for any i. Since $\Psi' \cap \mathfrak{N}_i \subseteq \Psi$ for each i, it follows that $\mathbf{P}$-a.s. $\Psi' \cap \mathfrak{N}_i = \emptyset$ for all i. □

Note that, by Proposition 5.1.1, the condition $\mathbf{E}^{\Psi}[k] < \infty$ appearing in Proposition 5.2.3 is automatically satisfied if the $\mathfrak{N}_i$ are pairwise disjoint, or more generally if there is a constant n such that almost surely no $X \in \Psi$ appears in more than n of the $\mathfrak{N}_i$, as this would imply $k \leqslant n$, $\mathbf{P}^{\Psi}$-a.s.

More information follows about the structure of the locally finite graph $\mathbf{\Gamma}^F$. In particular, cycles in components are unique, infinite components are acyclic, foils in infinite components can be ordered like $\mathbb{N}$ or $\mathbb{Z}$ in a flow-adapted way, and F acts bijectively on the primeval elements.

Lemma 5.2.4. *$\mathbf{P}$-a.s. a connected component C of $\mathbf{\Gamma}^F$ is either an infinite tree or has exactly one (directed) cycle $K(C)$ for which for all $Y \in C$ there is $n \in \mathbb{N}$ such that $F^n(Y) \in K(C)$. Moreover, $\mathbf{P}$-a.s. there are no infinite components with a cycle.*

Proof. The fact that all elements in C are connected and have out-degree 1 implies there can be at most one cycle. If there are no cycles, then C must be infinite since applying F to any element repeatedly must never repeat an element. Otherwise there is one cycle $K(C)$ and connectedness implies for every $Y \in C$ there is $n \in \mathbb{N}$ with $F^n(Y) \in K(C)$.

Let $\mathfrak{N}$ be the set of infinite components of $\mathbf{\Gamma}^F$ with a cycle, and let $\Psi' \subseteq \bigcup \mathfrak{N}$ be the union of all the cycles of these components. Since cycles are finite, it follows that $\Psi' \cap C$ is finite for all components $C \in \mathfrak{N}$. By Proposition 5.2.3, $\Psi' = \emptyset$ and hence there are no infinite components with a cycle **P**-a.s. □

Within an infinite acyclic connected component $C \in \mathcal{C}^F$, it is possible to define an order, called the **foil order**, on the foils $\mathcal{L}^F(C)$ that are subsets of C. This is accomplished by declaring $L^F(X) < L^F_+(X)$ for all $X \in C$. When thinking of $F(X)$ as being the father of X, the order is that of seniority.

Lemma 5.2.5. *The foil order on an infinite acyclic component C is a total order on C isomorphic to either the order of $\mathbb{Z}$ or $\mathbb{N}$.*

Proof. Fix any $X \in C$. Let $L_0 := L^F(X)$ and recursively define $L_{n+1} := (L_n)_+$ and if it exists $L_{-n-1} := (L_{-n})_-$ for $n > 0$. Let L be a foil in C, then it must be that $L = L_i$ for some i. Indeed, let $Y \in L$ and by definition of connectedness choose n, m such that $F^n(Y) = F^m(X) \in L_m$. It then follows by induction that $Y \in L_{m-n}$, and hence $L = L^F(Y) = L_{m-n}$. Next it is shown that $i \mapsto L_i$ is injective. Suppose for contradiction that $L_j = L_{j+N}$. Then there are N pairs

(X_i, Y_{i+1}) with $X_i \in L_i, Y_{i+1} \in L_{i+1}$ such that $F(X_i) = Y_{i+1}$ for $j \leqslant i \leqslant j+N-1$. Since $L_j = L_{j+N}$ it follows that $X_j, Y_{j+N} \in L_j$. Hence it is possible to choose n such that $F^n(X_j) = F^n(Y_{j+N})$ and $F^n(X_i) = F^n(Y_i)$ for all $j+1 \leqslant i \leqslant j+N-1$. Assume by induction that for some k one has $F^N(F^n(X_j)) = F^{N-k}(F^n(Y_{j+k}))$. Then as long as $k+1 \leqslant N$,

$$\begin{aligned} F^N(F^n(X_j)) &= F^{N-k}(F^n(Y_{j+k})) \\ &= F^{N-k}(F^n(X_{j+k})) \\ &= F^{N-k-1}(F^n(F(X_{j+k}))) \\ &= F^{N-k-1}(F^n(Y_{j+k+1})). \end{aligned}$$

Since $F^N(F^n(X_j)) = F^{N-1}(F^n(F(X_j))) = F^{N-1}(F^n(Y_{j+1}))$ shows the base case $k = 1$ holds, the induction is complete. Therefore, one finds $F^N(F^n(X_j)) = F^0(F^n(Y_{j+N})) = F^n(X_j)$, contradicting that C is acyclic. Thus $i \mapsto L_i$ is injective. If there is a smallest foil L_{i_0} then $i \mapsto L_{i_0+i}$ is an order isomorphism with $\mathbb{N}$, otherwise $i \mapsto L_i$ is an order isomorphism with $\mathbb{Z}$. □

Lemma 5.2.6. *F restricts to a bijective point-shift $F|_{\Psi'}$ on the flow-adapted subprocess $\Psi' := F^\infty(\Psi)$ of primeval elements.*

Proof. F naturally restricts to a point-shift $F|_{\Psi'}$ on Ψ' because if $X \in F^\infty(\Psi)$ then $F(X) \in F^\infty(\Psi)$. By definition, primeval elements are in the image $F(\Psi)$, but moreover they are in the image $F(\Psi')$. Indeed, by Proposition 5.2.2, points in Ψ have only finitely many children. If $X \in \Psi'$ were such that none of its children were primeval, then there would be $n \in \mathbb{N}$ large enough

that none of X's children are in the image $F^n(\Psi)$. But then X would not be in $F^{n+1}(\Psi)$, contradicting that $X \in F^\infty(\Psi)$. Thus the restricted point-shift $F|_{\Psi'}$ is surjective. If Ψ' is not the empty process **P**-a.s. then it has nonzero and finite intensity so that surjectivity and injectivity are equivalent by Proposition 5.1.4 □

The main result of this chapter follows.

Theorem 5.2.7 (Cardinality Classification of a Component)**.** **P**-*a.s. each connected component C of $\mathbf{\Gamma}^F$ is in one of the three following classes:*

(i) ***Class F/F:*** *C is finite, and hence so is each of its F-foils. In this case, when denoting by $1 \leqslant n = n(C) < \infty$ the number of its foils:*

- *C has a unique cycle of length n.*
- *$F^\infty(\Psi) \cap C$ is the set of vertices of this cycle.*

(ii) ***Class I/F:*** *C is infinite and each of its F-foils is finite. In this case:*

- *C is acyclic.*
- *Each foil has a junior foil.*
- *$F^\infty(\Psi) \cap C$ is a unique **bi-infinite** path, i.e. a sequence $(X_n)_{n\in\mathbb{Z}}$ of points of Ψ such that $F(X_n) = X_{n+1}$ for all n.*

(iii) ***Class I/I:*** *C is infinite and all its F-foils are infinite. In this case:*

- *C is acyclic.*

- $F^\infty(\Psi) \cap C = \emptyset$.

Proof. The properties of finite components C are immediate, so only infinite components are considered. Recall that by Lemma 5.2.4 **P**-a.s. all infinite components are acyclic. Consider the collection $\mathfrak{N}$ of all infinite components that have both finite and infinite foils. Suppose $C \in \mathfrak{N}$. According to Proposition 5.2.2, all $X \in \Psi$ have only finitely many children, so that if L is an infinite foil, then L_+ is also infinite. It follows that there is a maximum finite foil L with respect to the foil order in C. Let $\Psi' \subseteq \bigcup \mathfrak{N}$ be the union of these maximum finite foils of each $C \in \mathfrak{N}$. By construction, $\Psi' \cap C$ is finite for each $C \in \mathfrak{N}$, so Proposition 5.2.3 implies $\Psi' = \emptyset$ and hence $\mathfrak{N} = \emptyset$, **P**-a.s. Thus **P**-a.s. each infinite component is either of class I/F or I/I.

Next, redefine $\mathfrak{N}$ to be the set of infinite foils L of Ψ, and let $\Psi' := F^\infty(\Psi)$. By construction $\Psi' \cap L$ is finite for each $L \in \mathfrak{N}$ because a foil cannot have multiple primeval elements. If $X \neq Y \in L$ were both primeval, then with n minimal such that $F^n(X) = F^n(Y)$ one finds the primeval element $F^n(X)$ is the image of two distinct primeval elements $F^{n-1}(X), F^{n-1}(Y)$, contradicting injectivity of $F|_{\Psi'}$ guaranteed by Lemma 5.2.6. Thus Proposition 5.2.3 implies **P**-a.s. $\Psi' \cap L = \emptyset$ for all infinite foils L, and hence **P**-a.s. $F^\infty(\Psi) \cap C \neq \emptyset$ implies C is of class I/F.

Conversely, it will be shown that if C is class I/F, then $F^\infty(\Psi) \cap C \neq \emptyset$. Indeed, redefine $\mathfrak{N}$ to be the collection of components C of class I/F that have a minimum foil in the foil order. Letting $\Psi' \subseteq \bigcup \mathfrak{N}$ be the union of

minimum foils in C, it holds that $\Psi' \cap C$ is the (finite) minimum foil in C for each $C \in \mathfrak{N}$. Thus Proposition 5.2.3 implies $\Psi' = \emptyset$ and hence $\mathfrak{N} = \emptyset$, **P**-a.s. Now consider a C of class I/F and an arbitrary foil L of C. Since L is finite there is a minimum n such that $F^n(L)$ is a single point. Let C_0 denote the subgraph of $\mathbf{\Gamma}^F$ of L together with all descendants of elements of L and all forefathers of elements of L up to $F^n(L)$. Then C_0 is an infinite connected graph with vertices of finite degree, and hence it contains an infinite simple path $(X_i)_{i \leqslant 0}$ with $F(X_i) = X_{i+1}$ for each $i < 0$ by König's infinity lemma (c.f. Theorem 6 in [K$\ddot{\text{o}}$90]). For $i > 0$, define $X_i := F^i(X_0)$. Then $(X_i)_{i \in \mathbb{Z}}$ is a bi-infinite path in C satisfying $F(X_i) = X_{i+1}$ for all $i \in \mathbb{Z}$, and thus $\{X_i\}_{i \in \mathbb{Z}} \subseteq F^\infty(\Psi) \cap C$, in particular showing $F^\infty(\Psi) \cap C \neq \emptyset$. It also holds that $F^\infty(\Psi) \cap C \subseteq \{X_i\}_{i \in \mathbb{Z}}$ since for any $X \in F^\infty(\Psi) \cap C$ it is possible to choose n, m such that $F^n(X) = F^m(X_0) = X_m$. Uniqueness of primeval children then implies $X = X_{m-n}$. It follows that $F^\infty(\Psi) \cap C = \{X_i\}_{i \in \mathbb{Z}}$.

Thus it is shown that **P**-a.s. infinite components C are class I/F if and only if $F^\infty(\Psi) \cap C \neq \emptyset$ and in this case $F^\infty(\Psi) \cap C$ is a unique bi-infinite sequence $\{X_i\}_{i \in \mathbb{Z}}$ satisfying $F(X_i) = X_{i+1}$. Since I/F and I/I are the only possible choices, by process of elimination it follows that **P**-a.s. infinite components C are of class I/I if and only if $F^\infty(\Psi) \cap C = \emptyset$. □

5.2.2. A Counterexample on a Non-unimodular Group

This example serves to show that the cardinality classification (Theorem 5.2.7) does not hold for non-unimodular spaces. It is an open question

whether a more general classification for such spaces exists. Recall the standard first example of a non-unimodular group: the $ax + b$ group. In this section,

$$G = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a > 0, b \in \mathbb{R} \right\}$$

with matrix multiplication and the topology inherited from $\mathbb{R}^4$. G is identified with the right half-plane in $\mathbb{R}^2$ by identifying (a, b) with $\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$. In this notation

$$(a, b)(c, d) = (ac, ad + b), \qquad (a, b)^{-1} = (\frac{1}{a}, -\frac{b}{a}).$$

Then, cf. [HR79] Example 15.17 (g), G has a left-invariant Haar measure

$$\lambda(B) = \iint_B \frac{1}{a^2}\, da\, db$$

and modular function

$$\Delta(a, b) = \frac{1}{a}.$$

Let Ψ be a homogeneous Poisson point process on G with intensity $\gamma \in (0, \infty)$. Necessarily Ψ is G-stationary and simple. For all $(a, b) \in G$ define the strip

$$S(a, b) := [a, \infty) \times [b - \delta a, b + \delta a]$$

for some fixed $\delta > 0$. Note that the definition is chosen so $(a, b)S(1, 0) = S(a, b)$, where here $(1, 0) = e \in G$. Moreover, for any $(a, b) \in G$,

$$\lambda(S(a, b)) = \int_{b-\delta a}^{b+\delta a} \int_a^\infty \frac{1}{x^2}\, dx\, dy = \frac{1}{a} \cdot ((b + \delta a) - (b - \delta a)) = 2\delta$$

so in particular $\Psi(S(a, b)) < \infty$ a.s. By the Slivnyak-Mecke theorem (see Theorem B.1.6 in Appendix B), $\Psi^! := \Psi - \delta_e$ is Poisson under $\mathbf{P}^\Psi$ with

$\mathbf{E}^{\Psi}[\Psi^!(B)] = \gamma\lambda(B)$. Hence $\mathbf{E}^{\Psi}[\Psi^!(S(1,0))] = 2\delta\gamma$ and therefore $\Psi(S(1,0)) < \infty$, $\mathbf{P}^{\Psi}$-a.s. Equivalently, $\mathbf{P}$-a.s. $\Psi(S(X)) < \infty$ for all $X \in \Psi$ by Proposition 5.1.1. This leads to the **strip point-shift** F where $F(X)$ is defined to be the right-most point of Ψ in $S(X)$ for each $X \in \Psi$. One may theoretically resolve ties for right-most point using the lexicographic order on $\mathbb{R}^2$, but the interested reader may note that results in Section 5.3.4 will show that there is no need because almost surely each point $X \in \Psi$ has a unique first coordinate.

Now suppose that $2\delta\gamma < 1$. It will be shown that $\mathbf{P}$-a.s. $F^n(X)$ eventually becomes constant as $n \to \infty$ for all $X \in \Psi$. It suffices to show that under $\mathbf{P}^{\Psi}$ it holds that $F^n(e)$ eventually becomes constant. Recall that the ***n*-th factorial power** of a counting measure μ with representation $\mu = \sum_i \delta_{x_i}$ is defined as $\mu^{(n)} := \sum_{i_1 \neq \cdots \neq i_n} \delta_{(x_{i_1}, \ldots, x_{i_n})}$, where the notation $i_1 \neq \cdots \neq i_n$ means

that $i_1, \ldots, i_n$ are all distinct. Then

$$\begin{aligned}
&\mathbf{E}^{\Psi} \sum_{k=0}^{\infty} \Psi(S(F^k(e)) \setminus \{F^k(e)\}) \\
&= \sum_{k=0}^{\infty} \mathbf{E}^{\Psi} \Psi(S(F^k(e)) \setminus \{F^k(e)\}) \\
&\leqslant \sum_{k=0}^{\infty} \mathbf{E}^{\Psi} \int_{G^{k+1}} 1_{\{x_1 \in S(e)\}} \cdots 1_{\{x_{k+1} \in S(x_k)\}} \, (\Psi^!)^{(k+1)}(dx_1 \times \cdots \times dx_{k+1}) \\
&= \sum_{k=0}^{\infty} \mathbf{E} \int_{G^{k+1}} 1_{\{x_1 \in S(e)\}} \cdots 1_{\{x_{k+1} \in S(x_k)\}} \, \Psi^{(k+1)}(dx_1 \times \cdots \times dx_{k+1}) \\
&= \sum_{k=0}^{\infty} \int_{G^{k+1}} 1_{\{x_1 \in S(e)\}} \cdots 1_{\{x_{k+1} \in S(x_k)\}} \gamma^{k+1} \, \lambda(dx_{k+1}) \cdots \lambda(dx_1) \\
&= \sum_{k=0}^{\infty} (2\delta)^{k+1} \gamma^{k+1} \\
&< \infty,
\end{aligned}$$

where here the Slivnyak-Mecke theorem is used again, along with the fact that the factorial moment measures of a Poisson point process are just powers of the intensity measure, cf. Example 9.5 (d) in [DVJ08]. Thus it must be that $\Psi(S(F^k(e)) \setminus \{F^k(e)\}) = 0$ for all k large, $\mathbf{P}^{\Psi}$-a.s. That is, there are no points of Ψ in $S(F^k(e))$ besides $F^k(e)$ itself. Consequently, $F^k(e)$ is a fixed point of F for large k and $F^k(e)$ is thus eventually constant in k. Equivalently, $\mathbf{P}$-a.s. for every $X \in \Psi$ it holds that $F^k(X)$ is eventually constant in k.

Next it will be shown that every fixed point of F is the image of infinitely many $X \in \Psi$. Again it is enough to show under $\mathbf{P}^{\Psi}$ that if $F(e) = e$ then e is the image of infinitely many $X \in \Psi$. This is accomplished by finding

a region of points $(x, y) \in G$ such that

(i) $(1, 0) \in S(x, y)$, and

(ii) $S(x, y) \cap ([1, \infty) \times \mathbb{R}) \subseteq S(1, 0)$,

which implies F would map a point of Ψ at (x, y) to $(1, 0)$. The condition (i) says $1 \geqslant x$ and $y - \delta x \leqslant 0 \leqslant y + \delta x$, i.e. $-\delta x \leqslant y \leqslant \delta x$. Condition (ii) is guaranteed if $[y - \delta x, y + \delta x] \subseteq [-\delta, \delta]$, i.e. if $y \geqslant \delta(x - 1)$ and $y \leqslant \delta(1 - x)$. The constraints

$$0 < x \leqslant 1, \quad -\delta x \leqslant y \leqslant \delta x, \quad y \leqslant \delta(1 - x), \quad y \geqslant \delta(x - 1),$$

bound a parallelogram D with corners

$$(0, 0), \quad (1/2, \delta/2), \quad (1, 0), \quad (1/2, -\delta/2).$$

Then

$$\mathbf{E}^{\Psi}[\Psi^{!}(D)] = \gamma\lambda(D) \geqslant \gamma \int_0^{1/2} \int_{-\delta x}^{\delta x} \frac{1}{x^2}\, dy\, dx = \gamma \int_0^{1/2} \frac{2\delta}{x}\, dx = \infty$$

so that the region D contains infinitely many points of Ψ, $\mathbf{P}^{\Psi}$-a.s. By construction, if $F(e) = e$ then every $X \in \Psi \cap D$ has $F(X) = e$, proving the claim.

Putting previous claims together, it holds that the foils and connected components are identical because every component contains a fixed point, and the foils and components are in bijection with the fixed points of F. The connected component of a fixed point Y of F is all $X \in \Psi$ that are eventually

sent to Y. Thus all components and foils are infinite (class I/I). However, the components are not acyclic and $F^\infty(\Psi) = \{X \in \Psi : F(X) = X\} \neq \emptyset$, contrary to what the classification theorem would suggest for unimodular G. It follows that the properties of the cardinality classification cannot be extended beyond the case of unimodular G.

5.3. Properties of Point-shifts

5.3.1. Point-maps

One may define a point-shift on a point process Ψ indirectly by specifying where e is mapped to under $\mathbf{P}^\Psi$ instead of where every point is mapped under $\mathbf{P}$. Some properties of point-shifts that are more easily phrased in terms of these so-called point-maps are collected in this section. Fix a point-shift F on a flow-adapted simple point process Ψ with intensity $\gamma \in (0, \infty)$ for the remainder of the section.

A **point-map** on Ψ is a measurable map $f : \Omega \to G$ such that $f(\omega) \in \Psi(\omega)$ for $\mathbf{P}^\Psi$-a.e. $\omega \in \Omega$. There is a natural correspondence between point-shifts and point-maps. Namely, if F is a point-shift, then $f(\omega) := F(\omega, e)$ is a point-map, and if f is a point-map, then $F(\omega, X) := Xf(\theta_X^{-1}\omega)$ is a point-shift, and these operations are inverses. Proposition 5.1.1 ensures that changing the definition of F on a $\mathbf{P}$-null set will only change the corresponding f on a $\mathbf{P}^\Psi$-null set, and vice-versa. The unfamiliar reader may see Example B.1.5 in Appendix B for the details of how to convert definitions under $\mathbf{P}$ and $\mathbf{P}^\Psi$ more generally.

Let $\boldsymbol{F}^{-}(x) := \{Y \in \Psi : F(Y) = x\}$ be the **in-neighbors** of x for each $x \in G$. Further define $\boldsymbol{f}^{-} := \boldsymbol{F}^{-}(e)$ to be the **in-neighbors** of e. These two definitions also contain the same information by Proposition 5.1.1. The boldface in $\boldsymbol{f}^{-}, \boldsymbol{F}^{-}$ is intended to indicate to the reader that the result is a subset of G and should not be confused with, e.g., an expression like f^{-1}, which is an element of G, namely the group inverse of the element f.

Suppose that F is bijective. In this case, it makes sense to run the point-shift backwards in time. Define the **reverse point-shift** $F^{\leftarrow}$ by letting $F^{\leftarrow}(X)$ be the unique element of $\boldsymbol{F}^{-}(X)$ for all $X \in \Psi$. Similarly define the **reverse point-map** $f^{\leftarrow}$ to be the point-map corresponding to $F^{\leftarrow}$. Note that **P**-a.s.

$$F(F^{\leftarrow}(X)) = F^{\leftarrow}(F(X)) = X, \qquad X \in \Psi. \tag{5.7}$$

That is, F and $F^{\leftarrow}$ are inverses (as point-shifts, not as group elements) on the support of Ψ.

With the mass-transport theorem and Proposition 5.1.1, the following may be obtained in a straightforward manner.

Proposition 5.3.1. *The following hold:*

(a) **P***-a.s. every $X \in \Psi$ is the image under F of at least (resp. at most) k distinct points of Ψ if and only if $\mathbf{P}^{\Psi}$-a.s. $\#\boldsymbol{f}^{-} \geqslant k$ (resp. $\leqslant k$).*

(b) **P***-a.s. every $X \in \Psi$ is the image under F of finitely (resp. infinitely) many distinct points of Ψ if and only if $\mathbf{P}^{\Psi}$-a.s. $\#\boldsymbol{f}^{-} < \infty$ (resp. $= \infty$).*

(c) **P**-*a.s. F is bijective (resp. surjective, injective) if and only if* $\mathbf{P}^{\Psi}$-*a.s.* $\#\boldsymbol{f}^{-} = 1$ *(resp.* $\geqslant 1, \leqslant 1$*).*

(d) *For all* $\zeta : \Omega \to \mathbb{R}_{\geqslant 0}$ *measurable,*

$$\mathbf{E}^{\Psi}[\zeta(\theta_f^{-1})\Delta(f^{-1})] = \mathbf{E}^{\Psi}[\zeta \cdot \#\boldsymbol{f}^{-}]. \tag{5.8}$$

In particular, the following mass-flow relationship for point-shifts holds

$$\mathbf{E}^{\Psi}[\Delta(f^{-1})] = \mathbf{E}^{\Psi}[\#\boldsymbol{f}^{-}]. \tag{5.9}$$

(e) **P**-*a.s. every* $X \in \Psi$ *is the image under F of at least (resp. at most) k points of* Ψ *if and only if for all* $\zeta : \Omega \to \mathbb{R}_{\geqslant 0}$ *measurable*

$$\mathbf{E}^{\Psi}[\zeta(\theta_f^{-1})\Delta(f^{-1})] \geqslant k\mathbf{E}^{\Psi}[\zeta] \qquad (\textit{resp.} \ \leqslant k\mathbf{E}^{\Psi}[\zeta]).$$

(f) *(Test for Bijectivity)*[1] *F is bijective if and only if for all* $\zeta : \Omega \to \mathbb{R}_{\geqslant 0}$ *measurable*

$$\mathbf{E}^{\Psi}[\zeta(\theta_f^{-1})]\Delta(f^{-1})] = \mathbf{E}^{\Psi}[\zeta]. \tag{5.10}$$

(g) *If F is bijective, also*

$$\mathbf{E}^{\Psi}\left[f(\theta_f^{-1})\right] = \mathbf{E}^{\Psi}\left[\frac{f}{\Delta(f^{\leftarrow})}\right]. \tag{5.11}$$

(h) *If* **P**-*a.s. every* $X \in \Psi$ *is the image under F of at least (resp. at most) k points of* Ψ*, then* $\mathbf{E}^{\Psi}[\Delta(f^{-1})] \geqslant k$ *(resp.* $\leqslant k$*).*

[1] G. Last also proves this and similar results, e.g. Corollary 10.1 in [Las10a].

(i) *If* $\mathbf{E}^{\Psi}[\Delta(f^{-1})] < \infty$, $\mathbf{P}$*-a.s. every* $X \in \Psi$ *is the image of only finitely many* $Y \in \Psi$ *under* F.

(j) *If* $\mathbf{E}^{\Psi}[\Delta(f^{-1})] = 1$, *then* F *is injective if and only if it is surjective. In particular, this is automatic if* G *is unimodular.*

(k) *If* $\mathbf{E}^{\Psi}[\Delta(f^{-1})] > 1$ *(resp.* < 1*), then* F *is not injective (resp. not surjective).*

Proof.

(a),(b),(c): Direct application of Proposition 5.1.1.

(d): Apply the mass-transport theorem with the diagonally invariant function

$$\tau(\omega, x, y) := f(\theta_y^{-1}\omega)1_{\{x,y\in\Psi(\omega),y=F(x)\}}\Delta(y^{-1}x).$$

(e): Apply *(a)* and *(d)*.

(f): Apply *(e)* with $k := 1$.

(g): Replace ζ with $\frac{\zeta}{\Delta(f^{\leftarrow})}$ in *(d)* and use the fact that $\mathbf{P}^{\Psi}$-a.s.

$$f^{\leftarrow}(\theta_f^{-1}) = f^{-1} \cdot (f f^{\leftarrow}(\theta_f^{-1})) = f^{-1}F^{\leftarrow}(f) = f^{-1}F^{\leftarrow}(F(e)) = f^{-1}.$$

(h),(i): Take $\zeta := 1$ in *(d)* and apply *(a)* or *(b)*.

(j): Take $\zeta := 1$ in *(d)*. Use *(a)*, *(c)*, and the fact that a random variable bounded above (or below) by 1 with expectation 1 must be constant 1 a.s.

(k): From (5.9), the hypothesis implies that $\#f^- > 1$ (resp. < 1) with positive $\mathbf{P}^\Psi$-probability. But if F were injective (resp. surjective), then one would have $\#f^- \leqslant 1$ (resp. $\geqslant 1$) $\mathbf{P}^\Psi$-a.s. □

5.3.2. Mecke's Invariance Theorem

In the case of $G = \mathbb{R}^d$, Mecke's invariance theorem shows that Palm probabilities are preserved under bijective point-shifts. Even stronger, a point-shift is bijective if and only if it preserves Palm probabilities. It will be shown in Corollary 5.3.2 that, if G is unimodular, then this still holds. However, for non-unimodular G this is not so. Precisely, the notion of isomodularity will be introduced, and it will be shown that, amongst bijective point-shifts, isomodular ones are exactly those that preserve Palm probabilities (Proposition 5.3.7).

For the rest of the section, fix a flow-adapted simple point process Ψ of intensity $\gamma \in (0, \infty)$, and a point-map f with associated point-shift F. The simple case of Mecke's invariance theorem when G is unimodular follows.

Corollary 5.3.2 (Mecke's Invariance Theorem)**.** *Suppose that G is unimodular. Then F preserves $\mathbf{P}^\Psi$ if and only if F is bijective. That is, $\mathbf{P}^\Psi(\theta_f^{-1} \in A) = \mathbf{P}^\Psi(A)$ for all $A \in \mathcal{F}$ if and only if F is bijective.*

Proof. Apply Proposition 5.3.1 (f), the test for bijectivity, and use the fact that $\Delta(x) = 1$ for all $x \in G$. □

With Mecke's invariance theorem for unimodular G in place, one

may ask about non-unimodular G. For these G, which bijective point-shifts preserve Palm probabilities? Equation (5.10) shows that the obstruction is the factor $\Delta(f^{-1})$. This motivates the definition of isomodularity, which says that a point-shift preserves the value of $\Delta(X)$ for each $X \in \Psi$.

Definition 5.3.3. The point-shift F is said to be **isomodular** if **P**-a.s. $\Delta(F(X)) = \Delta(X)$ for all $X \in \Psi$.

Isomodularity is a special case of invariance of a subgroup under F, which is defined presently.

Definition 5.3.4. A measurable subgroup $H \in \mathcal{B}(G)$ of G is called **F-invariant** if **P**-a.s. $F(X)$ is in the same coset as X for all $X \in \Psi$.

Isomodularity of F is the same as the assumption that the subgroup $\{\Delta = 1\}$ is F-invariant. Also note that if G is unimodular, then F is automatically isomodular.

A brief detour is taken to go through the equivalent descriptions of F-invariance under **P** and $\mathbf{P}^\Psi$.

Proposition 5.3.5. *Let $H \in \mathcal{B}(G)$ a measurable subgroup of G, and for each $x \in G$ let $[x] := xH$ denote the coset of x. Then the following are equivalent:*

(a) H is F-invariant, i.e. **P***-a.s. $[F(X)] = [X]$ for all $X \in \Psi$.*

(b) $\mathbf{P}^\Psi$*-a.s. $[f] = [e]$.*

If F is bijective, the previous statements are also equivalent to:

(c) $\mathbf{P}$*-a.s.* $[F^{\leftarrow}(X)] = [X]$ *for all* $X \in \Psi$.

(d) $\mathbf{P}^{\Psi}$*-a.s.* $[f^{\leftarrow}] = [e]$.

Proof.

(a) $\iff$ *(b)*: The equivalence follows from Proposition 5.1.1, so that $\mathbf{P}^{\Psi}$-a.s. $[f] = [e]$ is equivalent to $\mathbf{P}$-a.s. $[f(\theta_X^{-1})] = [e]$ for all $X \in \Psi$, which is the same as $[F(X)] = [X]$ after multiplying by X.

(a) $\iff$ *(c)*: Using that F and $F^{\leftarrow}$ are inverses, replace X with $F^{\leftarrow}(X)$ in *(b)* to get *(c)* or replace X with $F(X)$ in *(c)* to get *(b)*.

(c) $\iff$ *(d)*: The proof is the same as *(a)* $\iff$ *(b)*. □

Since isomodularity plays an important role in what follows, the previous result is restated for $H := \{\Delta = 1\}$ in the bijective case.

Corollary 5.3.6. *Let F be bijective, then the following are equivalent:*

(a) F is isomodular, i.e. $\mathbf{P}$*-a.s.* $\Delta(F(X)) = \Delta(X)$ *for all* $X \in \Psi$.

(b) $\mathbf{P}^{\Psi}$*-a.s.* $\Delta(f) = 1$.

(c) $\mathbf{P}$*-a.s.* $\Delta(F^{\leftarrow}(X)) = \Delta(X)$ *for all* $X \in \Psi$.

(d) $\mathbf{P}^{\Psi}$*-a.s.* $\Delta(f^{\leftarrow}) = 1$. □

Now the question of which bijective point-shifts preserve Palm probabilities is answerable.

Proposition 5.3.7. *Suppose F is bijective. Then F preserves $\mathbf{P}^\Psi$ if and only if F is isomodular. That is, $\mathbf{P}^\Psi(\theta_f^{-1} \in A) = \mathbf{P}^\Psi(A)$ for all $A \in \mathcal{F}$ if and only if F is isomodular.*

Proof. Suppose F is isomodular. Then $\Delta(f^\leftarrow) = 1$, $\mathbf{P}^\Psi$-a.s. by Corollary 5.3.6. Hence (5.11) immediately implies F preserves $\mathbf{P}^\Psi$. If F is not isomodular, at least one of $\mathbf{P}^\Psi(\Delta(f^\leftarrow) > 1)$ and $\mathbf{P}^\Psi(\Delta(f^\leftarrow) < 1)$ is strictly positive. The cases are nearly identical, so assume $\mathbf{P}^\Psi(\Delta(f^\leftarrow) > 1) > 0$ and let $A := \{\Delta(f^\leftarrow) > 1\}$. Then take $f := 1_A$ in (5.11) to find

$$\mathbf{P}^\Psi(\theta_f^{-1} \in A) = \mathbf{E}^\Psi\left[\frac{1_{\{\Delta(f^\leftarrow)>1\}}}{\Delta(f^\leftarrow)}\right] < \mathbf{E}^\Psi\left[1_{\{\Delta(f^\leftarrow)>1\}}\right] = \mathbf{P}^\Psi(A),$$

showing that $\mathbf{P}^\Psi$ is not preserved. □

5.3.3. Reciprocal and Reverse of a Point-map

In this section, a curious interplay between the reverse $f^\leftarrow$ and the reciprocal f^{-1} of a point-map is investigated, and a characterization of when the two have the same law under $\mathbf{P}^\Psi$ is given. The notation of the previous section is retained. That is, Ψ is a flow-adapted simple point process of intensity $\gamma \in (0, \infty)$, and F is a point-shift with associated point-map f. Next follows another result along the lines of Proposition 5.3.1 (f) and (g) which sparks interest in the distributional relationship between f^{-1} and $f^\leftarrow$.

Corollary 5.3.8. *Suppose F is bijective. For all $\zeta : G \to \mathbb{R}_{\geqslant 0}$ measurable it holds that*

$$\mathbf{E}^{\Psi}\left[\zeta(f^{-1})\Delta(f^{-1})\right] = \mathbf{E}^{\Psi}\left[\zeta(f^{\leftarrow})\right], \tag{5.12}$$

$$\mathbf{E}^{\Psi}\left[\zeta(f^{-1})\right] = \mathbf{E}^{\Psi}\left[\frac{\zeta(f^{\leftarrow})}{\Delta(f^{\leftarrow})}\right]. \tag{5.13}$$

Proof. Use the fact that $\mathbf{P}^{\Psi}$-a.s.

$$f^{\leftarrow}(\theta_f^{-1}) = f^{-1} \cdot (f f^{\leftarrow}(\theta_f^{-1})) = f^{-1}F^{\leftarrow}(f) = f^{-1}F^{\leftarrow}(F(e)) = f^{-1}$$

and replace ζ by $\zeta(f^{\leftarrow})$ in each of (5.8) and (5.11). □

One sees in (5.13) that non-unimodularity of G is, as usual, an obstruction. Two more results relating the distributions of $\Delta(f^{\leftarrow})$ and $\Delta(f^{-1})$ are given. Then it is shown in Proposition 5.3.11 that, amongst bijective point-shifts, the isomodular ones are precisely those for which f^{-1} and $f^{\leftarrow}$ have the same distribution under $\mathbf{P}^{\Psi}$. Recall that this is also the class of point-shifts that preserve Palm probabilities by Proposition 5.3.7.

Corollary 5.3.9. *Let F be bijective, then for all $r > 0$ it holds that*

$$r\mathbf{P}^{\Psi}(\Delta(f^{-1}) = r) = \mathbf{P}^{\Psi}(\Delta(f^{\leftarrow}) = r),$$

and, if this number is strictly positive, then for all $A \in \mathcal{F}$

$$\mathbf{P}^{\Psi}(\theta_f^{-1} \in A \mid \Delta(f^{-1}) = r) = \mathbf{P}^{\Psi}(A \mid \Delta(f^{\leftarrow}) = r).$$

Proof. Fix $r > 0$ and take $\zeta(x) := 1_{\{\Delta(x)=r\}}$ in (5.13). One finds

$$\mathbf{P}^{\Psi}(\Delta(f^{-1}) = r) = \frac{1}{r}\mathbf{P}^{\Psi}(\Delta(f^{\leftarrow}) = r) =: p$$

showing the first claim. Supposing that $p > 0$, take $\zeta := 1_A 1_{\{\Delta(f^{\leftarrow})=r\}}$ in (5.11) and use that $\mathbf{P}^{\Psi}$-a.s. $f^{\leftarrow}(\theta_f^{-1}) = f^{-1}$ to find

$$\mathbf{P}^{\Psi}(\theta_f^{-1} \in A, \Delta(f^{-1}) = r) = \frac{1}{r}\mathbf{P}^{\Psi}(A, \Delta(f^{\leftarrow}) = r).$$

Division by p finishes the proof. □

Lemma 5.3.10. *Let F be bijective, then for all $\alpha \in \mathbb{R}$ and $0 \leqslant r \leqslant s \leqslant \infty$ it holds that*

$$\mathbf{E}^{\Psi}\left[\Delta(f^{-1})^{\alpha} 1_{\{r \leqslant \Delta(f^{-1}) \leqslant s\}}\right] = \mathbf{E}^{\Psi}\left[\Delta(f^{\leftarrow})^{\alpha-1} 1_{\{r \leqslant \Delta(f^{\leftarrow}) \leqslant s\}}\right]. \tag{5.14}$$

Proof. Take $\zeta(x) := \Delta(x)^{\alpha} 1_{\{r \leqslant \Delta(x) \leqslant s\}}$ in (5.13). □

Proposition 5.3.11. *Let F be bijective, then f^{-1} and $f^{\leftarrow}$ have the same law under $\mathbf{P}^{\Psi}$ if and only if F is isomodular.*

Proof. Suppose F is isomodular. Then by Corollary 5.3.6, $\mathbf{P}^{\Psi}$-a.s. $\Delta(f) = \Delta(f^{\leftarrow}) = 1$ and thus (5.13) shows that f^{-1} and $f^{\leftarrow}$ have the same law under $\mathbf{P}^{\Psi}$.

Next suppose that f^{-1} and $f^{\leftarrow}$ have the same law under $\mathbf{P}^{\Psi}$. Then

$$\mathbf{E}^{\Psi}[\Delta(f^{-1})^{\alpha} 1_{\{r \leqslant \Delta(f^{-1}) \leqslant s\}}] = \mathbf{E}^{\Psi}[\Delta(f^{\leftarrow})^{\alpha} 1_{\{r \leqslant \Delta(f^{\leftarrow}) \leqslant s\}}] \tag{5.15}$$

for all $\alpha \in \mathbb{R}$ and all $0 \leqslant r \leqslant s \leqslant \infty$. But then for all $\alpha \in \mathbb{R}$ and all $0 \leqslant r \leqslant s \leqslant \infty$

$$\begin{aligned}
\mathbf{E}^{\Psi}[\Delta(f^{-1})^{\alpha+1} 1_{\{r \leqslant \Delta(f^{-1}) \leqslant s\}}] &= \mathbf{E}^{\Psi}[\Delta(f^{\leftarrow})^{\alpha} 1_{\{r \leqslant \Delta(f^{\leftarrow}) \leqslant s\}}] && \text{(by (5.14))} \\
&= \mathbf{E}^{\Psi}[\Delta(f^{-1})^{\alpha} 1_{\{r \leqslant \Delta(f^{-1}) \leqslant s\}}] && \text{(by (5.15))} \\
&= \mathbf{E}^{\Psi}[\Delta(f^{\leftarrow})^{\alpha-1} 1_{\{r \leqslant \Delta(f^{-1}) \leqslant s\}}]. && \text{(by (5.14))}
\end{aligned}$$

Taking $\alpha := 1, r := 1, s := \infty$

$$\mathbf{E}^{\Psi}[\Delta(f^{-1})^2 1_{\{1 \leqslant \Delta(f^{-1})\}}] = \mathbf{E}^{\Psi}[\Delta(f^{-1}) 1_{\{1 \leqslant \Delta(f^{-1})\}}]$$

which is absurd unless $\Delta(f^{-1}) \leqslant 1$, $\mathbf{P}^{\Psi}$-a.s. It also holds that with $\alpha := 1, r := 0, s := 1$,

$$\mathbf{E}^{\Psi}[\Delta(f^{-1})^2 1_{\{\Delta(f^{-1}) \leqslant 1\}}] = \mathbf{E}^{\Psi}[\Delta(f^{-1}) 1_{\{\Delta(f^{-1}) \leqslant 1\}}],$$

which is absurd unless $\Delta(f^{-1}) \geqslant 1$, $\mathbf{P}^{\Psi}$-a.s. It follows that $\Delta(f^{-1}) = 1$, $\mathbf{P}^{\Psi}$-a.s. By Corollary 5.3.6 the result follows. □

5.3.4. Separating Points of a Point Process

In this section a notion of a function separating points of a point process is introduced. For the remainder of the section, Ψ is a simple and flow-adapted point process of intensity $\gamma \in (0, \infty)$. Let S be a set, $\zeta : G \to S$, and suppose that $\mathbf{P}$-a.s. no distinct $X, Y \in \Psi$ have $\zeta(X) = \zeta(Y)$. In this case one says that ζ **separates points** of Ψ. Similarly, say that a fixed partition $\{B_i\}_{i \in J}$ of G **separates points** of Ψ if $\mathbf{P}$-a.s. no B_i contains more than 1 point of Ψ.

When separation of points occurs is studied by proving a general result concerning when there cannot be an n-tuple of distinct points of Ψ satisfying a given constraint. Recall again that $\mu^{(n)} = \sum_{i_1 \neq \cdots \neq i_n} \delta_{(x_{i_1}, \ldots, x_{i_n})}$ denotes the n-th factorial power of a measure $\mu = \sum_i \delta_{x_i}$, $\mu^! := \mu - \delta_e$ is μ with a point at e removed, and a measurable space is a set together with a σ-algebra.

Proposition 5.3.12. *Let (S, Σ) be a measurable space and fix $M \in \Sigma$. Let $Z : G \times G^n \to S$ be measurable, and suppose that for all $y = (y_1, \dots, y_n) \in (G \setminus \{e\})^n$, or more generally that for $\mathbf{E}^\Psi[(\Psi^!)^{(n)}]$-a.e. $y \in G^n$,*

$$\lambda(x \in G : Z(x, xy) \in M) = 0.$$

Then $\mathbf{P}$-a.s. no $n+1$ distinct $X, Y_1, \dots, Y_n \in \Psi$ have $Z(X, Y_1, \dots, Y_n) \in M$.

Proof. By straight calculations,

$$\begin{aligned}
&\mathbf{P}(\exists X \in \Psi, Y \in \Psi^{(n)} : (X, Y) \in \Psi^{(n+1)}, Z(X, Y) \in M) \\
&\leqslant \mathbf{E} \int_G 1_{\{\exists Y \in \Psi^{(n)} : \forall i, Y_i \neq x, Z(x,Y) \in M\}} \, \Psi(dx) \\
&\leqslant \mathbf{E} \int_G \Psi^{(n)}(\theta_e, \{y \in G^n : \forall i, y_i \neq x, Z(x, y) \in M\}) \, \Psi(dx) \\
&= \gamma \mathbf{E}^\Psi \int_G \Psi^{(n)}(\theta_x, \{y \in G^n, \forall i, y_i \neq x, Z(x, y) \in M\}) \, \lambda(dx) \\
&= \gamma \mathbf{E}^\Psi \int_G \Psi^{(n)}(\theta_e, \{x^{-1}y : y \in G^n, \forall i, y_i \neq x, Z(x, y) \in M\}) \, \lambda(dx) \\
&= \gamma \mathbf{E}^\Psi \int_G \Psi^{(n)}(\theta_e, \{y \in G^n, \forall i, xy_i \neq x, Z(x, xy) \in M\}) \, \lambda(dx) \\
&= \gamma \mathbf{E}^\Psi \int_G \int_G 1_{\{Z(x,xy) \in M\}} \, (\Psi^!)^{(n)}(dy) \, \lambda(dx) \\
&= \gamma \mathbf{E}^\Psi \int_G \lambda(x \in G : Z(x, xy) \in M) \, (\Psi^!)^{(n)}(dy) \\
&= 0,
\end{aligned}$$

where in the first equality the refined Campbell theorem, stated in Appendix B as Theorem B.1.1, is used. This proves the claim. □

Proposition 5.3.12 immediately gives a condition for separating points of Ψ.

Corollary 5.3.13 (Condition for Separating Points)**.** *Let* (S, Σ) *be a measurable space,* $\zeta : G \to S$ *measurable, and suppose for all* $y \neq e$*, or more generally for* $\mathbf{E}^{\Psi}[\Psi^!]$*-a.e.* $y \in G$*,*

$$\lambda(x \in G : \zeta(x) = \zeta(xy)) = 0.$$

Then ζ *separates points of* Ψ*. Implicit in the previous line is the assumption that the sets* $\{x \in G : \zeta(x) = \zeta(xy)\}$ *are measurable for all* $y \in G$*. This is automatic if* (S, Σ) *is a standard measurable space, i.e., if there is a bi-measurable bijection of* S *with a Polish space, or more generally if* $S \times S$ *has measurable diagonal.*

Proof. Take $n := 1$, $Z(x, y) := (\zeta(x), \zeta(y))$ for all $x, y \in G$, and take M to be the diagonal of $S \times S$, then apply Proposition 5.3.12. □

Corollary 5.3.13 generalizes the well-known theorem in $G = \mathbb{R}^d$ that a stationary point process has not two points equidistant from 0. That would be the case of $\zeta(x) := |x|$. Not all G have this property though. Indeed, if G is a countable group with the discrete distance $d(x, y) := 1_{\{x \neq y\}}$, then $\lambda(x \in G : d(x, e) = d(xy, e)) > 0$ for all $y \neq e$ so the result does not apply if G has more than one element.

The next results can be used to show that there is no need to resolve ties when defining a point-shift in some situations. Intuitively, if a set B is small from the typical point's perspective, then no shift of B will contain more than one point of Ψ.

Proposition 5.3.14. *Let $B \in \mathcal{B}(G)$ with $e \in B$. If $\mathbf{E}^{\Psi}[\Psi^{!}(B)] = 0$, then $\mathbf{P}$-a.s. for all $X \in \Psi$ it holds that $\Psi(Xb : b \in B) = 1$, i.e. X is the unique point of Ψ inside $\{Xb : b \in B\}$.*

Proof. The hypotheses imply $\mathbf{P}^{\Psi}$-a.s. $\Psi(B \setminus \{e\}) = 0$. By Proposition 5.1.1, $\mathbf{P}$-a.s. all $X \in \Psi$ are such that $T_{X^{-1}}\Psi(B \setminus \{e\}) = 0$, i.e. $\Psi(\{Xb : b \in B\} \setminus \{X\}) = 0$, and hence $\Psi(Xb : b \in B) = 1$. □

For example, recall the strip point-shift on the $ax + b$ group of Section 5.2.2. It was defined by sending a point X to the right-most point in a certain strip in the plane. In that case, take $B := \{1\} \times \mathbb{R}$ in the previous result to find that the points of Ψ have unique first coordinates. Hence there are no ties for right-most point.

Finally, the previous result is restated in the case that B is a subgroup and applied to see that the only way for F to preserve a small subgroup from the typical point's perspective is to act as the identity.

Corollary 5.3.15. *Let $H \in \mathcal{B}(G)$ a subgroup of G. If $\mathbf{E}^{\Psi}[\Psi^{!}(H)] = 0$, then the cosets of H separate points of Ψ.* □

Corollary 5.3.16. *Let $H \in \mathcal{B}(G)$ a subgroup of G. If $\mathbf{E}^{\Psi}[\Psi^{!}(H)] = 0$ but H is F-invariant for some point-shift F, then F is the identity point-shift $\mathbf{P}$-a.s.*

Proof. H being F-invariant means $F(X)$ and X are in the same coset for $X \in \Psi$, then by Corollary 5.3.15 F is the identity point-shift. □

Corollary 5.3.17. *Let $H \in \mathcal{B}(G)$ a subgroup of G. If $\lambda(H) = 0$ and Ψ is Poisson with intensity $\gamma \in (0, \infty)$, then the only F for which H is F-invariant is the identity.*

Proof. The Slivnyak-Mecke theorem, Theorem B.1.6 in Appendix B, implies that one has $\mathbf{E}^{\Psi}[\Psi^!(H)] = \gamma\lambda(H) = 0$ and Corollary 5.3.16 applies. □

5.4. Connections with Unimodular Networks

A general situation is given under which a point process, seen under its Palm probability measure and rooted at the identity, is a unimodular network. The reader may wish to glance again at Section 2.4 to recall the notation used here for random networks. Next, the appropriate notion of flow-adaptedness for networks must be given, then the result follows.

Suppose β is a map on Ω such that for all $\omega \in \Omega$, $\beta(\omega)$ is a network whose vertex set $V(\beta) \subseteq G$. Note that β is *not* a random network as defined in Section 2.4. It is, for each $\omega \in \Omega$, a network, not an isomorphism class of rooted networks. Then β is called **flow-adapted** if for all $z \in G$, $\beta(\theta_z\omega)$ is the **shift** $T_z\beta(\omega)$ of $\beta(\omega)$ by z, i.e. $V(\beta(\theta_z\omega)) = \{zX : X \in V(\beta(\omega))\}$, and $E(\beta(\theta_z\omega)) = \{\{zX, zY\} : \{X, Y\} \in E(\beta(\omega))\}$, and all marks are preserved.

Theorem 5.4.1. *Suppose G is unimodular. Let Ψ be a flow-adapted simple point process with intensity $\gamma \in (0, \infty)$. Let β be a map on Ω such that for all $\omega \in \Omega$, $\beta(\omega)$ is a network, and such that β satisfies*

(i) $V(\beta) = \Psi$,

(ii) β *is flow-adapted,*

(iii) $\omega \mapsto [\beta(\omega), e]$ *is measurable on the event* $\{e \in V(\beta)\}$.

Then $[\beta, e]$ *is a unimodular network under* $\mathbf{P}^\Psi$ *on the event* $\{e \in V(\beta)\} = \{e \in \Psi\}$.

Proof. The assumption (iii) implies $[\beta, e]$ is a random network under $\mathbf{P}^\Psi$ on $\{e \in V(\beta)\}$, so one only needs to check unimodularity. Let $g : \mathcal{G}_{**} \to \mathbb{R}_{\geqslant 0}$ be given. Then

$$\begin{aligned}
\mathbf{E}^\Psi \sum_{v \in V(\beta)} g[\beta, e, v] &= \mathbf{E}^\Psi \int_G g[\beta, e, x]\, \Psi(dx) \\
&= \mathbf{E}^\Psi \int_G g[\beta(\theta_y^{-1}), e, y^{-1}]\, \Psi(dy) \qquad \text{(mass-transport)} \\
&= \mathbf{E}^\Psi \int_G g[T_{y^{-1}}\beta, e, y^{-1}]\, \Psi(dy) \\
&= \mathbf{E}^\Psi \int_G g[\beta, y, e]\, \Psi(dy) \\
&= \mathbf{E}^\Psi \sum_{v \in V(\beta)} g[\beta, v, e],
\end{aligned}$$

showing unimodularity. □

Recall that the symbol T_z for $z \in G$, used in the previous result as the shift operator on networks, is also used as the shift operator on counting measures and point processes, which is how it is used in the following.

Definition 5.4.2. Let $[\boldsymbol{\Gamma}, \boldsymbol{o}]$ be a unimodular network. A G**-embedding** of $[\boldsymbol{\Gamma}, \boldsymbol{o}]$ with respect to a probability measure $\mathcal{P}$ on Ω is a map $\eta : \mathcal{G}_* \to \mathbf{M}$ with the following properties:

(i) η is measurable.

(ii) $e \in \eta[\Gamma, o]$ for all rooted networks (Γ, o).

(iii) There is a measurable function $\mathfrak{t} : \mathcal{G}_{**} \to G$ such that for every rooted network (Γ, o) and every vertex $v \in V(\Gamma)$,

$$\eta[\Gamma, v] = T_{\mathfrak{t}[\Gamma,o,v]^{-1}}\eta[\Gamma, o],$$
$$\mathfrak{t}[\Gamma, v, o] = \mathfrak{t}[\Gamma, o, v]^{-1}.$$

(iv) $\mathcal{P}$-almost surely, the map defined on $V(\boldsymbol{\Gamma})$ by $v \mapsto \mathfrak{t}[\boldsymbol{\Gamma}, \boldsymbol{o}, v]$ is a bijection between $V(\boldsymbol{\Gamma})$ and the support of $\eta[\boldsymbol{\Gamma}, \boldsymbol{o}]$.

Say that the Palm version of a point process Ψ is a G**-embedding** of $[\boldsymbol{\Gamma}, \boldsymbol{o}]$ if there is a G-embedding η with respect to $\mathbf{P}^{\Psi}$ such that $\mathbf{P}^{\Psi}$-a.s. $\Psi = \eta[\boldsymbol{\Gamma}, \boldsymbol{o}]$.

Two of the motivating open questions of this research are:

(i) For a fixed Ψ, is the Palm version of Ψ a G-embedding of some $[\boldsymbol{\Gamma}, \boldsymbol{o}]$?

(ii) For a fixed $[\boldsymbol{\Gamma}, \boldsymbol{o}]$, is there a Ψ such that the Palm version of Ψ is a G-embedding of $[\boldsymbol{\Gamma}, \boldsymbol{o}]$?

When G is unimodular, the answer to the first question is "yes" for Ψ if it is possible to draw a connected graph on Ψ in a flow-adapted way.

Definition 5.4.3. Call a flow-adapted simple point process Ψ **connectible in a flow-adapted way** if there exists a connected flow-adapted locally finite

graph $\mathbf{\Gamma} = \mathbf{\Gamma}(\omega)$ such that $V(\mathbf{\Gamma}) = \Psi$ almost surely and such that for some measurable enumeration $(Y_i)_{i\in\mathbb{N}}$ of the points of Ψ (where Y_i is defined arbitrarily if i exceeds $\#\Psi$), $1_{\{Y_i,Y_j\}\in E(\mathbf{\Gamma})}$ is measurable for all $i, j \in \mathbb{N}$ on the event $\{\#\Psi \geqslant i, j\}$.

Theorem 5.4.4. *Suppose G is unimodular and that Ψ is a flow-adapted simple point process with intensity $\gamma \in (0, \infty)$ that is connectible in a flow-adapted way. Then the Palm version of Ψ is a G-embedding of some unimodular network.*

Proof. Choose a random graph $\mathbf{\Gamma}$ witnessing the fact that Ψ is connectible. Consider some edge $g \in E(\mathbf{\Gamma})$ from $X \in \Psi$ to $Y \in \Psi$. Without loss of generality, assume the mark space $\Xi = G$. Let the mark of (X, g) be $X^{-1}Y$, and let $\beta(\omega)$ be the network with underlying graph $\mathbf{\Gamma}(\omega)$ and marks as just specified. Choose $\boldsymbol{o} := e$ and let $[\mathbf{\Gamma}, \boldsymbol{o}] := [\beta, e]$. Let ψ a rooted automorphism of $(\mathbf{\Gamma}, v)$ be given, where v is any vertex of $\mathbf{\Gamma}$. It will be shown that ψ is the identity. Assume that ψ fixes all vertices less than graph distance k from v. For each $Y \in \Psi$ of distance $k + 1$ from v, there is an $X \in \Psi$ of distance k from v and an edge g from X to Y. The mark of (X, g) is $X^{-1}Y$ and this must equal the mark of $\psi(X) = X$ between X and $\psi(Y)$. But this mark is $X^{-1}\psi(Y)$. Thus $X^{-1}Y = X^{-1}\psi(Y)$ so that $Y = \psi(Y)$. By induction and using that $\mathbf{\Gamma}$ is connected, one finds that ψ is the identity automorphism. By construction $V(\beta) = \Psi$, β is flow-adapted, and $\omega \mapsto [\beta(\omega), e]$ is measurable on the set $\{e \in \Psi\}$. By Theorem 5.4.1, $[\beta, e]$ is a unimodular network under $\mathbf{P}^{\Psi}$ on the set $\{e \in \Psi\}$. It will be shown that Ψ is a G-embedding of $[\mathbf{\Gamma}, \boldsymbol{o}]$.

On the set $\{e \in \Psi\}$ one has that Ψ can be reconstructed from $[\boldsymbol{\Gamma}, \boldsymbol{o}]$. The reconstruction procedure will be used to define a G-embedding η. Indeed, let $\mathfrak{t}[\Gamma, o, v] := \prod_{i=0}^{k-1} \xi_E(v_i, g_i)$ where $v_0 v_1 \cdots v_k$ is a path between the two arbitrary vertices o and v of Γ, and g_i is the edge $\{v_i, v_{i+1}\}$, assuming the product is independent of the path chosen between. Let $\mathfrak{t}[\Gamma, o, v] := e$ otherwise. The fact that $\prod_{i=0}^{k-1} \xi_E(v_i, g_i)$ is required to be independent of path implies that for any $v_1, v_2, v_3 \in V(\Gamma)$ one has $\mathfrak{t}[\Gamma, v_1, v_3] = \mathfrak{t}[\Gamma, v_1, v_2]\mathfrak{t}[\Gamma, v_2, v_3]$ and in particular that $\mathfrak{t}[\Gamma, v, o] = \mathfrak{t}[\Gamma, o, v]^{-1}$ for any $o, v \in V(\Gamma)$.

For $[\Gamma, o] \in \mathcal{G}_*$ such that (Γ, v) has no rooted automorphisms for any $v \in V(\Gamma)$, define

$$\eta[\Gamma, o] := \{\mathfrak{t}[\Gamma, o, v] : v \in V(\Gamma)\},$$

otherwise define $\eta[\Gamma, o] := \{e\}$. Also for $[\Gamma, o] \in \mathcal{G}_*$ such that (Γ, v) has no rooted automorphisms for any $v \in V(\Gamma)$, one has for each $v \in V(\Gamma)$ that

$$\begin{aligned}
\eta[\Gamma, v] &= \{\mathfrak{t}[\Gamma, v, v'] : v' \in V(\Gamma)\} \\
&= \left\{\mathfrak{t}[\Gamma, o, v]^{-1}\mathfrak{t}[\Gamma, o, v]\mathfrak{t}[\Gamma, v, v'] : v' \in V(\Gamma)\right\} \\
&= \left\{\mathfrak{t}[\Gamma, o, v]^{-1}\mathfrak{t}[\Gamma, o, v'] : v' \in V(\Gamma)\right\} \\
&= T_{\mathfrak{t}[\Gamma, o, v]^{-1}}\eta[\Gamma, o].
\end{aligned}$$

If $[\Gamma, o]$ is such that some $v \in V(\Gamma)$ is such that (Γ, v) has a rooted automorphism, then then same is true of $[\Gamma, v]$, so that $\eta[\Gamma, v] = \{e\} = T_{\mathfrak{t}[\Gamma, o, v]^{-1}}\eta[\Gamma, o]$ for the only vertex $v = o \in V(\Gamma)$.

One may then recover Ψ from $[\boldsymbol{\Gamma}, \boldsymbol{o}]$ on the set $\{e \in \Psi\}$ in the following way. Consider a path $v_0 v_1 \cdots v_k$ in $[\boldsymbol{\Gamma}, \boldsymbol{o}]$ starting and ending at the root.

Since there are no rooted automorphisms of $(\boldsymbol{\Gamma}, \boldsymbol{o})$ on the set $\{e \in \Psi\}$ one may uniquely choose $X_0 := e, X_1, \ldots, X_{k-1}, X_k := e \in G$ such that $\xi_E(v_i, g_i) = X_i^{-1} X_{i+1}$ for each i. Then $\prod_{i=0}^{k-1} \xi_E(v_i, g_i) = X_0^{-1} X_k = e$ regardless of the choice of $v_0 v_1 \cdots v_k$ so long as the path starts and ends at the root. It follows that for an arbitrary path $v_0 v_1 \cdots v_k$ one has that $\prod_{i=0}^{k-1} \xi_E(v_i, g_i)$ depends only on the endpoints v_0 and v_k. Thus, for any $X \in \Psi = V(\boldsymbol{\Gamma})$, consider a path starting at the root $\boldsymbol{o} \in \boldsymbol{\Gamma}$ and ending at $v := X$. Then $\mathrm{t}[\boldsymbol{\Gamma}, \boldsymbol{o}, v] = e^{-1} X = X$. Thus $\Psi = \{\mathrm{t}[\boldsymbol{\Gamma}, \boldsymbol{o}, v] : v \in V(\boldsymbol{\Gamma})\} = \eta[\boldsymbol{\Gamma}, \boldsymbol{o}]$ almost surely on the set $\{e \in \Psi\}$. Hence η is a witness to the fact that Ψ is a G-embedding of $[\boldsymbol{\Gamma}, \boldsymbol{o}] = [\beta, e]$ under $\mathbf{P}^{\Psi}$. □

The problem of finding which point processes Ψ have Palm versions that are G-embeddings of some unimodular network now reduces to finding which Ψ admit a connected flow-adapted locally finite graph on Ψ. It is conjectured that the requirement that Ψ be connectible in a flow-adapted way is automatic.

Conjecture 5.4.5. *Suppose G is unimodular. Then all flow-adapted simple point processes Ψ on G are connectible in a flow-adapted way.*

Finally, some special cases of the conjecture are known to hold.

Proposition 5.4.6. *Let Ψ be a flow-adapted simple point process of intensity $\gamma \in (0, \infty)$. Then Ψ is connectible in a flow-adapted way in any of the following situations:*

(a) G is compact.

(b) $G = \mathbb{R}^d$.

(c) There exists a point-shift F such that $\mathbf{\Gamma}^F$ is connected.

Proof. If G is compact then $\mathbf{P}$-a.s. $\Psi(G) < \infty$ and the complete graph on Ψ suffices. If $G = \mathbb{R}^d$, then the Delaunay graph of Ψ suffices. If there exists a point-shift F such that $\mathbf{\Gamma}^F$ is connected, then the graph $\mathbf{\Gamma}^F$ suffices. □

5.5. Bibliographical Comments

The starting points for this chapter were Last's [Las10a] for Palm theory on groups and an early version of Baccelli and Haji-Mirsadeghi's work [BHM18] for the original cardinality classification theorem for components of point-shifts of point processes on $\mathbb{R}^d$. The classification theorem was generalized to the case of vertex-shifts on unimodular networks in [BHMK18].

Appendices

Appendix A

Doeblin Trees

A.1. Postponed Proofs

In this appendix, all the notation of Chapter 4 is retained. Proofs that were only sketched in the main text are collected in full detail here.

Proof of Theorem 4.1.2. For each $t \in \mathbb{Z}$, let μ_t be the distribution of X_t. It is enough to show the existence of $(\Omega', \mathcal{F}', \mathbf{P}')$ on which there is a process $X' := (X'_t)_{t\in\mathbb{Z}}$ and some i.i.d. $\xi' := (\xi'_t)_{t\in\mathbb{Z}}$ such that

(i) $X'_t \sim \mu_t$ for all $t \in \mathbb{Z}$,

(ii) $\xi'_t \sim \xi_t$ for all $t \in \mathbb{Z}$,

(iii) X'_t is independent of $(\xi'_s)_{s\geqslant t}$ for all $t \in \mathbb{Z}$, and

(iv) $X'_{t+1} = h_{\text{gen}}(X_t, \xi'_t)$ for all $t \in \mathbb{Z}$.

Items (i) to (iv) and Lemma 4.1.1 will imply the result. Note that items (i) to (iv) are sufficient to characterize the joint finite dimensional distributions of $(X'_t)_{t\in\mathbb{Z}}$ and $(\xi'_t)_{t\in\mathbb{Z}}$. To see this fix $t_0 \leqslant t_1$. The joint distribution of $(X'_t)_{t_0\leqslant t\leqslant t_1}$ and $(\xi'_t)_{t_0\leqslant t\leqslant t_1}$ is determined because, conditional on $(\xi'_t)_{t_0\leqslant t\leqslant t_1}$, X'_{t_0} is still distributed as μ_{t_0} by items (i) and (iii), and, conditional on both

X'_{t_0} and $(\xi'_t)_{t_0 \leqslant t \leqslant t_1}$, one has that $(X'_t)_{t_0 \leqslant t \leqslant t_1}$ is deterministic by item (iv). Thus it suffices to show that $(X'_t)_{t \in \mathbb{Z}}$ and $(\xi'_t)_{t \in \mathbb{Z}}$ satisfying items (i) to (iv) exist. Also note that items (i) to (iv) with $t_0 \leqslant t \leqslant t_1$ are sufficient for determining the joint distribution of $(X'_s)_{t_0 \leqslant s \leqslant t_1}$ and $(\xi'_s)_{t_0 \leqslant s \leqslant t_1}$.

The proof will proceed by the Kolmogorov extension theorem. Suppose, by extending $(\Omega, \mathcal{F}, \mathbf{P})$ if necessary, that $(X_t)_{t \in \mathbb{Z}}$ and ξ are defined on the same space and are independent of each other. Consider for each $t \in \mathbb{Z}$, the state path $F^{(t, X_t)}$ in $\mathbf{G}$ started at (t, X_t). Then for all $s, t \in \mathbb{Z}$ with $s \leqslant t$ and all $x \in S$,

$$\begin{aligned} \mathbf{P}(F_t^{(s,X_s)} = x) &= \sum_{y \in S} \mathbf{P}(X_s = y, F_t^{(s,X_s)} = x) \\ &= \sum_{y \in S} \mu_s(y) P^{t-s}(y, x) \\ &= \mu_s P^{t-s}(x), \end{aligned}$$

where in the previous line P is treated as a transition kernel with powers P^k ($k = 0, 1, 2, \ldots$). Since $(X_t)_{t \in \mathbb{Z}}$ exists and is a Markov chain with transition matrix P, one has

$$\mu_s P^{t-s} = \mu_r P^{s-r} P^{t-s} = \mu_r P^{t-r} = \mu_t \tag{A.1}$$

for all $r \leqslant s \leqslant t$. Moreover, for all $s \leqslant t$, $F_t^{(s,X_s)}$ is $\sigma(X_s, (\xi_{t'})_{s \leqslant t' < t})$-measurable, hence it is independent of $(\xi_{t'})_{t' \geqslant t}$. Now fix $s_0, t_0, t_1 \in \mathbb{Z}$ with $s_0 \leqslant t_0 \leqslant t_1$ and consider the joint distribution of $(F_t^{(s_0, X_{s_0})})_{t_0 \leqslant t \leqslant t_1}$ and $(\xi_t)_{t_0 \leqslant t \leqslant t_1}$. One has $F^{(s_0, X_{s_0})}$ and ξ satisfy

(i′) $F_t^{(s_0, X_{s_0})} \sim \mu_t$ for all $t_0 \leqslant t \leqslant t_1$,

(ii′) $\xi \sim \xi$,

(iii′) $F_t^{(s_0, X_{s_0})}$ is independent of $(\xi_{t'})_{t' \geqslant t}$ for all $t_0 \leqslant t \leqslant t_1$, and

(iv′) $F_{t+1}^{(s_0, X_{s_0})} = h_{\text{gen}}(F_t^{(s_0, X_{s_0})}, \xi_t)$ for all $t_0 \leqslant t$.

As mentioned before, items (i′) to (iv′) are sufficient to determine the joint distribution of $(F_t^{(s_0, X_{s_0})})_{t_0 \leqslant t \leqslant t_1}$ and $(\xi_t)_{t_0 \leqslant t \leqslant t_1}$, so the joint distribution of $(F_t^{(s_0, X_{s_0})})_{t_0 \leqslant t \leqslant t_1}$ and $(\xi_t)_{t_0 \leqslant t \leqslant t_1}$ does not depend on s_0 as long as $s_0 \leqslant t_0$. Thus a consistent set of finite dimensional distributions is determined by taking $s_0, t_0 \to -\infty$ and $t_1 \to \infty$ while maintaining $s_0 \leqslant t_0 \leqslant t_1$. It follows by the Kolmogorov extension theorem that there is a space $(\Omega', \mathcal{F}', \mathbf{P}')$ and processes $X' = (X'_t)_{t \in \mathbb{Z}}$ and $\xi' = (\xi'_t)_{t \in \mathbb{Z}}$ satisfying items (i) to (iv), completing the proof. □

Call P **strongly recurrent** if all its recurrent classes are positive recurrent and call P **recurrent-attracting** if any Markov chain with transition matrix P eventually enters a recurrent state. These conditions are both automatic if P is irreducible and positive recurrent.

Proposition A.1.1 (Subsumes Proposition 4.1.3)**.** *Let $S = T \cup \bigcup (R^i)_{0 \leqslant i < N}$ decompose S into its transient states and $N \in \mathbb{N} \cup \{\infty\}$ recurrent communication classes for P. Assume that P is strongly recurrent and recurrent-attracting. Let $d(i)$ be the period of R^i, and let $R^i = C_0^i \cup \cdots \cup C_{d(i)-1}^i$ be a cyclic decomposition.*

If $\mathbf{G}$ *has fully independent transitions, then the components* $(C_j^i)_{0\leqslant i<N,0\leqslant j<d(i)}$ *of* $\mathbf{G}$ *are in bijection with* $(C_j^i)_{0\leqslant i<N,0\leqslant j<d(i)}$*, and for* $x \in C_j^i$*,* $(t,x) \in V(C_{j'}^i)$ *if and only if* $j - t = j' \pmod{d(i)}$*, and for* $x \in T$*,* $(t,x) \in V(C_j^i)$ *where* $(t',y) \in V(C_j^i)$ *is any vertex on the path of* (t,x) *for which* y *is recurrent. That is,* C_j^i *is the set of all vertices of all paths in* $\mathbf{G}$ *that pass through an element of* C_j^i *at any time* $t = 0 \pmod{d(i)}$*.*

Proof. Fix i, j, t and let $x, y \in C_j^i$. Then $P^{d(i)}$ restricted to C_j^i is irreducible, aperiodic, and positive recurrent. Thus the product chain $P^{d(i)} \otimes P^{d(i)}$ restricted to $C_j^i \times C_j^i$ is too. Strictly before the hitting time to the diagonal, $(F_{t+sd(i)}^{(t,x)}, F_{t+sd(i)}^{(t,y)})_{s\geqslant 0}$ is distributed the same as the product chain $P^{d(i)} \otimes P^{d(i)}$ on $C_j^i \times C_j^i$, and thus the hitting time to the diagonal is a.s. finite because the product chain is irreducible, aperiodic, and positive recurrent. It follows that (t,x) and of (t,y) are in the same component of $\mathbf{G}$. If $x \in C_j^i$ and $y \in C_{j'}^{i'}$ with $i' \neq i$, then $F^{(t,x)}$ and $F^{(t,y)}$ cannot merge because the states of $F^{(t,x)}$ are contained in R^i and the states of $F^{(t,y)}$ is contained in $R^{i'}$. If $x \in C_j^i$ and $y \in C_{j'}^i$ with $j' \neq j \pmod{d(i)}$, then $F^{(t,x)}$ and $F^{(t,y)}$ cannot merge because $F_{t+s}^{(t,x)} \in C_{j+s}^i$ but $F_{t+s}^{(t,y)} \in C_{j'+s}^i$ with indices taken modulo $d(i)$ as necessary. Thus, the set of $y \in S \setminus T$ such that $F^{(t,x)}$ eventually merges with $F^{(t,y)}$ is precisely C_j^i. If $x \in C_j^i$, $y \in C_{j'}^{i'}$ and $t \leqslant t'$, then $F_{t'}^{(t,x)} \in C_{j+(t'-t)}^i$, so it follows that $F^{(t,x)}$ and $F^{(t',y)}$ eventually merge if and only if $i' = i$ and $j + (t' - t) = j' \pmod{d(i)}$, or equivalently $j' - t' = j - t \pmod{d(i)}$. It follows that for any $x, y \in S \setminus T$ and any $t, t' \in \mathbb{Z}$, the two vertices $(t,x), (t',y) \in V(\mathbf{G})$ are in the same component of $\mathbf{G}$ if and only if there are i, j, j' such that $x \in C_j^i$, $y \in C_{j'}^i$,

and $j' - t' = j - t \pmod{d(i)}$. If $x \in C_j^i$, then $F_s^{(t,x)} \in C_{j-t+s}^i$ for all $s \geqslant t$. Thus $F_s^{(t,x)} \in C_{j-t}^i$ for all $s = 0 \pmod{d(i)}$ with $s \geqslant t$. Call C_{j-t}^i the **time-zero class of** (t, x). Then for $x \in C_j^i$, $y \in C_{j'}^i$ and $t, t' \in \mathbb{Z}$, the condition that $j' - t' = j - t \pmod{d(i)}$ is equivalent to the fact that (t, x) and (t, y) have the same time-zero class. Thus the components of $\mathbf{G}$ are exactly the equivalence classes of vertices in the same time-zero class, except possibly ignoring (t, x) for transient x. By the assumption that P is recurrent-attracting, if $x \in T$, then $F^{(t,x)}$ eventually hits some recurrent class and so does not form a new component of $\mathbf{G}$, and the path $F^{(t,x)}$ is in the component of the first (and every) (t', y) it hits with y recurrent. □

Proof of Lemma 4.1.14. Fix $k \in \mathbb{N}$. For every $v \in V = \mathbb{Z} \times S$, the event that $v \in V(\mathbf{\Gamma})$ and $d_{\mathbf{\Gamma}}(\boldsymbol{o}, v) \leqslant k$ is measurable. Indeed, there are at most countably many paths $(v_0, v_1, v_2, \ldots, v_n)$ in V with $n \leqslant k$, and the desired event is the union over all such paths of any length $n \leqslant k$ ending at v of the event

$$\{\boldsymbol{o} = v_0\} \cap \bigcap_{i=1}^{n} \left(\left\{ f_V(v_i) = 1 \right\} \cap \left\{ f_E(v_{i-1}, v_i) = 1 \right\} \right).$$

From here one sees that event that the r-neighborhood around $\boldsymbol{o}$ is exactly some fixed finite graph Γ is measurable. Indeed,

$$\{N_{\mathbf{\Gamma}}(\boldsymbol{o}, r) = \Gamma\} = \bigcap_{v \in V} \left\{ (f_V(v) = 1 \text{ and } d_{\mathbf{\Gamma}}(\boldsymbol{o}, v) \leqslant r) \iff v \in V(\Gamma) \right\}.$$

Enhancing Γ with marks $\xi_u, \xi_{v,w}$ for each $u \in V(\Gamma)$ and all $\{v, w\} \in E(\Gamma)$,

for any $\epsilon > 0$ and $o \in V(\Gamma)$, one sees that the event

$$\begin{aligned} D_{r,\epsilon}(\Gamma, o) := \{ & \boldsymbol{o} = o, \\ & N_{\boldsymbol{\Gamma}}(\boldsymbol{o}, r) = \Gamma, \\ & \forall u \in V(\Gamma), d_{\Xi_{\text{univ}}}(\xi_V(u), \xi_u) < \epsilon, \\ & \forall \{v, w\} \in E(\Gamma), d_{\Xi_{\text{univ}}}(\xi_E(v, w), \xi_{v,w}) < \epsilon \} \end{aligned}$$

is measurable. Since V is countable and Γ is a finite graph, there are at most countably many rooted isomorphic copies of (Γ, o) that can be made with vertices in V. It follows that the event $\left\{ d_{\mathcal{G}_*}([\boldsymbol{\Gamma}, \boldsymbol{o}], [\Gamma, o]) < \epsilon \right\}$ is a countable union of the events $D_{\lceil \frac{1}{\epsilon} \rceil, \epsilon}(\rho(\Gamma, o))$ with ρ ranging over the countable collection of such rooted isomorphisms of (Γ, o). Hence $\omega \mapsto [\boldsymbol{\Gamma}(\omega), \boldsymbol{o}(\omega)]$ is measurable. □

Proof of Proposition 4.1.15. For Item (a), suppose $[\mathbf{G}, (0, X_0)]$ is unimodular. Let η be a vertex-shift that follows the arrows in $\mathbf{G}$. For example, define for each network Γ and $u \in V(\Gamma)$ the vertex-shift by $\eta_\Gamma(u) := v$ if there is a unique outgoing edge from u and this edge terminates at v, or $\eta_\Gamma(u) := u$ if this condition is not met for any v. Since $\mathbf{G}$ is connected, its η-foils are $(\mathbf{G}_t)_{t \in \mathbb{Z}}$. Let the mark of a vertex v be denoted $(s(v), \xi(v))$, and let $v \sim w$ denote that v and w are in the same η-foil. Fix $x, y \in S$ and let $g[\Gamma, v, w] := 1_{\{s(v)=x, s(w)=y, v \sim w\}}$. Then the mass-transport principle implies

$$\mathbf{P}(X_0 = x) = \mathbf{E} \sum_{v \in V(\mathbf{G})} g[\mathbf{G}, (0, X_0), v] = \mathbf{E} \sum_{v \in V(\mathbf{G})} g[\mathbf{G}, v, (0, X_0)] = \mathbf{P}(X_0 = y),$$

so X_0 is uniformly distributed on S.

For Item (b), first suppose that X_0 is independent of $\mathbf{G}$ and uniformly distributed on a finite S. Let N be the cardinality of S. Let $g : \mathcal{G}_{**} \to \mathbb{R}_{\geqslant 0}$ be given. Then

$$\begin{aligned}
\mathbf{E} \sum_{v \in V(\mathbf{G})} g[\mathbf{G}, (0, X_0), v] &= \frac{1}{N} \sum_{x \in S} \mathbf{E} \sum_{(t,y) \in \mathbb{Z} \times S} g[\mathbf{G}, (0, x), (t, y)] \\
&= \frac{1}{N} \sum_{t \in \mathbb{Z}} \sum_{x,y \in S} \mathbf{E}[g[\mathbf{G}, (0, x), (t, y)]] \\
&= \frac{1}{N} \sum_{t \in \mathbb{Z}} \sum_{x,y \in S} \mathbf{E}[g[\mathbf{G}, (-t, x), (0, y)]] \\
&= \frac{1}{N} \sum_{y \in S} \mathbf{E} \sum_{(t,x) \in \mathbb{Z} \times S} g[\mathbf{G}, (-t, x), (0, y)] \\
&= \mathbf{E} \sum_{v \in V(\mathbf{G})} g[\mathbf{G}, v, (0, X_0)],
\end{aligned}$$

where in the third equality time-homogeneity of $\mathbf{G}$ is used. It follows that in this case $\mathbf{G}$ is unimodular.

For Item (c), let X_0 be the output of the CFTP algorithm in the standard CFTP setup. Suppose $[\mathbf{G}, (0, X_0)]$ is unimodular. Since $(0, X_0)$ has one outgoing edge in $\mathbf{G}$, unimodularity implies that on average it has one incoming edge. But, being the output of the CFTP algorithm, $(0, X_0)$ a.s. has at least one incoming edge. Hence $(0, X_0)$ a.s. has exactly one incoming edge. By unimodularity, it follows that a.s. every vertex in $\mathbf{G}$ has exactly one incoming edge. Since $\mathbf{G}$ is a tree, this is only possibly if S has a single element. If S has only a single element unimodularity is immediate. □

A.2. List of Mass-transports

As mentioned in Section 4.1.5, the proof style of Proposition 4.1.13 can be used to prove many equalities and inequalities in mean. A list is provided giving mass-transports, followed by the results they give after applying the boilerplate proof style with these mass-transports. Drawing a picture for each transport helps significantly in computing w^+ and w^- for the given transports. In all of the following, β is the union of all bi-recurrent paths in $\mathbf{B}$.

(i) Send mass 1 from each s to all times t strictly after s and strictly before $F^{(s,x^*)}$ returns to x^*.

- $\mathbf{E}[\#\mathbf{B}_0] \leqslant \mathbf{E}[\sigma^{(0,x^*)}(x^*)]$, where $\sigma^{(0,x^*)}(x^*)$ is the time until return of $F^{(0,x^*)}$ to x^*.

(ii) Fix $y \in S$. For each s, if $y \in \mathbf{B}_s$, send mass 1 to the first time $t > s$ that $F^{(s,y)}$ hits x^*.

- $\mathbf{P}(y \in \mathbf{B}_0) = \mathbf{E}[\#\mathbf{R}(0)^y]$, where $\mathbf{R}(0) \subseteq \mathbf{B}$ is the subgraph of vertices that first return to x^* at time 0, i.e., the (possibly empty) subgraph of $\mathbf{B}$ of all $(t, y) \in V(\mathbf{B})$ such that $\tau^{(t,y)}(x^*) = 0$, where $\tau^{(t,y)}(x^*)$ is the return time of $F^{(t,y)}$ to x^*.

- Summing over $y \in S$, one finds $\mathbf{E}[\#V(\mathbf{R}(0))] = \mathbf{E}[\#\mathbf{B}_0]$.

(iii) Send mass 1 from each s to the first time $t > s$ that $F_t^{(s,x^*)} = F_t^{(s',x^*)}$ for some $s' > t$.

- $\mathbf{E}[C(0)] = 1$, where $C(0)$ is the total number of paths $F^{(s,x^*)}$ that merge with a younger $F^{(s',x^*)}$ (i.e. with $s' > s$) for the first time at time 0.

- $\mathbf{P}(\#\mathbf{B}_1 \leqslant \#\mathbf{B}_0 - k) \leqslant \mathbf{P}(C(1) \geqslant k+1) \leqslant \frac{1}{k+1}$ for all $k \in \mathbb{N}$.

(iv) Fix $y \in S$. For each s, send mass 1 to each time t that $F_t^{(s,x^*)} = y$ and t is strictly before $F^{(s,x^*)}$ merges with the unique bi-recurrent path in its component of $\mathbf{B}$.

- $\mathbf{E}[N_0^{(0,x^*)}(y;\beta)] = \mathbf{E}[1_{\{y \in \mathbf{B}_0 \setminus \beta_0\}} \# V^{x^*}(D^{(0,y)} \cap V(\mathbf{B}))]$, where $N_0^{(0,x^*)}(y;\beta)$ denotes the number of visits (potentially 0) of $F^{(0,x^*)}$ to y strictly before merging with β.

- Summing over $y \in S$, $\mathbf{E}[\sigma_0^{(0,x^*)}(\beta)] = \mathbf{E}[\# V^{x^*}(D^{V_0(\mathbf{B}) \setminus V_0(\beta)} \cap V(\mathbf{B}))]$, where $\sigma_0^{(0,x^*)}(\beta)$ is the number of steps (potentially 0) before $F^{(0,x^*)}$ merges with β, and $D^{V_0(\mathbf{B}) \setminus V_0(\beta)}$ is the set of all descendants of all $v \in V_0(\mathbf{B}) \setminus V_0(\beta)$.

(v) Fix $y \in S$. For each s, if $y \in \mathbf{B}_s$ send mass 1 to the first time t that $F^{(s,y)}$ is on the bi-recurrent path in its component of $\mathbf{B}$.

- $\mathbf{P}(y \in \mathbf{B}_0) = \mathbf{E}[\# V^y(D^{V_0(\beta),M} \cap V(\mathbf{B}))]$, where $D^{V_0(\beta),M}$ denotes the union of $V_0(\beta)$ with their **mortal descendants**, i.e. those descendants with only finitely many descendants and whose first ancestor in β is at time 0.

- Summing over $y \in S$, one finds $\mathbf{E}[\#\mathbf{B}_0] = \mathbf{E}[\#(D^{V_0(\beta),M} \cap V(\mathbf{B}))]$.

(vi) Fix $y \in S$ and suppose $\mathbf{G}$ is an EFT and $(\beta_t)_{t\in\mathbb{Z}}$ is the bi-recurrent path in $\mathbf{G}$. For each t, if $\beta_t = y$ send mass 1 backwards to the most recent time $s < t$ such that $\beta_s = x^*$.

- $\mathbf{E}[N^{(0,x^*)}(y;x^*)1_{\beta_0=x^*}] = \mathbf{P}(\beta_0 = y)$, where $N^{(0,x^*)}(y;x^*)$ denotes the number of visits of $F^{(0,x^*)}$ to y before returning to x^*, including the initial visit if $y = x^*$.
- Summing over $y \in S$, one finds $\mathbf{E}[\sigma^{(0,x^*)}(x^*)1_{\{\beta_0=x^*\}}] = 1$.
- If $\mathbf{G}$ is also Markovian, then the previous points reduce to the classical cycle formulas, $\mathbf{E}[N^{(0,x^*)}(y;x^*)]\pi(x^*) = \pi(y)$ and $\mathbf{E}[\sigma^{(0,x^*)}(x^*)]\pi(x) = 1$, where π is the stationary distribution of the Markov chain.

Instead of using the unimodularity of $\mathbb{Z}$ and specifying a mass-transport $w = w(s,t)$ for $s,t \in \mathbb{Z}$, one may also use the unimodular version of $\mathbf{B}$ (that is, the random network with distribution $\mathbf{P}^{\square}$) and specify a mass-transport $w = w[\Gamma,u,v]$ for all networks Γ and all $u,v \in V(\Gamma)$. Some mass-transports are much easier to write in this way. For example, the mass-transport in item (iii) above also follows from the mass-transport $w[\Gamma,u,v] = 1$ if v is the unique out-neighbor of u in Γ. However, strictly speaking, there are no results using a mass-transport on $\mathbf{B}$ that could not also be proved with a mass-transport on $\mathbb{Z}$. Indeed, if w is a mass-transport defined for all networks Γ, then with $[\overline{\mathbf{B}},\square]$ denoting the identity mapping

under $\mathbf{P}^{\square}$,

$$\mathbf{E}^{\square}\left[\sum_{v\in V(\mathbf{B})} w[\overline{\mathbf{B}},\square,v]\right] = \mathbf{E}^{\square}\left[\sum_{v\in V(\mathbf{B})} w[\overline{\mathbf{B}},v,\square]\right]$$

may be rewritten as

$$\frac{1}{\mathbf{E}[\#\mathbf{B}_0]}\mathbf{E}\left[\sum_{t\in\mathbb{Z}} \hat{w}(0,t)\right] = \frac{1}{\mathbf{E}[\#\mathbf{B}_0]}\mathbf{E}\left[\sum_{t\in\mathbb{Z}} \hat{w}(t,0)\right]$$

where

$$\hat{w}(s,t) := \sum_{u\in V_s(\mathbf{B})}\sum_{v\in V_t(\mathbf{B})} w[\overline{\mathbf{B}},u,v], \qquad s,t\in\mathbb{Z}$$

is a mass-transport on $\mathbb{Z}$. That being said, the reader is encouraged the ponder the sequence of mass-transports on $\mathbb{Z}$ that would be required to prove a result like the classification theorem, Theorem 2.5.1, for the network $[\overline{\mathbf{B}},\square]$ directly. It seems more elegant to call upon the machinery of unimodular networks when convenient instead.

Appendix B

Point-shifts of Point Processes

B.1. Palm Calculus

In this appendix, fix a flow-adapted point process Ψ of intensity $\gamma \in (0, \infty)$. The necessity of this appendix is mostly to prove Proposition 5.1.1 and show how it may be used to translate definitions under $\mathbf{P}$ and $\mathbf{P}^{\Psi}$, a technique that is used extensively is this research.

The connection between $\mathbf{P}$ and $\mathbf{P}^{\Psi}$ is given by the refined Campbell theorem, abbreviated to C-L-M-M for Campbell, Little, Mecke, and Matthes.

Theorem B.1.1 (C-L-M-M)**.** *[Las10a] For all $\zeta : \Omega \times G \to \mathbb{R}_{\geqslant 0}$ measurable,*

$$\mathbf{E} \int_G \zeta(\theta_x^{-1}, x)\, \Psi(dx) = \gamma \mathbf{E}^{\Psi} \int_G \zeta(\theta_e, x)\, \lambda(dx).$$

It is possible to recover $\mathbf{P}$, up to the set on which Ψ is the zero measure, via the following inversion formula. The zero measure on G is denoted $\mathbf{0}$.

Theorem B.1.2 (Inversion Formula)**.** *[Las10a] There exists a bounded measurable $K : \Omega \times G \to \mathbb{R}_{\geqslant 0}$ such that*

$$\int_G K(\theta_e, x)\, \Psi(dx) = 1_{\{\Psi \neq \mathbf{0}\}}, \tag{B.1}$$

and for all $K : \Omega \times G \to \mathbb{R}_{\geqslant 0}$ *(not necessarily bounded)* $\mathbf{P}$*-a.s. satisfying* (B.1)*, it holds that*

$$\mathbf{E}[1_{\{\Psi \neq \mathbf{0}\}}\zeta] = \gamma \mathbf{E}^{\Psi} \int_G \zeta(\theta_x) K(\theta_x, x)\, \lambda(dx) \tag{B.2}$$

for all measurable $\zeta : \Omega \to \mathbb{R}_{\geqslant 0}$*.*

Proposition B.1.3. *If* $A \in \mathcal{F}$ *is* ***shift-invariant*** *in the sense that* $A = \theta_x^{-1} A$ *for all* $x \in G$*, then*

$$\mathbf{P}(A) = 1 \implies \mathbf{P}^{\Psi}(A) = 1 \implies \mathbf{P}(A \mid \Psi \neq \mathbf{0}) = 1.$$

In particular, if $\{\Psi = \mathbf{0}\} \subseteq A$ *then*

$$\mathbf{P}(A) = 1 \iff \mathbf{P}^{\Psi}(A) = 1.$$

Proof. Suppose $\mathbf{P}(A) = 1$. From the definition of Palm probabilities, for $B \in \mathcal{B}(G)$ such that $\lambda(B) \in (0, \infty)$,

$$\begin{aligned}
\mathbf{P}^{\Psi}(A) &= \frac{1}{\gamma \lambda(B)} \mathbf{E} \int_G 1_{\{x \in B\}} 1_{\{\theta_x^{-1} \in A\}}\, \Psi(dx) \\
&= \frac{1}{\gamma \lambda(B)} \mathbf{E} \int_G 1_{\{x \in B\}} 1_A\, \Psi(dx) && \text{(shift-invariance of } A\text{)} \\
&= \frac{1}{\gamma \lambda(B)} \mathbf{E}[1_A \Psi(B)] \\
&= \frac{1}{\gamma \lambda(B)} \mathbf{E}[\Psi(B)] && (\mathbf{P}(A) = 1) \\
&= 1.
\end{aligned}$$

Next suppose $\mathbf{P}^{\Psi}(A) = 1$. Then from Theorem B.1.2 there is measur-

able $K : \Omega \times G \to \mathbb{R}$ such that

$$\begin{aligned}
\mathbf{P}(A \cap \{\Psi \neq \mathbf{0}\}) &= \mathbf{E}[1_{\{\Psi \neq \mathbf{0}\}} 1_A] \\
&= \gamma \mathbf{E}^{\Psi} \int_G 1_{\{\theta_x \in A\}} K(\theta_x, x)\, \lambda(dx) && \text{(inversion formula)} \\
&= \gamma \mathbf{E}^{\Psi} \left[1_A \int_G K(\theta_x, x)\, \lambda(dx) \right] && \text{(shift-invariance of } A\text{)} \\
&= \gamma \mathbf{E}^{\Psi} \left[\int_G K(\theta_x, x)\, \lambda(dx) \right] && (\mathbf{P}^{\Psi}(A) = 1) \\
&= \mathbf{E}[1_{\{\Psi \neq \mathbf{0}\}} \cdot 1] && \text{(inversion formula)} \\
&= \mathbf{P}(\Psi \neq \mathbf{0}).
\end{aligned}$$

Dividing by $\mathbf{P}(\Psi \neq \mathbf{0}) > 0$ gives $\mathbf{P}(A \mid \Psi \neq \mathbf{0}) = 1$, and if $\{\Psi = \mathbf{0}\} \subseteq A$, then

$$\mathbf{P}(A) = \mathbf{P}(A \cap \{\Psi \neq \mathbf{0}\}) + \mathbf{P}(A \cap \{\Psi = \mathbf{0}\}) = \mathbf{P}(\Psi \neq \mathbf{0}) + \mathbf{P}(\Psi = \mathbf{0}) = 1. \quad \square$$

Lemma B.1.4. *Let $A \in \mathcal{F}$. Then*

$$\mathbf{P}^{\Psi}(A) = 1 \iff \mathbf{P}(\Psi(x \in G : \theta_x^{-1} \notin A) = 0) = 1.$$

Proof. By replacing A with its complement it is equivalent to show $\mathbf{P}^{\Psi}(A) = 0$ if and only if it holds that $\mathbf{P}(\Psi(x \in G : \theta_x^{-1} \in A) > 0) = 0$. Note that it is the joint measurability of the action $(\omega, x) \mapsto \theta_x \omega$ that lets one conclude for $B \in \mathcal{B}(G)$ that sets like

$$\{\Psi(x \in G : x \in B, \theta_x^{-1} \in A) > 0\}$$

are measurable.

If $\mathbf{P}^{\Psi}(A) = 0$, then for $B \in \mathcal{B}(G)$ such that $\lambda(B) \in (0, \infty)$,

$$\begin{aligned} 0 &= \mathbf{P}^{\Psi}(A) \\ &= \frac{1}{\gamma\lambda(B)} \mathbf{E} \int_G 1_{\{x \in B\}} 1_{\{\theta_x^{-1} \in A\}} \, \Psi(dx) \\ &= \frac{1}{\gamma\lambda(B)} \mathbf{E}[\Psi(x \in G : x \in B, \theta_x^{-1} \in A)]. \end{aligned}$$

Thus $\mathbf{E}[\Psi(x \in G : x \in B, \theta_x^{-1} \in A)] = 0$ and taking relatively compact B increasing to G one finds that $\mathbf{E}[\Psi(x \in G : \theta_x^{-1} \in A)] = 0$, so $\mathbf{P}(\Psi(x \in G : \theta_x^{-1} \in A) > 0) = 0$.

Conversely, suppose $\mathbf{P}(\Psi(x \in G : \theta_x^{-1} \in A) > 0) = 0$. Then for $B \in \mathcal{B}(G)$ with $\lambda(B) \in (0, \infty)$,

$$\begin{aligned} \mathbf{P}^{\Psi}(A) &= \frac{1}{\gamma\lambda(B)} \mathbf{E} \int_G 1_{\{x \in B\}} 1_{\{\theta_x^{-1} \in A\}} \, \Psi(dx) \\ &= \frac{1}{\gamma\lambda(B)} \mathbf{E}[\Psi(x \in G : x \in B, \theta_x^{-1} \in A)] \\ &\leqslant \frac{1}{\gamma\lambda(B)} \mathbf{E}[\Psi(x \in G : \theta_x^{-1} \in A)] \\ &= 0, \end{aligned}$$

completing the proof. □

It is now possible to prove Proposition 5.1.1.

Proof of Proposition 5.1.1.

(a) $\Longleftrightarrow$ *(b)*: This is the content of Lemma B.1.4.

(b) $\iff$ *(c)*: This follows from Proposition B.1.3 and the fact that the event $\{\Psi(x \in G : \theta_x^{-1} \notin A) = 0\}$ contains $\{\Psi = \mathbf{0}\}$ and is shift-invariant. To wit, for all $y \in G$,

$$\begin{aligned}
&\theta_y^{-1}\omega \in \{\Psi(x \in G : \theta_x^{-1} \notin A) = 0\} \\
\iff &\Psi(\theta_y^{-1}\omega, \{x \in G : \theta_x^{-1}\theta_y^{-1}\omega \notin A\}) = 0 \\
\iff &\Psi(\omega, \{yx : x \in G, \theta_{yx}^{-1}\omega \notin A\}) \\
\iff &\Psi(\omega, \{x \in G, \theta_x^{-1}\omega \notin A\}) = 0 \\
\iff &\omega \in \{\Psi(x \in G, \theta_x^{-1}\omega \notin A) = 0\}.
\end{aligned}$$

□

Example B.1.5. Fix some measurable space (S, Σ) and a measurable $\zeta : \Omega \to S$. Define $Z : \Omega \times G \to S$ by $Z(\omega, x) := \zeta(\theta_x^{-1}\omega)$ for all $\omega \in \Omega, X \in \Psi(\omega)$, and $Z(\omega, x)$ may be defined arbitrarily otherwise. It will be shown that knowing Z up to a $\mathbf{P}$- or $\mathbf{P}^\Psi$-null set on the support of Ψ is equivalent to knowing ζ up to a $\mathbf{P}^\Psi$-null set. Indeed, suppose $\zeta = \zeta'$, $\mathbf{P}^\Psi$-a.s., then it will be shown that the corresponding Z, Z' agree $\mathbf{P}, \mathbf{P}^\Psi$-a.s. on the support of Ψ. By Proposition 5.1.1, $\mathbf{P}$- and $\mathbf{P}^\Psi$-a.e. $\omega \in \Omega$ has for all $X \in \Psi(\omega)$ that $\zeta(\theta_X^{-1}\omega) = \zeta'(\theta_X^{-1}\omega)$, i.e. $Z(\omega, X) = Z'(\omega, X)$. Similarly, if either $\mathbf{P}$-a.e. or $\mathbf{P}^\Psi$-a.e. $\omega \in \Omega$ is such that $Z(\omega, X) = Z'(\omega, X)$ for all $X \in \Psi(\omega)$, then

$$\begin{aligned}
\zeta(\theta_X^{-1}\omega) &= Z(\omega, X) \\
&= Z'(\omega, X) \\
&= \zeta'(\theta_X^{-1}\omega),
\end{aligned}$$

for $\mathbf{P}$-a.e. or $\mathbf{P}^\Psi$-a.e. $\omega \in \Omega, X \in \Psi(\omega)$, so by Proposition 5.1.1 one finds that

$\zeta = \zeta'$, $\mathbf{P}^\Psi$-a.s. Thus, ζ may be defined under $\mathbf{P}^\Psi$ or Z may be defined under $\mathbf{P}$ or $\mathbf{P}^\Psi$, whichever is more convenient.

Finally, the following standard result is needed in Section 5.2.2.

Theorem B.1.6 (Slivnyak-Mecke Theorem)**.** *[DVJ08] The distribution of Ψ under $\mathbf{P}^\Psi$ is the same as the distribution of $\Psi + \delta_e$ under $\mathbf{P}$ if and only if Ψ is a homogeneous Poisson point process with intensity γ under $\mathbf{P}$.*

Bibliography

[AL07] D. Aldous and R. Lyons. Processes on unimodular random networks. *Electron. J. Probab.*, 12:no. 54, 1454–1508, 2007. doi:10.1214/EJP.v12-463.

[Arn10] Ö. Arnaldsson. On coupling of discrete random walks on the line. Master's thesis, Citeseer, 2010.

[Ber79] H. C. P. Berbee. *Random walks with stationary increments and renewal theory*, volume 112 of *Mathematical Centre Tracts*. Mathematisch Centrum, Amsterdam, 1979.

[BF92] A. A. Borovkov and S. G. Foss. Stochastically recursive sequences and their generalizations. *Siberian Adv. Math.*, 2(1):16–81, 1992. Siberian Advances in Mathematics.

[BF94] A. Borovkov and S. G. Foss. Two ergodicity criteria for stochastically recursive sequences. *Acta Appl. Math.*, 34(1-2):125–134, 1994. doi:10.1007/BF00994261.

[BHM18] F. Baccelli and M.-O. Haji-Mirsadeghi. Point-shift foliation of a point process. *Electron. J. Probab.*, 23:Paper No. 19, 25, 2018. doi:10.1214/17-EJP123.

[BHMK18] F. Baccelli, M.-O. Haji-Mirsadeghi, and A. Khezeli. Eternal family trees and dynamics on unimodular random graphs. *arXiv preprint arXiv:1608.05940*, 2018. To appear in the speccial issue of Contemporary Mathematics on Unimodularity in Randomly Generated Graphs.

[Bog07] V. I. Bogachev. *Measure theory. Vol. I, II*. Springer-Verlag, Berlin, 2007. doi:10.1007/978-3-540-34514-5.

[Bor13] A. A. Borovkov. *Probability theory*. Universitext. Springer, London, 2013. doi:10.1007/978-1-4471-5201-9. Translated from the 2009 Russian fifth edition by O. B. Borovkova and P. S. Ruzankin, Edited by K. A. Borovkov.

[BS18] F. Baccelli and A. Sodre. Renewal population dynamics and their eternal family trees. *arXiv preprint arXiv:1803.08081*, 2018.

[CK07] S. B. Connor and W. S. Kendall. Perfect simulation for a class of positive recurrent Markov chains. *Ann. Appl. Probab.*, 17(3):781–808, 2007. doi:10.1214/105051607000000032.

[Coh13] D. L. Cohn. *Measure theory*. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser/Springer, New York, second edition, 2013. doi:10.1007/978-1-4614-6956-8.

[Doe38] W. Doeblin. Exposé de la théorie des chaınes simples constantes de markov á un nombre fini d'états. *Mathématique de l'Union Interbalkanique*, 2(77-105):78–80, 1938.

[DVJ08] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. II*. Probability and its Applications (New York). Springer, New York, second edition, 2008. doi:10.1007/978-0-387-49835-5. General theory and structure.

[Fil99] J. A. Fill. An interruptible algorithm for perfect sampling via Markov chains. In *STOC '97 (El Paso, TX)*, pages 688–695. ACM, New York, 1999.

[FK03] S. Foss and T. Konstantopoulos. Extended renovation theory and limit theorems for stochastic ordered graphs. *Markov Process. Related Fields*, 9(3):413–468, 2003.

[FT98] S. G. Foss and R. L. Tweedie. Perfect simulation and backward coupling. *Comm. Statist. Stochastic Models*, 14(1-2):187–203, 1998. doi:10.1080/15326349808807466. Special issue in honor of Marcel F. Neuts.

[Her65] H. Herrmann. Glättungseigenschaften der Faltung. *Wiss. Z. Friedrich-Schiller-Univ. Jena/Thüringen*, 14:221–234, 1965.

[HL05] M. Heveling and G. Last. Characterization of Palm measures via bijective point-shifts. *Ann. Probab.*, 33(5):1698–1715, 2005.

doi:10.1214/009117905000000224.

[HN99] O. Häggström and K. Nelander. On exact simulation of Markov random fields using coupling from the past. *Scand. J. Statist.*, 26(3):395–411, 1999. doi:10.1111/1467-9469.00156.

[HR79] E. Hewitt and K. A. Ross. *Abstract harmonic analysis. Vol. I*, volume 115 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin-New York, second edition, 1979. Structure of topological groups, integration theory, group representations.

[HvLM99] O. Häggström, M.-C. N. M. van Lieshout, and J. Møller. Characterization results and Markov chain Monte Carlo algorithms including exact simulation for some spatial point processes. *Bernoulli*, 5(4):641–658, 1999. doi:10.2307/3318694.

[KÖ90] D. König. *Theory of finite and infinite graphs*. Birkhäuser Boston, Inc., Boston, MA, 1990. doi:10.1007/978-1-4684-8971-2. Translated from the German by Richard McCoart, With a commentary by W. T. Tutte and a biographical sketch by T. Gallai.

[Ken98] W. S. Kendall. Perfect simulation for the area-interaction point process. In *Probability towards 2000 (New York, 1995)*, volume 128 of *Lect. Notes Stat.*, pages 218–234. Springer, New York, 1998. doi:10.1007/978-1-4612-2224-8_13.

[Ken04] W. S. Kendall. Geometric ergodicity and perfect simulation. *Electron. Comm. Probab.*, 9:140–151, 2004. doi:10.1214/ECP.v9-1117.

[Khe17] A. Khezeli. Shift-coupling of random rooted graphs and networks. *arXiv preprint arXiv:1703.06619*, 2017.

[KM00] W. S. Kendall and J. Møller. Perfect simulation using dominating processes on ordered spaces, with application to locally stable point processes. *Adv. in Appl. Probab.*, 32(3):844–865, 2000. doi:10.1239/aap/1013540247.

[Las10a] G. Last. Modern random measures: Palm theory and related models. In *New perspectives in stochastic geometry*, pages 77–110. Oxford Univ. Press, Oxford, 2010.

[Las10b] G. Last. Stationary random measures on homogeneous spaces. *J. Theoret. Probab.*, 23(2):478–497, 2010. doi:10.1007/s10959-009-0231-9.

[Lin01] E. Lindenstrauss. Pointwise theorems for amenable groups. *Invent. Math.*, 146(2):259–295, 2001. doi:10.1007/s002220100162.

[Lin02] T. Lindvall. *Lectures on the coupling method*. Dover Publications, Inc., Mineola, NY, 2002. Corrected reprint of the 1992 original.

[LP17] D. A. Levin and Y. Peres. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2017. Second edition

of [MR2466937], With contributions by Elizabeth L. Wilmer, With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson.

[Mec67] J. Mecke. Stationäre zufällige Masse auf lokalkompakten Abelschen Gruppen. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 9:36–58, 1967. doi:10.1007/BF00535466.

[Men00] X.-L. Meng. Towards a more general Propp-Wilson algorithm: multistage backward coupling. In *Monte Carlo methods (Toronto, ON, 1998)*, volume 26 of *Fields Inst. Commun.*, pages 85–93. Amer. Math. Soc., Providence, RI, 2000.

[MG98] D. J. Murdoch and P. J. Green. Exact sampling from a continuous state space. *Scand. J. Statist.*, 25(3):483–502, 1998. doi:10.1111/1467-9469.00116.

[Møl99] J. Møller. Perfect simulation of conditionally specified models. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 61(1):251–264, 1999. doi:10.1111/1467-9868.00175.

[Møl01] J. Møller. A review of perfect simulation in stochastic geometry. In *Selected Proceedings of the Symposium on Inference for Stochastic Processes (Athens, GA, 2000)*, volume 37 of *IMS Lecture Notes Monogr. Ser.*, pages 333–355. Inst. Math. Statist., Beachwood, OH, 2001. doi:10.1214/lnms/1215090699.

[MT09] S. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second edition, 2009. doi:10.1017/CBO9780511626630. With a prologue by Peter W. Glynn.

[Mur17] J. T. Murphy. Exact coupling of random walks on polish groups. *Journal of Theoretical Probability*, pages 1–17, 2017.

[Orn69] D. S. Ornstein. Random walks. I, II. *Trans. Amer. Math. Soc. 138 (1969), 1-43; ibid.*, 138:45–60, 1969. doi:10.2307/1994897.

[PW96] J. G. Propp and D. B. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. In *Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995)*, volume 9, pages 223–252, 1996. doi:10.1002/(SICI)1098-2418(199608/09)9:1/2<223::AID-RSA14>3.3.CO;2-R.

[PW98a] J. Propp and D. Wilson. Coupling from the past: a user's guide. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, volume 41 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 181–192. Amer. Math. Soc., Providence, RI, 1998.

[PW98b] J. G. Propp and D. B. Wilson. How to get a perfectly random sample from a generic Markov chain and generate a random spanning tree of a directed graph. *J. Algorithms*, 27(2):170–217,

1998. doi:10.1006/jagm.1997.0917. 7th Annual ACM-SIAM Symposium on Discrete Algorithms (Atlanta, GA, 1996).

[Sta66] A. J. Stam. On shifting iterated convolutions. I. *Compositio Math.*, 17:268–280 (1966), 1966.

[Str72] K. Stromberg. An elementary proof of Steinhaus's theorem. *Proc. Amer. Math. Soc.*, 36:308, 1972. doi:10.2307/2039082.

[Tho00] H. Thorisson. *Coupling, stationarity, and regeneration*. Probability and its Applications (New York). Springer-Verlag, New York, 2000. doi:10.1007/978-1-4612-1236-2.

[Tho11] H. Thorisson. Open problems in renewal, coupling and Palm theory. *Queueing Syst.*, 68(3-4):313–319, 2011. doi:10.1007/s11134-011-9241-2.

[Wil00] D. B. Wilson. How to couple from the past using a read-once source of randomness. *Random Structures Algorithms*, 16(1):85–113, 2000. doi:10.1002/(SICI)1098-2418(200001)16:1<85::AID-RSA6>3.0.CO;2-H.