Doeblin Trees

François Baccelli[*], Mir-Omid Haji-Mirsadeghi[†], and
James T. Murphy III[‡]

November 25, 2018

Abstract

This paper is centered on the random graph generated by a Doeblin-type coupling of discrete time processes on a countable state space 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 at any time. The bridge graph is made into a unimodular network by marking it and selecting a root in a specified fashion. The unimodularity of this network is leveraged to discern global properties of the larger Doeblin graph. Bi-recurrence, i.e., recurrence both forwards and backwards in time, is introduced and shown to be a key property in uniquely distinguishing paths in the Doeblin graph, and also a decisive property for Markov chains indexed by $\mathbb{Z}$. Properties related to simulating the bridge graph are also studied.

Contents

[*]The University of Texas at Austin, `baccelli@math.utexas.edu`
[†]Sharif University of Technology, `mirsadeghi@sharif.edu`
[‡]The University of Texas at Austin, `james@intfxdx.com`

1.   Introduction

The first incarnation of the Propp and Wilson [24] coupling from the past (CFTP) algorithm was designed to build a perfect sample from the stationary distribution π of an irreducible, aperiodic, and positive recurrent 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.

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 [6, 4] laid out the framework of stochastically recursive sequences (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 [24] 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 [6] for SRS to the Markov case aiming at perfect simulation. Foss and Tweedie [10] then gave a necessary and sufficient condition for the CFTP algorithm to converge a.s. From 1996 to 2000, many papers [8, 23, 22, 25, 13, 11, 12, 20, 15, 26, 18] investigated how to improve CFTP implementations or how to apply CFTP or a CFTP-inspired algorithm to obtain a perfect sample from

2

a particular Markov chain's stationary distribution. Of particular importance is Wilson's read-once CFTP algorithm [26], 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 [21]. Since then, [14, 7] showed that (possibly impractical) generalizations of the CFTP algorithm can be applied under weaker conditions, and [9] gives a CFTP-like algorithm that applies even in the non-Markovian setting.

In this paper, 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 [1]. 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** and is roughly inspired by the population process in [3]. 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 2. Random networks and how to view subgraphs like the bridge graph as random networks are handled in Section 3. The main theorem is then proved in Section 4.

Section 4.1 proves the main theorem, identifying the unimodular structure in the bridge graph. Section 4.2 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 [2], which is reviewed in Section 3. The most interesting case is when S is infinite and the Doeblin graph is connected. In this case (see Corollary 4.7), 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 4.10 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

3

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 4.3 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. The final section gives several bibliographical comments, which make connections of the present research to other works.

2. The Doeblin Graph

2.1. 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 document. 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

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

Equation (1) is the defining property of a stochastically recursive sequence (SRS) in the sense of Borovkov and Foss [6]. 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 [5]). Many processes in this paper 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 document. The notation for the transition matrix $P = (p_{x,y})_{x,y \in S}$ is also fixed for the remainder of the document, 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

$$V_t(\Gamma) := \{(s, y) \in V(\Gamma) : s = t\}, \quad V^x(\Gamma) := \{(s, y) \in V(\Gamma) : y = x\}. \tag{2}$$

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

$$\Gamma_t := \{x \in S : (t, x) \in V_t(\Gamma)\}, \quad \Gamma^x := \{t \in \mathbb{Z} : (t, x) \in V^x(\Gamma)\}. \tag{3}$$

4

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

$$f_+(t, x) := (t + 1, h_{\text{gen}}(x, \xi_t)), \qquad (t, x) \in \mathbb{Z} \times S. \tag{4}$$

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 parent vertex.

2.2. Modeling

When dealing with paths in $\mathbf{G}$, 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 which takes values in S and is such that $(t, X_t)_{t \in I}$ is a.s. a path in $\mathbf{G}$ 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 $\mathbf{G}$: 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 2.1. *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 $\mathbf{G}$. Then $(X_t)_{t \in I}$ is a state path in $\mathbf{G}$.*

Proof. One must check that $(t, X_t)_{t \in I}$ is a.s. a path in $\mathbf{G}$. Fix $t \in I$. Since $V(\mathbf{G}) = \mathbb{Z} \times S$, (t, X_t) is certainly a vertex of $\mathbf{G}$. 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 $\mathbf{G}$. The edges of $\mathbf{G}$ 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 $\mathbf{G}$. $\qquad \square$

In particular, Lemma 2.1 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 document, let $F^{(t,x)} := (F_s^{(t,x)})_{s \geqslant t}$ 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

$$(s, F_s^{(t,x)}) = f_+^{s-t}(t, x), \qquad (t, x) \in \mathbb{Z} \times S, \quad s \geqslant t. \tag{5}$$

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
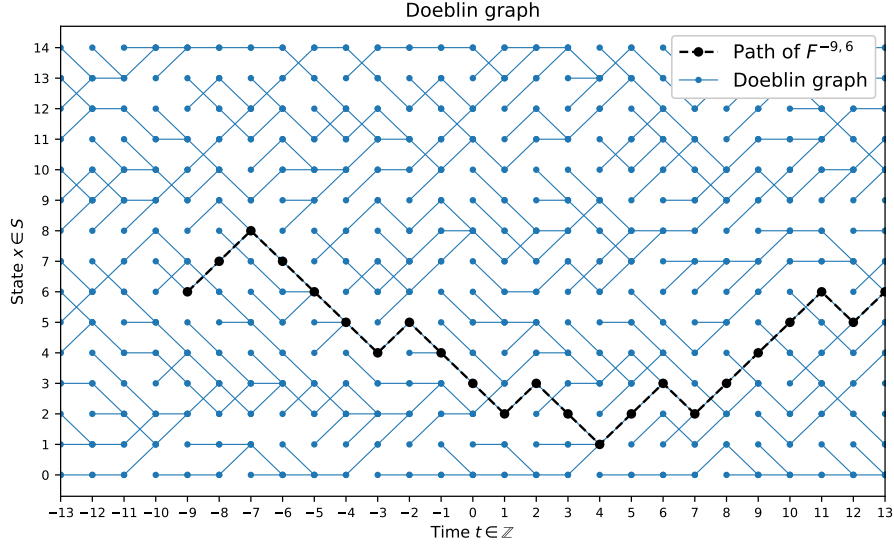
Figure 1: An example of a Doeblin graph with the path corresponding to the state path $F^{(t,x)}$ distinguished. All edges are directed from left to right.

paper, 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 1.

It has already been noted (see [5]) 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 2.2. *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 $\mathbf{G}$. 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 $\xi' = (\xi'_t)_{t \in \mathbb{Z}} \sim \xi$ 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 $\mathbf{G}$. Then consider for each $t \in \mathbb{Z}$ the state path in $\mathbf{G}$ 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 2.2 is given in the appendix.

6

2.3. 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 [2] and the word eternal refers to the fact that every vertex of **G** has a unique outgoing edge. That is, there is no individual that is an ancestor of all other individuals. 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 2.3. *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 **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 2.3 is given in the appendix.

A ξ-measurable subgraph $\Gamma = \Gamma((\xi_t)_{t \in \mathbb{Z}})$ of **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 **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 occurence 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: **G** is locally finite, **G** contains no cycles, **G** is connected, **G** has

7

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 $\mathbf{G}$ is a Markovian, one needs to be cautious that not all state paths in $\mathbf{G}$ are Markov chains with transition matrix P.

Example 2.4. Let $S := \mathbb{Z}$ and suppose $\mathbf{G}$ 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 2.4, 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 $\mathbf{G}$, proper state paths inherit a Markov transition structure.

Lemma 2.5. *Suppose $\mathbf{G}$ is Markovian. If $(X_t)_{t \in I}$ is a proper state path in $\mathbf{G}$ 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. $\qquad \square$

2.4. Connections with CFTP

Consider the following structural result that will be expanded upon in Section 4.2.1. It is a special case of Proposition 4.6 and Corollary 4.7, which will be proved later.

Proposition 2.6. *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 **G**, and in doing so one sees that is has much the same spirit as Proposition 2.6.

Proposition 2.7 (Perfect Sampling [24])**.** *If* S *is finite and* **G** *is Markovian and an EFT (which, since* S *is finite, necessitates that* P *is irreducible and aperiodic), then there is an a.s. finite time* τ *such that all paths in* **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 2.8. *In fact, the* β_0 *appearing in Proposition 2.7 and the* β_0 *appearing in Proposition 2.6 are the same. That is, the perfect sampling algorithm* A *is ultimately computing the point in* **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* **G** *must also pass through* $(0, \beta_0)$. *However, what is notably absent in Proposition 2.6 is any mention of an algorithm to compute* β_0. *Whether such an algorithm exists in general is not studied in the present research.*

2.5. Bridge Graphs

The primary tool used in this document will be the theory of unimodular networks in the sense of [1]. Local finiteness is essential in the theory of unimodular networks, but the Doeblin graph **G** may not be locally finite, as the following result shows.

Proposition 2.9. *If* $\sum_{x \in S} p_{x,y} < \infty$ *for all* $y \in S$, *then* **G** *is a.s. locally finite. If* **G** *has fully independent transitions and for some* $y \in S$, $\sum_{x \in S} p_{x,y} = \infty$, *then* **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 $\{f_+(t - 1, x) = (t, y)\}_{x \in S}$ occur, showing (t, y) has finite in-degree, and hence finite degree, in **G**. On the other hand, if **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 $\{f_+(t - 1, x) = (t, y)\}_{x \in S}$ 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

$$\tau^{(t,x)}(y) := \inf \left\{ s > t : F_s^{(t,x)} = y \right\}, \qquad \sigma^{(t,x)}(y) := \tau^{(t,x)}(y) - t \qquad (6)$$

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

$$V(\mathbf{B}(x)) := \bigcup_{t \in \mathbb{Z}} \left\{ (s, F_s^{(t,x)}) : s \geqslant t \right\} = \bigcup_{t \in \mathbb{Z}} \left\{ f_+^n(t, x) : n \geqslant 0 \right\}. \qquad (7)$$

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 2.10. *For the remainder of the document, 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 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 **descendants** of (t, y) in $\mathbf{G}$ to be

$$D^{(t,y)} := \left\{ (s, x) \in \mathbb{Z} \times S : F_t^{(s,x)} = y \right\}. \qquad (8)$$

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 (3),

$$y \in \mathbf{B}_t \iff \exists s, \ x^* \in D_s^{(t,y)}, \qquad (t, y) \in \mathbb{Z} \times S. \qquad (9)$$

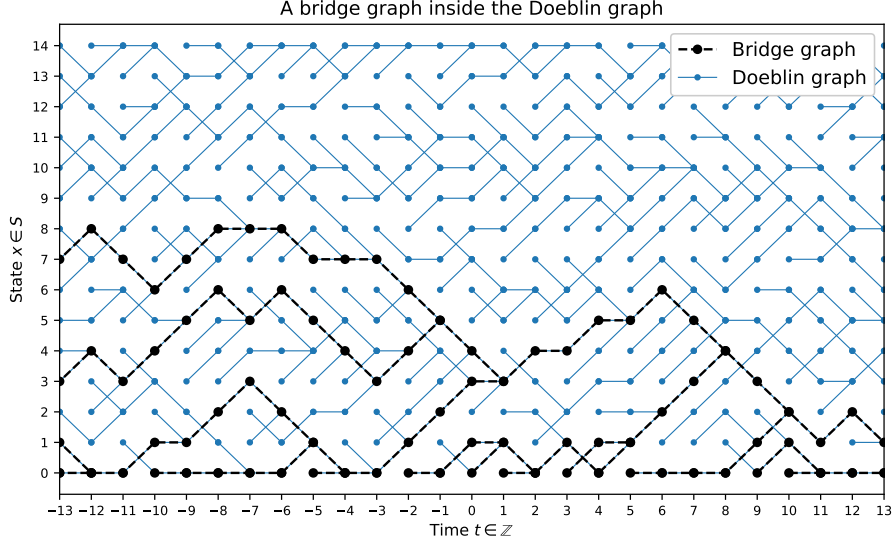Lemma 2.11 shows that if $\mathbf{G}$ is a.s. connected, then $\mathbf{B}$ is too.

Figure 2: An example bridge graph, in this case for state $x^* = 0$, sitting inside the Doeblin graph.

Lemma 2.11. *If $u, v \in V(\mathbf{B})$ are in the same component of $\mathbf{G}$, then they are in the same component of $\mathbf{B}$. In particular, if $\mathbf{G}$ is an EFT, then $\mathbf{B}$ is an EFT.*

Proof. Consider times $s, t \in \mathbb{Z}$. Suppose (s, x^*) and (t, x^*) are in the same component of $\mathbf{G}$. 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 $\mathbf{B}$. Hence (s, x^*) and (t, x^*) are in the same component of $\mathbf{B}$. Now if $u, v \in V(\mathbf{B})$ are in the same component of $\mathbf{G}$, u is in the same component in $\mathbf{G}$ as some (s, x^*) and v is in the same component of $\mathbf{G}$ as some (t, x^*), and (s, x^*) and (t, x^*) are in the same component of $\mathbf{B}$ by the previous part. Hence u, v are in the same component of $\mathbf{B}$. $\square$

The condition that $\mathbf{B}$ is an EFT is equivalent to strong coupling convergence (defined and studied in [6, 4, 9]) of $F^{(0, x^*)}$ to a stationary version of the SRS. However, simple conditions for $\mathbf{B}$ to be an EFT are not known outside of the Markovian case, where Proposition 2.3 showed that if P is irreducible, aperiodic, and positive recurrent, then $\mathbf{G}$ is an EFT. Another (not neccessarily easy to check) condition for $\mathbf{B}$ to be an EFT will be given in Corollary 4.8.

The main tool used in this paper 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 2.12 (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*

$$\mathbf{E}[w^+] = \mathbf{E}[w^-]. \tag{10}$$

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. $\qquad\square$

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

Proposition 2.13. *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 $w(s,t) := 1_{\left\{\sigma^{(s,x^*)}(x^*) > t-s > 0\right\}}$. 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. $\qquad\square$

The proof style of Proposition 2.13 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 2.13 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. $\qquad\square$

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 6.2 in the appendix. 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 reviews the more general theory of random networks and unimodularity, then shows how to embed subgraphs of **G** as random networks, so that eventually one may find a unimodular structure inside **G**.

3. Random Networks

3.1. Definition and Basic Properties

See [1, 16] 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 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 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 $\mathbb{N}^{\mathbb{N}}$ or the Hilbert cube, since all possible mark spaces are homeomorphic to a subset of such a Ξ_{univ}. 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. [1]). 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 balls 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.

A **random (rooted) network** is a random element in $\mathcal{G}_*$ equipped with its Borel σ-algebra $\mathcal{B}(\mathcal{G}_*)$. A random network $[\mathbf{\Gamma}, \mathbf{o}]$ is called **unimodular** if for all measurable $g : \mathcal{G}_{**} \to \mathbb{R}_{\geqslant 0}$, the following **mass transport principle** is

13

satisfied:

$$\mathbf{E} \sum_{v \in V(\mathbf{\Gamma})} g[\mathbf{\Gamma}, \boldsymbol{o}, v] = \mathbf{E} \sum_{v \in V(\mathbf{\Gamma})} g[\mathbf{\Gamma}, v, \boldsymbol{o}]. \tag{11}$$

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 (11) 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 [2] 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 [2].

Theorem 3.1 (Foil Classification in Unimodular Networks [2]). *Let $[\mathbf{\Gamma}, \boldsymbol{o}]$ be a unimodular network and Φ a vertex-shift. Almost surely, every vertex has finite degree in the Φ-graph of $\mathbf{\Gamma}$. In addition, each component C of the Φ-graph of $\mathbf{\Gamma}$ 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 [2] is the so-called no infinite/finite inclusion lemma, which is used heavily in the proof of Theorem 3.1. To state it, the following definitions are needed. A **covariant subset (of the set of vertices)** is a map C which 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 $\{[G, o, v] : v \in \Pi_G(o)\} \subseteq \mathcal{G}_{**}$ is measurable, where $\Pi_\Gamma(o)$ denotes the partition element in Π_Γ containing o. Then one has the following.

14

Lemma 3.2 (No Infinite/Finite Inclusion [2]). *Let $[\mathbf{\Gamma}, \mathbf{o}]$ 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 $E \cap C_{\mathbf{G}}$ is finite and nonempty.*

3.2. 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 $\mathbf{o}$ 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.

Let $V := \mathbb{Z} \times S$. Suppose that

$$\mathbf{\Gamma} : \Omega \to \{0,1\}^V \times \Xi^V \times \{0,1\}^{V \times V} \times \Xi^{V \times V} =: (f_V, \xi_V, f_E, \xi_E) \qquad (12)$$

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) $f_E((t, x), (t+1, h(x, \xi_t))) = f_E((t+1, h(x, \xi_t)), (t, x)) = 1$ 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 3.3. *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 3.3 given in the appendix.

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 3.4. *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*

- *if* $[\mathbf{G}, (0, X_0)]$ *is unimodular, then* S *is finite and* X_0 *is uniformly distributed on* S,

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

- *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

16

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 3.4 is given in the appendix.

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 3.5. *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 2.13 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 3.3 finishes the claim. $\square$

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

4. Unimodularizability and its Consequences

4.1. Unimodularizability of the Bridge Graph

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

Theorem 4.1. *Any random network with distribution*

$$\mathbf{P}^{\square}(A) := \frac{1}{\mathbf{E}[\#\mathbf{B}_0]} \mathbf{E} \left[\sum_{w \in V_0(\mathbf{B})} 1_{\left\{[\overline{\mathbf{B}}, w] \in A\right\}} \right], \qquad A \in \mathcal{B}(\mathcal{G}_*). \qquad (13)$$

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 3.5, $\overline{\mathbf{B}}$ 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

$$
\int_{\mathcal{G}_*} \sum_{v \in V(\Gamma)} g[\Gamma, o, v] \, \mathbf{P}^{\square}(d[\Gamma, o])
$$

$$
= \frac{1}{\mathbf{E}[\#\mathbf{B}_0]} \mathbf{E} \sum_{y \in \mathbf{B}_0} \sum_{v \in V(\mathbf{B})} g[\overline{\mathbf{B}}, (0, y), v]
$$

$$
= \frac{1}{\mathbf{E}[\#\mathbf{B}_0]} \sum_{y, y' \in S, t \in \mathbb{Z}} \mathbf{E}\left[1_{\{(0,y),(t,y') \in V(\mathbf{B})\}} g[\overline{\mathbf{B}}, (0, y), (t, y')] \right].
$$

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

$$
\frac{1}{\mathbf{E}[\#\mathbf{B}_0]} \sum_{y, y' \in S, t \in \mathbb{Z}} \mathbf{E}\left[1_{\{(-t,y),(0,y') \in V(\mathbf{B})\}} g[\overline{\mathbf{B}}, (-t, y), (0, y')] \right]
$$

$$
= \frac{1}{\mathbf{E}[\#\mathbf{B}_0]} \sum_{y, y' \in S, t \in \mathbb{Z}} \mathbf{E}\left[1_{\{(t,y),(0,y') \in V(\mathbf{B})\}} g[\overline{\mathbf{B}}, (t, y), (0, y')] \right]
$$

$$
= \frac{1}{\mathbf{E}[\#\mathbf{B}_0]} \mathbf{E} \sum_{y' \in \mathbf{B}_0} \sum_{v \in V(\mathbf{B})} g[\overline{\mathbf{B}}, v, (0, y')]
$$

$$
= \int_{\mathcal{G}_*} \sum_{v \in V(\Gamma)} g[\Gamma, v, o] \, \mathbf{P}^{\square}(d[\Gamma, o]).
$$

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

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

Proposition 4.2. *Let 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 $[\overline{\mathbf{B}}, o]$ 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 $C := \mathbf{E}[\#\mathbf{B}_0]$,

$$
\hat{\mathbf{P}}([\overline{\mathbf{B}}, o] \in A)
$$

$$
= \frac{1}{C} \mathbf{E}[\#\mathbf{B}_0 1_{\{[\overline{\mathbf{B}}, o] \in A\}}]
$$

$$
= \frac{1}{C} \sum_V |V| \mathbf{P}(V_0(\mathbf{B}) = V) \mathbf{P}([\overline{\mathbf{B}}, o] \in A \mid V_0(\mathbf{B}) = V)
$$

$$
= \frac{1}{C} \sum_V \sum_{v \in V} |V| \mathbf{P}(V_0(\mathbf{B}) = V) \mathbf{P}(o = v, [\overline{\mathbf{B}}, v] \in A \mid V_0(\mathbf{B}) = V)
$$

18

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

$$= \frac{1}{C} \sum_V \sum_{v \in V} |V| \, \mathbf{P}(V_0(\mathbf{B}) = V) \frac{1}{|V|} \mathbf{P}([\overline{\mathbf{B}}, v] \in A \mid V_0(\mathbf{B}) = V)$$

$$= \frac{1}{C} \mathbf{E} \left[\sum_V \sum_{v \in V} 1_{\left\{ [\overline{\mathbf{B}}, v] \in A, V_0(\mathbf{B}) = V \right\}} \right]$$

$$= \mathbf{P}^{\square}(A)$$

as claimed. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

4.2. 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

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

This immediately gives the following.

Lemma 4.3. *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.* $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

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. From now on, 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 Sections 4.2.1 and 4.2.2.

4.2.1. 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 $\mathbf{G}$ because $\mathbf{G}$ is disconnected.

Example 4.4. 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 $\mathbf{G}$ has two components a.s. Each component is itself a bi-infinite path.

Moreover, even when $\mathbf{G}$ is connected, it it still possible to have multiple bi-infinite paths in $\mathbf{G}$.

Example 4.5. 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 $\mathbf{G}$ 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 $\mathbf{G}$ is an EFT, which follows from Proposition 2.3.

In Example 4.5, even though $\mathbf{G}$ is connected, $\mathbf{G}$ 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. 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-reccurent (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 4.6. *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 3.1, $\mathbf{P}^{\square}$-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 **B**, **P**-a.s. every component of **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 3.2, 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 **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 **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 **G** because, by definition, a bi-recurrent state path in **G** will lie in **B** since it visits x^* at arbitrarily large negative times. $\square$

The next result applies Proposition 4.6 to the nicest case, where **G** is a tree.

Corollary 4.7. *Suppose that* **G** *is an EFT. Then* **G** *contains a unique (up to measure zero modifications) state path* $(\beta_t)_{t \in \mathbb{Z}}$ *that is bi-recurrent for* x^*. *Moreover,* $(\beta_t)_{t \in \mathbb{Z}}$ *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 4.6 shows that a.s. there is a unique bi-infinite path in each component of **B**, and the corresponding state paths are bi-recurrent for x^*. Since **G** a.s. has only one component, **B** does too. The second part of Proposition 4.6 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. 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 4.7 shows that, like in the standard CFTP setup, there is a β_0 living at time 0 in **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 **B** is that one can bound the number of components of **B**.

Corollary 4.8. *The a.s. constant number n of components of* **B** *is no larger than* $\min\{k : \mathbf{P}(\#\mathbf{B}_0 = k) > 0\} < \infty$. *In particular,* **B** *has finitely many connected components, even if* **G** *has infinitely many components, and if* $\mathbf{P}(\#\mathbf{B}_0 = 1) > 0$, *then* **B** *is an EFT.*

Proof. The number of components of **B** is shift-invariant and hence a.s. constant. Each component of **B** contains a bi-recurrent path by Proposition 4.6. Each bi-recurrent path intersects $V_0(\mathbf{B})$ in a different element since they are in different components of **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 4.8 can be achieved for each n. In general, any bi-infinite stationary process on S (or any countable set) must be bi-recurrent.

Proposition 4.9. *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 4.10. *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 2.2, 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 2.3, **G** is an EFT and therefore Corollary 4.7 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 $t \in \mathbb{Z}$,
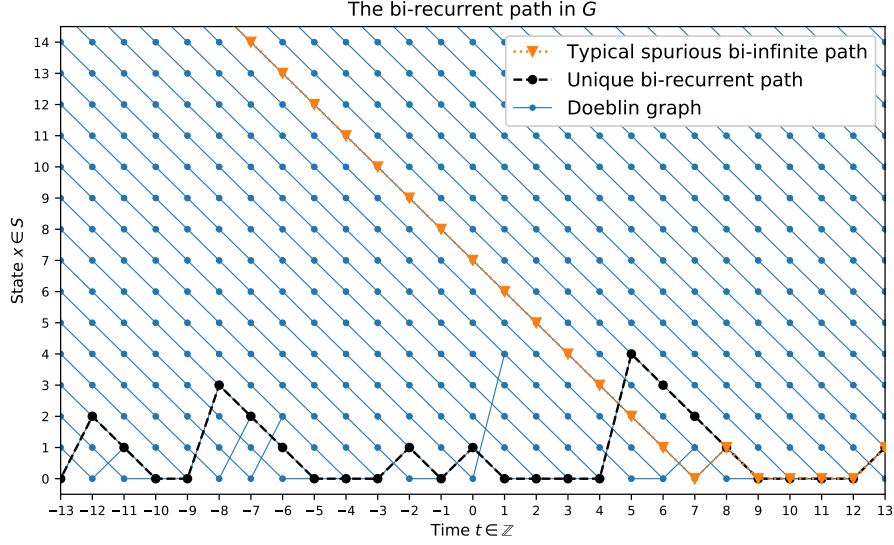
Figure 3: The Doeblin graph from Example 4.5 with the bi-recurrent path and a spurious path distinguished.

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 4.9 and irreducibility. $\square$

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 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 4.5.

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 $\sup_{x\in S} \|P^n(x,\cdot) - \pi\| \to 0$ 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. [19] 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. [10] 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 dis-

tribution by any finite time.

Proposition 4.11. *Suppose P^n converges uniformly to π as $n \to \infty$ and $\mathbf{G}$ has fully independent transitions. Then $\mathbf{G}$ 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 $\mathbf{G}$ collapse to a single state by time t. Note that $\mathbf{P}(C_{s,t})$ depends only on $t - s$. Since $\mathbf{G}$ has fully independent transitions and P^n converges to π uniformly as $n \to \infty$, by e.g. Theorem 5.2 in [10], 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 $\mathbf{G}$, and thus any bi-infinite path in $\mathbf{G}$ must coincide with the unique bi-recurrent path guaranteed to exist by Corollary 4.7. $\square$

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 conveges uniformly to its stationary distribution, then it must be that $X_t \sim \pi$ for all $t \in \mathbb{Z}$.

Proposition 4.12. *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 2.2, 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 2.3. Since P^n converges uniformly to π as $n \to \infty$, $\mathbf{G}$ contains no spurious bi-infinite paths by Proposition 4.11, and hence $(X_t)_{t \in \mathbb{Z}}$ must be the bi-recurrent state path. Theorem 4.10 then implies $(X_t)_{t \in \mathbb{Z}}$ is stationary. $\square$

Proposition 4.12 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 4.5, 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 4.13. *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

$$\bigcap_{x \in S} V(\mathbf{B}(x)) = \{(t, \beta_t) : t \in \mathbb{Z}\}, \tag{14}$$

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.

$$\bigcap_{x \in S} V(\mathbf{B}(x)) \supsetneq \{(t, \beta_t) : t \in \mathbb{Z}\}. \tag{15}$$

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

$$L_0 := \{(t, x_t) : t = 0, \ldots, n\}, \qquad L_1 := \{(t + 1, x_t) : t = 0, \ldots, n\} \tag{16}$$

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 $\mathbf{G}$ and therefore L_0 makes up a segment of the bi-infinite path in $\mathbf{G}$. 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 $\mathbf{G}$. $\qquad\square$

4.2.2. 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 4.2.1, which centered around bi-recurrent paths in $\mathbf{B}$. However, there are many other potential things to say about $\mathbf{B}$ 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 $\mathbf{B}$ 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 $\mathbf{B}$, 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 $\mathbf{B}$ 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 4.3.1 and is the main topic of Section 4.3.2.

Additionally, the list of mass transports given in the appendix 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 $\mathbf{B}$ 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 $\mathbf{B}$ as a sequence of vertical slices, the vertices are partitioned by which slice they are in. In the view of $\mathbf{B}$ 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 the appendix 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 the appendix for a more detailed description of these results and other finer-grained results.

4.3. Applications to Simulating the Bridge Graph

4.3.1. Local Weak Convergence to the Bridge Graph

It was shown in Proposition 4.2 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 [1]. 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 $V_I(\mathbf{B}) := \cup_{t \in I} V_t(\mathbf{B})$ and let $\overline{\mathbf{B}} \cap I$ denote the subgraph of $\overline{\mathbf{B}}$ induced by $V_I(\mathbf{B})$. Also define $V_I'(\mathbf{B}) := \cup_{t \in I} \left\{ (s, F_s^{(t, x^*)}) : t \leqslant s \leqslant \sup I \right\}$ 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 4.14. *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*

$$\frac{1}{\#I_n \mathbf{E}[\#\mathbf{B}_0]} \sum_{v \in V_{I_n}(\mathbf{B})} g[\overline{\mathbf{B}}, v] \to \mathbf{E}^\square[g] \tag{17}$$

and

$$\frac{1}{\#V_{I_n}(\mathbf{B})} \sum_{v \in V_{I_n}(\mathbf{B})} g[\overline{\mathbf{B}}, v] \to \mathbf{E}^\square[g], \tag{18}$$

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

$$\frac{\#V_{I_n}(\mathbf{B})}{\mathbf{E}[\#V_{I_n}(\mathbf{B})]} = \frac{\#V_{I_n}(\mathbf{B})}{\#I_n \mathbf{E}[\#\mathbf{B}_0]} \to 1. \tag{19}$$

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. [17]) then proves the claim. $\square$

Proposition 4.15. *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*

$$\frac{1}{\#V_{I_n}(\mathbf{B})} \sum_{v \in V_{I_n}(\mathbf{B})} g[\overline{\mathbf{B}} \cap I_n, v] \to \mathbf{E}^\square[g], \qquad \mathbf{P}\text{-}a.s. \tag{20}$$

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

$$\mathbf{P}([\overline{\mathbf{B}} \cap I_n, \boldsymbol{o}_n] \in \cdot) \to \mathbf{P}^\square, \qquad n \to \infty \tag{21}$$

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 $n \to \infty$, $\mathbf{P}$-a.s., by Lemma 4.14. But also $c_n \to \mathbf{E}^{\square}[g]$ and $a_n \to 1$, $\mathbf{P}$-a.s., also by Lemma 4.14. Hence $a_n b_n \to \mathbf{E}^{\square}[g]$, $\mathbf{P}$-a.s., as claimed. $\qquad\square$

Proposition 4.16. *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 $\boldsymbol{o}'_n$ 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*

29

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

$$\frac{1}{\#(V'_{I_n}(\mathbf{B}) \cap [0, b_n])} \sum_{v \in V'_{I_n}(\mathbf{B}) \cap [0, b_n]} g[\overline{\mathbf{B}} \sqcap I_n, v] \to \mathbf{E}^{\square}[g], \qquad \mathbf{P}\text{-}a.s. \qquad (22)$$

In particular,

$$\mathbf{P}([\overline{\mathbf{B}} \sqcap I_n, \boldsymbol{o}'_n] \in \cdot) \to \mathbf{P}^{\square}, \qquad n \to \infty \qquad (23)$$

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

$$\frac{1}{\#(V'_{I_n}(\mathbf{B}) \cap [0, b_n])} \sum_{v \in V'_{I_n}(\mathbf{B}) \cap [0, b_n]} g[\overline{\mathbf{B}} \sqcap I_n, v]$$

$$= \frac{1}{\#V_{[0, b_n]}(\mathbf{B})} \sum_{v \in V_{[0, b_n]}(\mathbf{B})} g[\overline{\mathbf{B}} \cap I_n, v]$$

$$+ \frac{1}{\#V_{[0, b_n]}(\mathbf{B})} \sum_{v \in V_{[b_n - N + 1, b_n]}(\mathbf{B})} (g[\overline{\mathbf{B}} \sqcap I_n, v] - g[\overline{\mathbf{B}} \cap I_n, v]).$$

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 Proposition 4.15, so it suffices to show that the last term goes to 0. Indeed,

$$\left| \frac{1}{\#V_{[0, b_n]}(\mathbf{B})} \sum_{v \in V_{[b_n - N + 1, b_n]}(\mathbf{B})} (g[\overline{\mathbf{B}} \sqcap I_n, v] - g[\overline{\mathbf{B}} \cap I_n, v]) \right|$$

$$\leqslant \frac{2\|g\|_\infty \#V_{[b_n - N + 1, b_n]}(\mathbf{B})}{\#V_{[0, b_n]}(\mathbf{B})}$$

$$= \frac{2\|g\|_\infty (\#V_{[0, b_n]}(\mathbf{B}) - \#V_{[0, b_n - N]}(\mathbf{B}))}{\#V_{[0, b_n]}(\mathbf{B})}$$

$$\to 2\|g\|_\infty (1 - 1)$$

$$= 0$$

as desired. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

4.3.2. 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 2.13, $S_\mathbf{B}$ consists only of finite subsets of S and is therefore countable.

Lemma 4.17. *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. $\square$

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

Lemma 4.18. *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 2.13, $\#\mathbf{B}_t$ 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 $\mathbf{G}$. $\square$

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 4.19. *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 2.13, $\#\mathbf{B}_t$ 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 4.17, $(\mathbf{B}_t)_{t \in \mathbb{Z}}$ 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 4.18 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. $\qquad\square$

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

Example 4.20. 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 4.21. *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 4.19 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 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.
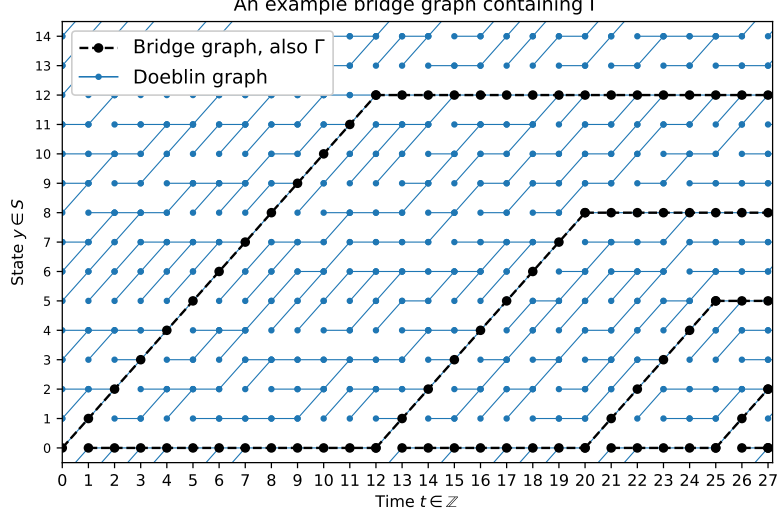
Figure 4: The graph Γ from the proof of Proposition 4.21 when $S = \mathbb{Z}/15\mathbb{Z}$, $x^* = 0$, and $E = \{0, 2, 5, 8, 12\}$, where the Markov dynamics are the lazy version of the deterministic cycle $x \mapsto x + 1$ on S. The graph Γ is constructed so that, as shown in the figure, if $\mathbf{B}_0 = \{0\}$ and $\Gamma \subseteq \mathbf{G}$, then $\mathbf{B}_{27} = \{0, 2, 5, 8, 12\}$.

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

33

(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 $y \in S$. $\qquad\square$

Proposition 4.22. *Suppose* $\mathbf{G}$ *has fully independent transitions. Extend the definition of* $P_{\mathbf{B}}$ *to*

$$P_{\mathbf{B}}(E, E') := \mathbf{P}\left(\{x^*\} \cup \left\{F_1^{(0,y)} : y \in E\right\} = E'\right), \tag{24}$$

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* $E' = \{x^*, y_1, \ldots, y_m\}$,

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

$$+ \sum_{i=1}^{m} p_{x_n, y_i} P_{\mathbf{B}}(E \setminus \{x_n\}, E' \setminus \{y_i\}), \tag{25}$$

with recursive depth at most n *and base cases*

$$\begin{cases} P_{\mathbf{B}}(E, E') = 0, & \#E' > \#E + 1 \\ P_{\mathbf{B}}(\{x^*\}, \{x^*, y\}) = p_{x^*, y}, & y \in S \\ P_{\mathbf{B}}(E, \{x^*\}) = \prod_{y \in E} p_{y, x^*}. \end{cases} \tag{26}$$

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

$$P_{\mathbf{B}}(E, E') = \mathbf{P}(\mathbf{B}_1 = E' \mid \mathbf{B}_0 = E)$$

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

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

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

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 $E' = \{x^*, y_1, \ldots, y_m\}$ 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

$$\{x^*\} \cup \left\{ F_1^{(0,y)} : y \in E \setminus \{x_n\} \right\} = E' \setminus \{y_i\},$$

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

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

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

which simplifies to

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

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 4.23. By implementing the recurrence of Proposition 4.22 in, e.g. Python, one may compute $P_{\mathbf{B}}$ explicitly. Then, given values for the $p_{x,y}$, one may compute the staitonary 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

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

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.

5. Bibliographical Comments

While this work may be the first time the Doeblin graph $\mathbf{G}$ 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 authors' 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 [6], 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 **B** is an EFT. That is, it is the condition that all paths in **B** eventually merge. It is conceivable that, in the EFT case, one could derive the existence of the bi-infinite path in **B** from the work in [6], 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 paper to distinguish this bi-infinite path from the potential others in **G**.

Another occurrence of a similar object to the bridge graph may be found in [3] 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 [3] 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 [3] is then similar in nature to the sequence of cardinalities of $(\mathbf{B}_t)_{t \in \mathbb{Z}}$ as considered in Section 4.3.2. It is proved in [3] 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 4.19 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.

6. Appendix

6.1. Postponed Proofs

Proofs that were only sketched in the main text are collected in full detail here.

Proof of Theorem 2.2. 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 2.1 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 $x \in S$,

$$\mathbf{P}(F_t^{(s, X_s)} = x) = \sum_{y \in S} \mathbf{P}(X_s = y, F_t^{(s, X_s)} = x)$$
$$= \sum_{y \in S} \mu_s(y) P^{t-s}(y, x)$$
$$= \mu_s P^{t-s}(x),$$

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

$$\mu_s P^{t-s} = \mu_r P^{s-r} P^{t-s} = \mu_r P^{t-r} = \mu_t \tag{27}$$

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 $t_0 \leqslant t$.

37

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. $\qquad \square$

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 6.1 (Subsumes Proposition 2.3). *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 $(\mathcal{C}_j^i)_{0 \leqslant i < N, 0 \leqslant j < d(i)}$ of $\mathbf{G}$ are in bijection with $(C_j^i)_{0 \leqslant i < N, 0 \leqslant j < d(i)}$, and for $x \in C_j^i$, $(t, x) \in V(\mathcal{C}_{j'}^i)$ if and only if $j - t = j' \pmod{d(i)}$, and for $x \in T$, $(t, x) \in V(\mathcal{C}_j^i)$ where $(t', y) \in V(\mathcal{C}_j^i)$ is any vertex on the path of (t, x) for which y is recurrent. That is, $\mathcal{C}_j^i$ 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. $\square$

Proof of Lemma 3.3. 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

$$\{\boldsymbol{o} = v_0\} \cap \bigcap_{i=1}^{n} \left(\{f_V(v_i) = 1\} \cap \{f_E(v_{i-1}, v_i) = 1\}\right).$$

From here one sees that event that the r-neighborhood around $\boldsymbol{o}$ is exactly some fixed finite graph Γ is measurable. Indeed,

$$\{N_{\mathbf{\Gamma}}(\boldsymbol{o}, r) = \Gamma\} = \bigcap_{v \in V} \{(f_V(v) = 1 \text{ and } d_{\mathbf{\Gamma}}(\boldsymbol{o}, v) \leqslant r) \iff v \in V(\Gamma)\}.$$

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

$$
\begin{aligned}
D_{r,\epsilon}(\Gamma, o) :=\{&\boldsymbol{o} = o, \\
&N_{\mathbf{\Gamma}}(\boldsymbol{o}, r) = \Gamma, \\
&\forall u \in V(\Gamma), d_{\Xi_{\text{univ}}}(\xi_V(u), \xi_u) < \epsilon, \\
&\forall \{v, w\} \in E(\Gamma), d_{\Xi_{\text{univ}}}(\xi_E(v, w), \xi_{v,w}) < \epsilon\}
\end{aligned}
$$

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 $\{d_{\mathcal{G}_*}([\mathbf{\Gamma}, \boldsymbol{o}], [\Gamma, o]) < \epsilon\}$ 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 $\omega \mapsto [\mathbf{\Gamma}(\omega), \boldsymbol{o}(\omega)]$ is measurable. $\square$

Proof of Proposition 3.4. 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 $g : \mathcal{G}_{**} \to$

$\mathbb{R}_{\geqslant 0}$ supported on directed neighbors be given. Then

$$
\begin{aligned}
\mathbf{E} \sum_{v \in V(\mathbf{G})} g[\mathbf{G}, (0, X_0), v] &= \frac{1}{N} \sum_{x \in S} \mathbf{E} \sum_{y \in S} g[\mathbf{G}, (0, x), (1, y)] \\
&= \frac{1}{N} \sum_{x, y \in S} \mathbf{E}[g[\mathbf{G}, (0, x), (1, y)]] \\
&= \frac{1}{N} \sum_{x, y \in S} \mathbf{E}[g[\mathbf{G}, (-1, x), (0, y)]] \\
&= \frac{1}{N} \sum_{y \in S} \mathbf{E} \sum_{x \in S} g[\mathbf{G}, (-1, x), (0, y)] \\
&= \mathbf{E} \sum_{v \in V(\mathbf{G})} g[\mathbf{G}, v, (0, X_0)],
\end{aligned}
$$

where in the third equality time-homogeneity of $\mathbf{G}$ is used. It follows that in this case $\mathbf{G}$ is unimodular.

Next 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 $g[G, v, w] := 1_{\{s(v)=x, s(w)=y, v \sim w\}}$. 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.

Next 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. $\qquad\square$

6.2. List of Mass Transports

As mentioned in Section 2.5, the proof style of Proposition 2.13 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 $\mathbf{B}$.

- $\mathbf{E}[N_0^{(0,x^*)}(y;\beta)] = \mathbf{E}[1_{\{y \in \mathbf{B}_0 \backslash \beta_0\}} \#V^{x^*}(D^{(0,y)} \cap V(\mathbf{B}))]$, 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 $y \in S$, $\mathbf{E}[\sigma_0^{(0,x^*)}(\beta)] = \mathbf{E}[\#V^{x^*}(D^{V_0(\mathbf{B})\backslash V_0(\beta)} \cap V(\mathbf{B}))]$, where $\sigma_0^{(0,x^*)}(\beta)$ is the number of steps (potentially 0) before $F^{(0,x^*)}$ merges with β, and $D^{V_0(\mathbf{B})\backslash V_0(\beta)}$ 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 $\beta_s = x^*$.

41

- $\mathbf{E}[N^{(0,x^*)}(y;x^*)1_{\beta_0=x^*}] = \mathbf{P}(\beta_0 = y)$, 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 map 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 3.1, for the network $[\overline{\mathbf{B}}, \square]$ directly. It seems more elegant to call upon the machinery of unimodular networks when convenient instead.

Acknowledgments

7. References

[1] D. Aldous and R. Lyons. Processes on unimodular random networks. *Electron. J. Probab.*, 12:no. 54, 1454–1508, 2007.

[2] 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.

[3] F. Baccelli and A. Sodre. Renewal population dynamics and their eternal family trees. *arXiv preprint arXiv:1803.08081*, 2018.

[4] A. Borovkov and S. G. Foss. Two ergodicity criteria for stochastically recursive sequences. *Acta Appl. Math.*, 34(1-2):125–134, 1994.

[5] A. A. Borovkov. *Probability theory*. Universitext. Springer, London, 2013. Translated from the 2009 Russian fifth edition by O. B. Borovkova and P. S. Ruzankin, Edited by K. A. Borovkov.

[6] 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.

[7] 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.

[8] 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.

[9] S. Foss and T. Konstantopoulos. Extended renovation theory and limit theorems for stochastic ordered graphs. *Markov Process. Related Fields*, 9(3):413–468, 2003.

[10] S. G. Foss and R. L. Tweedie. Perfect simulation and backward coupling. *Comm. Statist. Stochastic Models*, 14(1-2):187–203, 1998. Special issue in honor of Marcel F. Neuts.

[11] 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.

[12] 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.

[13] 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.

[14] W. S. Kendall. Geometric ergodicity and perfect simulation. *Electron. Comm. Probab.*, 9:140–151, 2004.

[15] 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.

[16] A. Khezeli. Shift-coupling of random rooted graphs and networks. *arXiv preprint arXiv:1703.06619*, 2017.

[17] E. Lindenstrauss. Pointwise theorems for amenable groups. *Invent. Math.*, 146(2):259–295, 2001.

[18] 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.

[19] S. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second edition, 2009. With a prologue by Peter W. Glynn.

[20] J. Møller. Perfect simulation of conditionally specified models. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 61(1):251–264, 1999.

[21] 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.

[22] D. J. Murdoch and P. J. Green. Exact sampling from a continuous state space. *Scand. J. Statist.*, 25(3):483–502, 1998.

[23] 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.

[24] 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.

[25] 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. 7th Annual ACM-SIAM Symposium on Discrete Algorithms (Atlanta, GA, 1996).

[26] D. B. Wilson. How to couple from the past using a read-once source of randomness. *Random Structures Algorithms*, 16(1):85–113, 2000.