Non-Markovian Interacting Particle Systems on Large Sparse Graphs:

Hydrodynamic Limits and Marginal Characterizations

by

Ankan Ganguly

B. Sc. in Mathematics and Economics, University of Minnesota, Twin Cities; 2015

B. Sc. in Computer Science (Magna Cum Laude), University of Minnesota, Twin Cities; 2015

M. Sc. in Applied Mathematics, Brown University; 2019

A dissertation submitted in partial fulfillment of the

requirements for the degree of Doctor of Philosophy

in the Department of Applied Mathematics at Brown University

Providence, Rhode Island

October 2022

This dissertation by Ankan Ganguly is accepted in its present form by the Department of Applied Mathematics as satisfying the dissertation requirement for the degree of Doctor of Philosophy.

Date ______________ ______________________________
Kavita Ramanan, Ph.D., Advisor

Recommended to the Graduate Council

Date ______________ ______________________________
Rick Durrett, Ph.D., Reader

Date ______________ ______________________________
Stuart Geman, Ph.D., Reader

Approved by the Graduate Council

Date ______________ ______________________________
Thomas A. Lewis
Dean of the Graduate School

Curriculum Vitae

Ankan Ganguly graduated Magna Cum Laude in 2015 from the University of Minnesota, Twin Cities with a Bachelor of Science in mathematics, economics and another Bachelor of Science in computer science. He received his Master of Science in applied mathematics from Brown in 2019. During his time at Brown university, he co-mentored two undergraduate students, Mitchell Wortsman and Timothy Sudijono, as they completed their undergraduate honors theses under his advisor Kavita Ramanan. He is also a recipient of the Dunmu Ji Award.

Acknowledgments

I am deeply grateful to my advisor, Kavita Ramanan for the immense support and advice she has given me over the years. She has spent countless hours reading and proofreading drafts of my work and I am a better writer and researcher for it. I will carry these lessons with me for the rest of my professional career.

I would also like to thank my thesis committee members, Stuart Geman and Rick Durrett, for taking the time to read my thesis and advise me as well as attending my thesis defense. I also would not be here without the tireless work by the department staff for their organization and for making the process of putting together and submitting this thesis go as smoothly as possible.

Thank you also to my friend and coauthor Alex McDonough for our many fascinating talks about the interface between probability, graph theory and combinatorics. I greatly enjoyed the experience of writing our paper [35] together, and I find it unfortunate that I was unable to highlight some of our findings from that paper within this thesis.

Lastly, I would like to thank my friends and family who had to put up with me while I stressed over deadlines and a seemingly endless editing process. Who gave me the support I needed to get where I am now and who I will continue to rely on moving forward.

My work was supported in part by the United States Army Research Office grant W911NF-20-1-0133.

Contents

List of Tables

List of Figures

Acronyms	
IPS	Interacting Particle Systems
ODE	Ordinary Differential Equation
SDE	Stochastic Differential Equation
(U)GW tree	(Unimodular) Galton-Watson Tree (Definitions 2.13, 2.16)
(SG)MRF	(Semi-Global) Markov Random Field (Definition
MA	Markov Approximation
PA	Pairwise Approximation
MFA	Mean-Field Approximation
Notation: Graph Notation	
V_H, E_H, $\varnothing_H$	the respective vertex set, edge set and root of a (rooted) graph H
$\|A\|$	the cardinality of a set A.
$\mathscr{C}_v(G)$	the largest connected subgraph of G containing v also called the *connected component* of G containing v. It is a rooted graph with v equipped as the root.
$d_G(u, v)$	graph distance: length of the shortest path from u to v in G
$\mathcal{N}_v$, $\mathcal{N}_v(G)$, $\mathcal{N}_v^\alpha$, $\mathcal{N}_v^\alpha(G)$	$u \in V_G \setminus \{v\}$ such that $d_G(u, v) = 1$ (respy. $d_G(u, v) \leq \alpha$)
$\mathcal{N}_U$, $\mathcal{N}_U(G)$, $\mathcal{N}_U^\alpha$, $\mathcal{N}_U^\alpha(G)$	$u \in V_G \setminus U$ such that $d_G(u, U) = 1$ (respy. $d_G(u, U) \leq \alpha$)
$\mathrm{cl}_v, \mathrm{cl}_U, \mathrm{cl}_v(G), \mathrm{cl}_U(G)$	$\{v\} \cup \mathcal{N}_v$, $U \cup \mathcal{N}_U$, $\{v\} \cup \mathcal{N}_v(G)$ and $U \cup \mathcal{N}_U(G)$
Λ_G	set of finite subsets of V_G
$G[U]$	(unrooted) subgraph of G induced by $U \subseteq V_G$
V_n^G	$\{v \in V_G : d_G(\varnothing, v) \leq n\}$

$B_n(G)$	$(G[V_n^G], \varnothing)$
Notation: Tree Notation	
$\mathbb{T}_d := (V^d, E^d, \varnothing)$	The rooted d-regular tree
V_n^d	$\{v \in V^d : d_{\mathbb{T}_d}(\varnothing, v) \leq n\}$
c_v, $c_v(\mathcal{T})$	the children of v in $\mathcal{T}$: $\{u \in \mathrm{cl}_v(\mathcal{T}) : d_{\mathcal{T}}(u, \varnothing) > d_{\mathcal{T}}(v, \varnothing)\}$
π_v	the parent of v in $\mathcal{T}$: $u \in \mathrm{cl}_v(\mathcal{T})$ s.t. $d_{\mathcal{T}}(\varnothing, u) < d_{\mathcal{T}}(\varnothing, v)$
$u \leq v$	a partial order on the vertices of a rooted tree. u is descended from v.
$\mathcal{L}(\mathcal{T})$	$\{v \in V_{\mathcal{T}} : c_v(\mathcal{T}) = \emptyset\}$
Notation: Spaces	All spaces are Polish unless otherwise stated
$\mathcal{Z}$	Generic Polish space
$\mathcal{S}$	Generic countable space (assumed to be discrete)
$\mathscr{N}(\mathcal{Z})$	Non-negative integer-valued locally finite measures on $\mathcal{Z}$
$\mathcal{P}(\mathcal{Z})$	Probability measures on $\mathcal{Z}$
$\mathcal{X}$	State space of the IPS
$\overline{\mathcal{K}}$, $\mathcal{K}$	respective sets of edge and vertex marks encoding the parameters of the model
$\mathcal{J}$	set of allowable jumps of the IPS
$\mathcal{D}(I, \mathcal{Z})$, I is an interval in $\mathbb{R}$	space of càdlàg functions from I to $\mathcal{Z}$ equipped with the J1 topology
$\mathcal{D}$, $\mathcal{D}_t$, $\mathcal{D}_{t-}$	$\mathcal{D}(\mathbb{R}_+, \mathcal{X})$, $\mathcal{D}([0,t], \mathcal{X})$, $\mathcal{D}([0,t), \mathcal{X})$
$\overline{\mathcal{D}}$, $\overline{\mathcal{D}}_t$, $\overline{\mathcal{D}}_{t-}$	$\mathcal{D}(\mathbb{R}, \mathcal{X})$, $\mathcal{D}((-\infty,t], \mathcal{X})$, $\mathcal{D}((-\infty,t), \mathcal{X})$
$\mathscr{G}_*[\overline{\mathcal{K}}, \mathcal{K}]$	space of isomorphism classes of rooted graphs with edge marks in $\overline{\mathcal{K}}$ and vertex marks in $\mathcal{K}$

$\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K}]$	space of marked, rooted graphs G with edge marks in $\overline{\mathcal{K}}$, vertex marks in $\mathcal{K}$ and $V_G \subseteq \mathbb{N}$
$\widehat{\mathscr{G}}^{\mathcal{S}}_{*,1}$	the discrete space of rooted graphs of radius 1 whose vertices lie in the countable, deterministic set $\mathcal{S}$.
$\widehat{\mathscr{G}}_{*,1}$	the space of rooted graphs of radius 1. This space is not Polish.
$\mathscr{G}_{*,1}[\overline{\mathcal{K}}, \mathcal{K}]$	the subspace of $\mathscr{G}_*[\overline{\mathcal{K}}, \mathcal{K}]$ whose graph representatives all have radius 1.
Notation: Miscellaneous	
$\mathcal{B}(\mathcal{Z})$	Borel σ-algebra on $\mathcal{Z}$
$\mathcal{H}^Z$	the smallest, complete σ-algebra containing $\sigma(Z)$
$\mathbb{H}^X$	if X is a $\mathcal{D}^U$-random element, then $\mathbb{H}$ is the smallest filtration satisfying the usual conditions such that X is $\mathbb{H}$-adapted.
$\mathrm{Disc}_t\,(x)$	If $x \in \mathcal{D}(I, \mathcal{X})$ for an interval $I \subseteq \mathbb{R}$, then $\mathrm{Disc}_t\,(x)$ is set of times $s \in I \cap (-\infty, t]$ such that $\Delta x(t) \neq 0$.
$\{t_x(k), v_x(k), j_x(k)\}$	jump characteristics of x (Definition 8.4). The subscript is omitted when x is clear from the context.

Table 1: Glossary of Notation

Part I

Introduction

CHAPTER 1

Introduction

An *interacting particle system* (IPS) is a class of mathematical models consisting of a collection of stochastic processes (called particles) indexed by the vertices of a graph. Each particle takes values in a countable state space and evolves according to a pure jump process whose transition kernels depend upon the state of the particle and the state of its neighbors in the underlying graph.

One simple example of an interacting particle system is the contact process [47] (see Example 3.15 for a definition of the contact process within our framework). This is a highly stylized model of the spread of infection through a population that is the topic of ongoing interest. In this model, the underlying graph G represents the contact network, in which vertices represent people and edges represent interactions by which an infection can spread. Then each infected particle recovers at unit rate while susceptible particle are infected at a rate proportional to the number of infected neighbors at any given time (here, the constant of proportionality is called the *infection rate*).

We represent IPS as solutions to a Poisson-driven stochastic differential equation (SDE). Our framework covers standard IPS such as the contact process, voter model and Glauber dynamics for the Ising model (and other Gibbs measures). In fact, our results are stated in a more general context that accounts for heterogeneous models such as the contact process with individual recovery and infection rates and non-Markovian processes such as the renewal contact process (see Section 3.2). Such models will be referred to as IPS for the remainder of the thesis. IPS generally arise as models in a wide variety of fields including statistical physics [70], epidemiology [15, 51, 77, 11], neuroscience

[87], opinion dynamics [65, 83, 23, 49, 4], engineering and operations research [39, 22]. Indeed, the field of interacting particle systems is enormous. See [65, 67, 12, 55] for more information.

The challenge of working with IPS is that the complexity of the process increases exponentially with the size of the graph. For example, the contact process on a rooted graph $G = (V, E, \varnothing)$ is a Markov process with a state space of size $2^{|V|}$. As a result, IPS on large graphs are often both numerically and analytically intractable. For this reason, we are interested in limiting theorems of IPS on (possibly random) sparse graph sequences. In particular, we assume that a sequence of finite, sparse rooted graphs $G_n := (V_n, E_n, \varnothing_n)$, $n \in \mathbb{N}$, with increasing vertex sets can be approximated by the infinite rooted graph G, and we let X^{G_n}, $n \in \mathbb{N}$, and X^G represent IPS on the respective graphs. Then this thesis addresses the following four questions, which we will address for the remainder of the thesis:

Question 1. Well-Posedness: is X^G even a well-defined quantity?

Question 2. Convergence: if the initial conditions converge and $G_n \to G$ in some sense, then when can we say that $X^{G_n} \to X^G$?

Question 3. Hydrodynamic Limit: if

$$\pi^{G_n} := \frac{1}{|G_n|} \sum_{v \in V_{G_n}} \delta_{X_v^{G_n}},$$

then does π^{G_n} have a deterministic limit as $n \to \infty$, and if so, what is it?

Question 4. Autonomous Characterization of the Marginals: is there some autonomous characterization of the behavior of a "typical" particle in the limiting IPS X^G?

In Part II, we focus on questions 1-3. Answering Question 1 reduces to proving that the SDE defining X^G is well-posed. We answer Question 1 in the affirmative given certain conditions on the transition kernels of the IPS (which we characterize using certain *jump rate functions*) and the underlying graph. We show that this is not trivial by constructing a graph G and a simple collection of jump rate functions such that the SDE defining X^G is not well-posed. We next show that Question 2 holds (under weak conditions) with respect to the topology of *local weak convergence* (first introduced by Benjamini and Schramm [9]). We also discuss two other modes of convergence under which the result holds. One is a topology on a novel space $\widehat{\mathscr{G}}_*[\cdot]$ first introduced in [36]

(Chapters 2-5 of this thesis replicate and extend the results of this paper). The other is *convergence in probability in the local weak sense* which was shown in [61, Lemma 2.7] to imply the following hydrodynamic limit:

$$\pi^{G_n} \to \mathrm{Law}(X^G_\varnothing) \text{ in probability.}$$

In Part III, we focus on Question 4. Note that the display above demonstrates why it is natural to want to understand the marginal distribution of X^G at its root (which we take to be the "typical" particle). However, it should be clear that this quantity is heavily dependent on the graph G itself, which raises the question of on what graph G should we study $\mathrm{Law}(X^G_\varnothing)$? It is well known that the majority of well-studied random graph models converge (up to appropriate scaling) locally weakly to random trees, so a natural starting point is the infinite regular tree $\mathbb{T}_d$. So that is the focus of this thesis.

In a recent, innovative paper, Lacker et al. derived the local equation: a non-Markovian, measure-dependent SDE depending only on its initial condition that characterizes the marginal dynamics at the root and its neighbors of an interacting diffusion on a unimodular Galton-Watson (UGW) tree (see Definition 2.16) [60]. Before this, there wasn't even a conjecture as to what such a characterization might look like. We derive the local equation for IPS on the regular tree $\mathbb{T}_d$ and prove that it describes the marginal dynamics of the IPS at the root and its neighbors. Our characterization of the local equation holds for a more general class of "strongly symmetric" initial conditions (see Definition 6.11) compared to the i.i.d. initial conditions studied in [60], which has key applications in the study of flows and stationary solutions to the local equation. This more general result requires a number of new results discussed later in this introduction and in Section 6.2. We finish by constructing a Markovianized version of the local equation, presenting some results regarding stationary solutions to this *Markov approximation* and running some numerical simulations comparing this new approximation to the more standard mean-field and pairwise approximations (introduced in the next section).

1.1 The Mean Field Approximation

The classical approach to this problem applies in the Markov case when G_n is the complete graph of n vertices (labeled 1 to n) and the IPS has exchangeable particles and is *weakly interacting*, meaning that interactions between pairs of particles are of strength $O(n^{-1})$. The assumption that the IPS is weakly interacting often manifests as an assumption that the jump rate functions of X_v^n only depend on $X_v^n(t-)$ and the empirical measure $\pi^n := \frac{1}{n}\sum_{u=1}^n \delta_{X_u^n(t-)}$, where δ is the Dirac measure, along with some Lipschitz-type bound on the jump rate functions (e.g. [21, 73] also see [8, 79]). It should be noted that many interacting particle systems of interest are not weakly interacting, so mean-field limits are taken by modifying the dynamics of such processes. See Section 7.1 for a more detailed discussion.

1. **Propagation of chaos:** there exists a measure μ such that for any $k \in \mathbb{N}$, $\text{Law}(X_1^n, \dots, X_k^n)$ converges in some sense to $\otimes_{i=1}^k \mu$. That is, any finite collection of particles is asympotically independent as $n \to \infty$.

2. **Law of large numbers:** the empirical measure π^n converges in some sense to μ.

3. **The limit measure μ:** μ can be described as the law of an integer-valued pure jump process X whose transitions are described by the same jump rate functions as X_v^n (for any $v \in V_n$) but in which π^n is replaced by $\text{Law}(X) = \mu$.

Formally speaking the idea behind the mean-field model is quite simple. When n is large, the particles $\{X_v^n\}_{v \in V_n}$ are exchangeable and almost independent. It can be shown that the dependence between particles is weak enough that π^n satisfies a law of large numbers limit. That is, when n is large, π^n is approximately equal to the law of X_1^n. Thus, the interaction between a typical particle X_v^n and the empirical distribution π^n can be approximated by an interaction between X_v^n and $\text{Law}(X_1^n)$. In the finite state case, μ can be described as the solution to a non-linear ODE, which is easily solved (see (7.1.2)). Therefore, the mean-field approximation, in which X_v^n are assumed to be independent with law μ, is heavily used due to its simplicity. For more information, see [73, 57, 24, 25, 85, 57].

Now consider the case in which G_n is no longer necessarily complete, and π^n is replaced by the local empirical measures $\pi^{n,v} := \frac{1}{|\mathcal{N}_v|}\sum_{\{u,v\}\in E_n}\delta_{X_u^n(t-)}$, $v \in V_n$. There has been a burgeoning interest in the question of under what conditions on G_n does the mean field limit, or some minor variation thereof, still hold (e.g. [29, 74, 7, 27, 10]). Although work in this area is still ongoing, it appears that as long as the sequence of graphs G_n, $n \in \mathbb{N}$, is sufficiently dense, the mean field limit still holds. To our knowledge, most work in this area is focused in the diffusion setting, and we only know of one paper, [22], that looks at such limits for an IPS (specifically, the supermarket model: a queuing model in queues with unit exponential service times are connected in a network; jobs arrive at each queue according to a rate lambda Poisson process and enter the shortest of the arrival queue and $d-1$ randomly chosen neighboring queues).

Even if the mean-field approximation holds asymptotically as $n \to \infty$, the mean-field approximation is not always the best approximation to work with in practice [42]. For example, n may be too small, the graph G_n may be insufficiently dense, or interactions between neighboring particles may not be sufficiently weak. As a result, neighboring particles may be strongly correlated which would result in the mean-field approximation yielding poor results. One fairly standard refinement of the mean-field approximation is the *pairwise approximation*. The pairwise approximation applies the philosophy of the mean-field approximation, but to pairs of neighboring particles rather than individual particles. The assumption of independence is replaced with the following assumption of conditional independence: if $\{v_1, v_2, v_3\}$ forms a path in G_n, then for any $t \in \mathbb{R}_+$, $X^n_{v_1}(t)$ is conditionally independent of $X^n_{v_3}(t)$ given $X^n_{v_2}(t)$. Just as for the mean-field approximation, the assumption of conditional independence can be used to derive a non-linear ODE that approximates the distribution of a typical pair of neighboring particles of G_n (see (7.1.7)). Unlike the mean-field approximation, the pairwise approximation takes pairwise correlations between neighboring particles into account. See [39, 40, 68, 52, 91, 84] for some examples of derivations and implementations of the pairwise approximation for specific IPS.

1.2 The Sparse Graph Case

Now consider the case in which the (expected) average degree of G_n is uniformly bounded with respect to n. In this case, the mean field approximation notably fails as interactions between neighboring particles remain significant for arbitrarily large n. In fact, as we demonstrate within this thesis, the specific topology of the graph plays a large role in determining the distribution of the resulting IPS, so limiting processes such as the mean-field approximation and the pairwise approximation, which are indifferent to the topology of the graph sequence $\{G_n\}_{n\in\mathbb{N}}$, will generally fail. This brings us back to the four questions stated in the beginning of the chapter. When the graph is complete, questions 1 and 4 hold by well-posedness of the IPS on finite graphs (which is easily proven) and by the well-posedness of the mean field limit (which is a classical result [73]). Question 2 is addressed by propagation of chaos while Question 3 is a statement of the law of large numbers result. Thus, all four questions are addressed. Even in the dense case, it can be surmised that for sufficiently dense graphs, similar results hold. However, in the sparse case, the natural limiting object of X^{G_n} is X^G where $G_n \to G$ in some sense. If G_n is an increasing sequence of graphs, then X^G will be an infinite particle process. This is where the local equation becomes useful. When $G = \mathbb{T}_d$ is the infinite d-regular tree, the local equation yields the $d+1$-particle process $X^G_{\mathrm{cl}_\varnothing}$.

1.2.1 Well-Posedness of IPS on Locally Finite Graphs

A key prerequisite to proving the convergence of $X^{G_n} \to X^G$ is to first establish that X^G is a well-defined quantity. Indeed, although several recent works studying IPS on random graphs provide intuitive descriptions of the IPS, there appears to be no general result that rigorously establishes well-posedness of even Markovian IPS on locally finite graphs of unbounded degree such as a standard Galton Watson tree with a Poisson offspring distribution. While well-posedness of IPS on finite graphs is standard under our assumptions, on infinite graphs the issue is more subtle and well-posedness can in fact fail to hold even for simple Markovian IPS, as illustrated by the simple example in Section 4.3. Until recently, previous well-posedness results for IPS on infinite graphs have been almost exclusively focused on graphs with uniformly bounded maximum degree. For example, on lattices, an analytical proof of well-posedness of a large class of Feller IPS via examination of their semigroups can be found in the seminal paper of Liggett [64] (see also [65]). Another approach to well-posedness involves a standard Picard iteration argument applied to the (jump) stochastic differential equation (SDE) representation of the IPS dynamics. This works when the jump rates of any individual particle satisfy a strong Lipschitz property, that is, when the jump rate functions of every particle are uniformly Lipschitz with respect to the state (or trajectory, in the non-Markovian setting) of each of the neighboring particles and the sum of the (single-neighbor) Lipschitz coefficients associated with each particle are uniformly bounded (see also [30] for a slightly weaker averaged version of this Lipschitz condition). However, for even standard Markovian IPS such as the majority process or the contact process (see Example 3.15), the Lipschitz constant of the jump rate function with respect to the states (or trajectories and marks) of each neighboring particle does not decrease, but rather increases with the degree of the vertex of the particle, or at best remains constant. In particular, strong Lipschitz continuity of the jump rates (in the sense described above) fails on infinite (random) graphs that have unbounded maximum degree such as Galton-Watson (GW) trees with Poisson offspring distribution. In Theorem 4.2 we establish *strong* well-posedness of a general class of possibly non-Markovian IPS on (almost surely) finitely dissociable graphs, which includes the GW trees mentioned above. Lastly, there is a probabilistic proof of well-posedness of IPS with nearest-neighbor interactions using percolation arguments can be found in the classical work of Harris [46]. The latter argument is easily extended to locally interacting IPS on translation invariant graphs, but crucially relies on the graph having uniformly bounded maximum degree. To the best of our knowledge, the only work other than [36] that uses percolation methods to prove well-posedness of a (jump) IPS on a graph with unbounded maximal degree appears to be the recent work of Gantert and Schmidt [38], which establishes well-posedness of the simple exclusion process on a GW tree whose offspring distribution has finite mean by crucially exploiting the special structure of the exclusion process to reduce the problem to the study of a standard bond percolation problem.

In this thesis, we present the well-posedness results and proofs of [36]. We prove strong well-posedness under three sets of assumptions of increasing generality. At the highest layer of generality, we introduce the notion of spatial localization of the IPS dynamics (see Definition 4.4), which also plays a key role in the proof of the hydrodynamic limit. Roughly speaking, an IPS with interaction graph G is said to be spatially localized if the IPS dynamics on any finite subset U of the graph G coincides with the marginal dynamics on U of the same IPS model on (the induced subgraph of G on) a finite random subset that contains U. Leveraging strong well-posedness of the IPS on finite graphs, we show that the IPS is strongly well-posed on any graph G that spatially localizes the IPS SDE (see Proposition 4.10). At the next layer of generality, we introduce a notion of finitely dissociable graphs which is defined in terms of a degree-dependent site percolation that is closely related to the percolation introduced in [46] (see Definition 4.15). Then, under mild conditions on the jump rates, we show that the IPS SDE is spatially localized by any finitely dissociable graph (see Proposition 4.22). The proof entails the analysis of so-called causal chains that capture the propagation of influence of the dynamics at a vertex in the IPS. To get to the lowest layer of generality, we show in Section 4.2.3 that finitely dissociable graphs include all bounded degree graphs (this is a classical percolation result), GW trees whose offspring distributions have finite first moments and consequently, UGW trees whose offspring distributions have finite second moments. Our result does not subsume that of [38], but is applicable to a wide class of possibly non-Markovian models and does not rely on specific features of the IPS.

1.2.2 Graph Limits and Limits of IPS

A natural notion of convergence of graphs, first introduced by Benjamini and Schramm [9] and formalized by Aldous and Steele [3] and Aldous and Lyons [2], is local convergence, which applies to sequences of isomorphism classes of connected, rooted graphs. Formally speaking, the local topology bases the distance between the isomorphism classes of two rooted graphs G_1 and G_2 on the size of the largest subgraphs of $H_1 \subset G_1$ and $H_2 \subset G_2$ containing the roots such that H_1 and H_2 are isomorphic. Many common random graphs have well known and well understood local weak limits (see Section 2.1.3). But more importantly, the local topology can be extended to isomorphism classes of *marked graphs* (sometimes called *networks*).

One weakness of the local topology is its inability to account for hydrodynamic limits such as the convergence of the empirical measure of a graphical model (e.g. [61, Proposition 7.7]). This is because empirical quantities may be influenced by particles that are arbitrarily far from the root. This concern can be alleviated by additionally establishing an asymptotic spatial correlation decay property (also called the second moment method) [61, 69, 88, 36], in which case the graph sequence can be said to converge in probability in the local weak sense (see Definition 2.2 and Lemma 2.5) from which the convergence of (for example) the empirical measures of a sequence of graphical models immediately follows (see [61, Lemma 2.7] and Lemma 2.4). Note that similar results can be achieved through the use of concentration inequalities [75].

For more than a decade, it has been known that graphical models such as the Ising model [69] and other factor models [31] respect the local topology. More recently, there has been some investigation into interactions between various interacting particle systems and the local topology [49, 11, 71].

Of particular interest are two recent papers by Lacker et al. and Oliveira et al. regarding the local convergence and hydrodynamic limits of trajectories interacting diffusions [61, 75]. The paper by Oliveira et al. applies to interacting diffusions with (possibly heterogeneous) pairwise interactions and with certain restrictions on the underlying graphs. The paper by Lacker et al. holds for both interacting diffusions and discrete time processes on all graphs with symmetric interactions. Both papers fundamentally rely on an analogue of the strong Lipschitz property described in the previous section for interacting diffusions. As mentioned above, for jump processes, even simple examples require weaker assumptions on the jump rates, and well-posedness need not hold for IPS on all graphs. In addition, our framework also allows for heterogeneities, which are of relevance in many applications, and the nature of correlation decay established is different (see the discussion in Section 5.2.1 for an elaboration of this point). Thus, several new tools are required for the analysis in comparison with the diffusion setting. Finally, even though the class of IPS we focus on already covers a large class of models (see the examples in Section 3.2), many auxiliary results are established under still weaker assumptions so as to facilitate further extensions of the main results to IPS described by more general Poisson-driven SDE, in particular those with simultaneous jumps (see Remarks 4.9, 4.11, 5.12 and 5.13).

There is one more problem with working with isomorphism classes. Often the underlying graph of an IPS is merely a convenient means of encoding the interactions between various particles of the IPS, in which case the isomorphism class of the marked graph representing the IPS is not an object of interest. For example, if $G = \mathbb{T}_d$, v is a non-root vertex of $\mathbb{T}_d$ and (G, z^n), $n \in \mathbb{N}$, (G, z), is a collection of graphs with vertex marks in $\mathcal{Z}$ such that $\langle G, z^n \rangle \to \langle G, z \rangle$ in $\mathscr{G}_*[\{1\}, \mathcal{Z}]$, then it is not necessarily true that $z^n_v \to z$. Thus, we need something more in order to establish componentwise convergence IPS defined on common vertex sets. Furthermore, we later establish that IPS with sufficiently heterogeneous dynamics may not even have well-defined isomorphism classes (see Remark 3.7). And yet, we would like a way to investigate the convergence of even such heterogeneous IPS. For this reason, we introduce a space of marked graphs that can also be equipped with a Polish topology, but in which convergence is arguably more natural than in the space of isomorphism classes. In addition, this space is highly compatible with the local topology which makes it easy to translate convergence results in this topology to analogous convergence results in the local topology. This space was first defined in [36], but we include several new lemmas describing properties of this space.

This thesis replicates and extends the results of [36], which establish local weak convergence and convergence in probability in the local weak sense of IPS. We further extend these results by reintroducing a space of marked graphs first introduced in [36] and proving that the trajectories of IPS additionally converge in this new space.

1.2.3 Autonomous Characterization of the Dynamics

As mentioned at the beginning of the chapter, Lacker et al. recently derived local equations that they use to autonomously characterize the marginal dynamics at the root and its neighbors of a UGW tree [60], thereby satisfying Question 4. Assuming i.i.d. initial conditions, they proved that the local equation is well-posed and that the solution to the local equation is equal in distribution to the marginal at the root and its neighbors of an interacting diffusion defined on the UGW tree.

Here we explain the basic intuition behind this result. To keep the explanation simple, assume that the graph $G = \mathbb{T}_2$ is the infinite path (or equivalently, the infinite cycle, the 1-dimensional lattice or the regular 2-tree). Assume the vertex set $V = \mathbb{Z}$, the edge set $E = \{\{v, v+1\} : v \in \mathbb{Z}\}$ and the root $\varnothing = 0$. Let X^G represent an IPS taking values in $\mathcal{X} := \{0, 1\}$, and that for any $v \in \mathbb{Z}$ and $t > 0$,

$$X_v^G(t) \mapsto 1 - X_v^G(t) \text{ at rate } \overline{r}(X_{\{v-1,v,v+1\}}^G(t)),$$

where $\overline{r} : \mathcal{X}^3 \to \mathbb{R}_+$ is a measurable function. For a concrete example of such a function, imagine that X^G is the contact process with infection rate $\lambda > 0$ so that,

$$\overline{r}(a_{\{0,1,2\}}) = \begin{cases} 1 & \text{if } a_1 = 1, \\ \lambda(a_0 + a_2) & \text{otherwise.} \end{cases}$$

So then how do we answer Question 4? As mentioned earlier, we provide an *autonomous characterization of Law*$(X_{\{-1,0,1\}}^G)$. How do we do this?

Step 1. Projection: describe the dynamics of the particles at ± 1 in terms of $X_{\{-1,0,1\}}^G$.

Consider the particle at 1, X_1^G, of the contact process at time t. To understand the rate at which particle 1 will be infected (assuming it is susceptible at time t), one must also understand whether or not particles 0 and 2 are infected at time t. How can we remove this dependence on particle 2? As we show in Section 9, we can simply average it out. That is, if (conditional on the full history up to time t of particles -1, 0 and 1) the probability that particle 2 is infected is given by $p \in (0,1)$, then the rate at which X_1^G will get infected is simply given by,

$$p\overline{r}(X_0^G(t), X_1^G(t), 1) + (1-p)\overline{r}(X_0^G(t), X_1^G(t), 0).$$

In general, the same principle applies. We can (up to certain regularity conditions) *average out* the impact of the particle at 2 without changing the law of the process. This *average* is taken with respect to the regular conditional probability $\mathrm{Law}(X_2^G[0,t)|X_{\{-1,0,1\}}^G[0,t))$. Specifically, we can replace $t \mapsto \overline{r}(X_{\{-1,0,1\}}^G(t-))$ (we use $t-$ instead of t for measurability reasons) by,

$$t \mapsto \rho^1(t, X^G) := \mathbb{E}\left[\overline{r}(X_{\{0,1,2\}}^G(t-))|X_{\{-1,0,1\}}^G[0,t)\right] \text{ a.s.,}$$

without changing the law of X_1^G. We may define ρ^{-1} analogously to give a characterization of the dynamics of X_{-1}^G that also depends only on $X_{\{-1,0,1\}}^G$.

Although this is a nice start, this introduces a problem. Once known, the functions $\rho^{\pm 1}$ only take $X_{\{-1,0,1\}}^G$ as inputs and therefore provide a way to describe the dynamics of $X_{\{-1,0,1\}}^G$ without having to know X_v^G for $|v| > 1$. However, to derive the functions $\rho^{\pm 1}$ requires that we already know $\mathrm{Law}(X_{\{-2,-1,0,1,2\}}^G)$, which defeats the purpose of finding a characterization of $\mathrm{Law}(X_{\{-1,0,1\}}^G)$.

Step 2. Conditional Structure: to resolve this problem, we now examine the conditional structure of X^G.

Again, let us restrict our consideration to the particle at 1. Then in step 1, we claimed that we could average out the impact of X_2^G on X_1^G by averaging it out with respect to the conditional measure $\mathrm{Law}(X_2^G[0,t)|X_{\{-1,0,1\}}^G[0,t))$. In fact, in Proposition 6.10, we show that

$$\mathrm{Law}(X_2^G[0,t)|X_{\{-1,0,1\}}^G[0,t)) = \mathrm{Law}(X_2^G[0,t)|X_{\{0,1\}}^G[0,t)) \text{ a.s..}$$

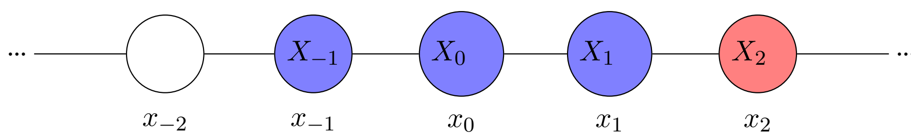
In fact, we show that X^G satisfies a certain *second order Markov random field* (MRF) structure that later becomes relevant in proving that the local equation characterizes the marginals of IPS for all strongly symmetric (instead of i.i.d.) initial data.

Step 3. Redefining ρ^1: we can now write $\rho^1(t,\cdot)$ in terms of $\mathrm{Law}(X^G_{\{-1,0,1\}}[0,t))$.

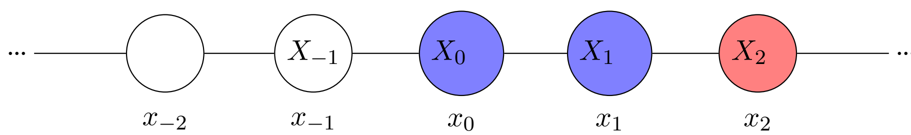
To do this, note that it suffices to understand the regular conditional probability $\mathrm{Law}(X^G_2[0,t)|X^G_{\{-1,0,1\}}[0,t))$ in terms of $\mathrm{Law}(X^G_{\{-1,0,1\}}[0,t))$. Let $X = X^G[0,t)$ and fix any x in the support of X. Then we can visually represent the regular conditional probability,

$$\mathrm{Law}(X_2|X_{\{-1,0,1\}} = x_{\{-1,0,1\}}),$$

as follows:



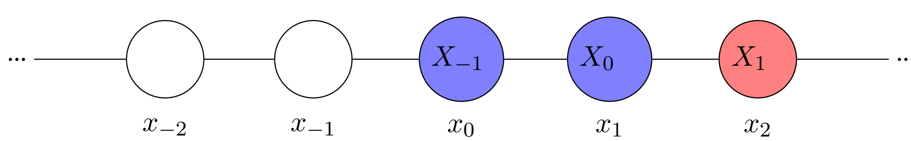
Above, the red node represents the law we are interested in while the blue nodes denote what we are conditioning on. Then by step 2, we do not have to condition on X_{-1}:



Thus,

$$\mathrm{Law}(X_2|X_{\{-1,0,1\}} = x_{\{-1,0,1\}}) = \mathrm{Law}(X_2|X_{\{0,1\}} = x_{\{0,1\}}).$$

Next, it is easily proven that X is translation-invariant in the sense that $\mathrm{Law}(X_v : v \in \mathbb{Z}) = \mathrm{Law}(X_{v-1} : v \in \mathbb{Z})$. Visually, this implies that we may shift the graph to the right:



Thus,

$$\mathrm{Law}(X_2|X_{\{-1,0,1\}} = x_{\{-1,0,1\}}) = \mathrm{Law}(X_1|X_{\{-1,0\}} = x_{\{0,1\}}),$$

or

$$\mathrm{Law}(X^G_2[0,t)|X^G_{\{-1,0,1\}}[0,t) = x_{\{-1,0,1\}}) = \mathrm{Law}(X^G_1[0,t)|X^G_{\{-1,0\}}[0,t) = x_{\{0,1\}}).$$

By defining ρ^1 as an average taken with respect to the regular conditional probability above, it is possible to explicitly define $\rho^1(t,\cdot)$ in terms of $\mathrm{Law}(X^G_{\{-1,0,1\}}[0,t))$.

Step 4. The Local Equation: we can now define the local equation.

Now consider a different IPS Y on the three particles $\{-1,0,1\}$ whose dynamics at ±1 are described by $\gamma^{\pm1}$ and whose dynamics at 0 are described by $t \mapsto \overline{r}(Y_{\{-1,0,1\}}(t-))$. Suppose the definition of $\gamma^{\pm1}$ exactly matches the definition of $\rho^{\pm1}$ except that $\mathrm{Law}(X^G_{\{-1,0,1\}})$ is replaced by $\mathrm{Law}(Y)$. Then Y can be described as the solution to a non-Markovian, law-dependent SDE called the local equation.

Step 5. Characterization: why does the local equation capture $\mathrm{Law}(X^G_{\{-1,0,1\}})$?

Let Z solve the local equation, except that $\gamma^{\pm1}$ are replaced by $\rho^{\pm1}$. Then by steps 1 and 3, it follows that $\mathrm{Law}(Z) = \mathrm{Law}(X^G_{\{-1,0,1\}})$. But then it follows from step 4 that $\gamma^{\pm1} = \rho^{\pm1}$, so Z is also a solution to the local equation. Therefore, if the local equation is well-posed (so that Z is the unique in law solution to the local equation), then the local equation fully characterizes $\mathrm{Law}(X^G_{\{-1,0,1\}})$.

1.2.4 Additional Results

We finish by describing some additional results we were able to obtain.

Result 1. The IPS is a 2nd-order Markov random field (Proposition 6.10):

A Markov random field (MRF) is a conditional structure satisfied by certain graphs with random vertex marks, and may be thought of as the generalization of a Markov chain to a process whose components are indexed by the vertices of a graph. When the graph is infinite, an MRF may be categorized as local or global (see [43] and Definition 6.5).

Although the use of Markov random fields to characterize diffusions (expressed as solutions to SDEs) is a topic of broad interest (see [59] and the references therein), the only paper (to our knowledge) that addresses the analogous result for general IPS is [37], the results of which are stated (with lesser generality) in Section 6.2 and proved in Sections 8.1, C.1 and C.2.

Lacker et al. uses a Clifford-Hammersley based proof and some clever graph theory to establish that interacting diffusions whose initial conditions form 2-local MRFs have 2-local MRF trajectories on both finite and infinite time intervals [59]. They assume that all finite dimensional marginals of the initial conditions are absolutely continuous with respect to a product measure. However, as observed in [59, Remark 3.6], this result is insufficient to establish the local equation. [60, Remark 3.16] states that MRF trajectories would need to be global. To deal with this, [60, Proposition 3.17] establishes specific conditional independence properties of the trajectories of the interacting diffusion that hold when the initial conditions are i.i.d..

We introduce the concept of a *semi-global* Markov random field (SGMRF), which generalizes the notion of tree-indexed Markov chains [41, Chapter 12] (see Lemma 6.8). It turns out that this notion of Markov random field is sufficient to establish the local equation for non-i.i.d. initial data. We then show that if the initial data forms a 2-SGMRF, then the trajectories of the IPS also form a 2-SGMRF on both finite and infinite time intervals. Crucially, we do not require the initial data (or any of its finite dimensional marginals) to be absolutely continuous with respect to any product measure, which is a key requirement to study stationary solutions to the local equation.

There are two intermediate steps in this process which may be of independent interest. First, we establish an infinite-dimensional change of measure theorem for IPS. We prove the theorem in very general terms even allowing for IPS defined on non-locally finite graphs, with partial results in the case that IPS is not well defined. Second, under the same conditions we establish a duality between the IPS and a point process, which provides a way to lift standard results from point process theory to the theory of IPS. The main challenge in both proofs when the underlying graph is infinite is that the resulting point process will typically be *explosive*. It is therefore necessary to first establish the respective results for finite graphs and then extend them to infinite graphs.

Result 2. A Tiling Proposition (Proposition 6.12):

Let $H = (\{-1,0,1\},\{\{-1,0\},\{0,1\}\},0) \subset G = \mathbb{T}_2$ be the path from -1 to 1 in G. The tiling proposition follows from the observation that for any two random marked graphs (G,z^1) and (G,z^2) with marks in the Polish space $\mathcal{Z}$ forming an automorphism invariant, 2-SGMRF, $z^1_{\{-1,0,1\}} \overset{(\mathrm{d})}{=} z^2_{\{-1,0,1\}}$ if and only if $(G,z^1) \overset{(\mathrm{d})}{=} (G,z^2)$. Furthermore, there is a bijection between the set of automorphism invariant, 2-SGMRF (with respect to G) probability measures $\eta \in \mathcal{P}(\mathcal{Z}^{\mathbb{Z}})$ and the set of *strongly symmetric* (see Definition 6.11) probability measures in $\mathcal{P}(\mathcal{Z}^{\{-1,0,1\}})$ (see Proposition 6.12). This bijection is useful because strong symmetry is a very simple property of probability measures to verify. Given any strongly symmetric probability measure η, we denote by

$\overline{\eta}$ the *tiled extension* of η, that is, $\overline{\eta}$ is the unique automorphism invariant, 2-SGMRF measure such that the marginal distribution of $\overline{\eta}$ at the vertices $\{-1, 0, 1\}$ is η.

To our knowledge, the tiling proposition is a novel result, and we provide a constructive proof of the tiling proposition in Section 8.3. Thus, given any strongly symmetric probability measure η, there exists a unique 2-SGMRF, has powerful applications to the analysis of the local equation. The tiling proposition can be used to extract the law of the IPS given the law of the solution to the local equation (Theorem 6.17). We also use the tiling proposition to extract information about how the flows generated by the IPS and by the solution to the local equation interact and to understand how the local equation behaves at stationarity.

Most importantly, the tiling proposition of provides a map between strongly symmetric solutions to the local equation and IPS. This plays a pivotal role in our ability to extend the characterization of the local equation from i.i.d. initial data to strongly symmetric initial data. We also show in Theorem 6.17 that the solution to the local equation for strongly symmetric initial data remains strongly symmetric. This allows us to investigate properties of flows generated by the IPS X^G and by the corresponding solution to the local equation. It also allows us to investigate limiting (in time) solutions to the local equation as well as stationary solutions to the local equation.

Result 3. The Markov Local Equation:

The local equation is defined in terms of expectations conditioned on path space. In practice, such conditional expectations are extremely difficult to compute directly. Therefore, we introduce a *Markov approximation* of the local equation and derive its forward equation, which is a simple ODE. We show that stationary solutions to this ODE capture the automorphism invariant, 2-SGMRF (with respect to state space rather than path space) stationary distributions of Markovian IPS. We present some simple numerical comparisons of the Markov approximation to the pairwise and mean-field approximations.

To our knowledge, this is the first general investigation of stationary solutions to the local equation, though [62] does analyze stationary solutions to the local equation for a specific class of reversible interacting diffusions with 1-SGMRF stationary solutions (although they identify these solutions as Gibbs measures satisfying the Markov property on trees, this is equivalent to the 1-SGMRF property by Remark 6.9).

1.3 Outline of this Thesis

Part I generally deals with preliminary matters such as the introduction and notation in this chapter. Chapter 2 discusses notions of graph convergence in greater detail. We discuss the local topology, convergence in probability in the local weak sense and finally we define the novel space $\widehat{\mathscr{G}}_*[\cdot]$ (first introduced in [36]) and state its properties. In Chapter 3, we define the model and provide examples of IPS within our framework.

Part II deals with questions of well-posedness and graph convergence. In Chapter 4, we investigate conditions under which the IPS X^G is well-defined (or equivalently, when the SDE defining X^G is well-posed). We also construct a counterexample demonstrating that well-posedness can fail even for IPS with bounded jump rates. In Chapter 5, we state and prove all of our limit theorems.

Part III focuses on the autonomous characterization of X^G when G is the d-regular tree. In Chapter 6, we state all of the results of Part III which we prove in Chapters 7 to 10. Here we derive the local equation, prove that it characterizes the marginal dynamics of X^G, study flow properties and properties of stationary solutions to the local equation and finally we examine some numerical results obtained from the Markov approximation to the local equation.

In Appendix A, we prove all results from Chapter 2 relating to the space $\widehat{\mathscr{G}}_*[\cdot]$. In Appendix B, we prove the projection result (Proposition 9.2), which we informally discuss in step 1 of the derivation of the local equation in the previous section. In Appendix C, we state and prove a few useful results about the properties of conditional independence and MRFs. Finally, in Appendix D we prove a series of minor miscellaneous results.

1.4 Notation and Conventions

Most of the notation we use in this thesis can be found in Table 1. Given any two real numbers $a, b \in \mathbb{R}$, $a \wedge b = \min\{a, b\}$ and $a \vee b = \max\{a, b\}$.

Graph Notation: A graph $G = (V, E)$ equipped with a distinguished vertex $\varnothing \in V$, denoted the root, is called a rooted graph, and denoted by $(G, \varnothing) := (V, E, \varnothing)$, although when the root is clear from context, we will simply write G instead of $(G, \varnothing)$. We may additionally use V_G, E_G and $\varnothing_G$ to denote the respective vertex set, edge set and root of G. For $U \subseteq V$, we denote by $G[U]$ the induced subgraph of G on U, that is, $G[U] = (U, E[U])$ where $E[U] = E \cap \{\{x,y\} : x, y \in U\}$. For $u, v \in V$, a path between u and v in G is defined to be a sequence of vertices $\Gamma = (u = v_0, v_1, \ldots, v_{n-1}, v_n = v)$ such that for all $i \in \{1, \ldots, n\}$, $\{v_{i-1}, v_i\} \in E$ and $v_i \neq v_j$ whenever $i \neq j$ except possibly when $\{i, j\} = \{n, 0\}$, in which case the path is said to be a cycle. A graph is said to be acylic if it has no cycles. The length of the path, denoted $|\Gamma|$, is the number of edges in the path. We let $d_G(u, v)$ denote the usual graph distance, which is the length of the shortest path between u and v in G, where $d_G(u, v) = \infty$ if no such path exists. If $d_G(u, v) < \infty$ for all $u, v \in V$, then G is said to be connected. For any vertex $v \in V$, $\mathscr{C}_v(G)$ is the largest connected subgraph of G containing v equipped with v as its root. Given three disjoint subsets $A, B, S \subset V$, and a positive integer α, S is said to α-separate A and B if any path between a vertex in A and a vertex in B contains a consecutive sequence of α vertices in S. When G is a finite rooted graph, its radius is the maximal distance from any vertex to the root. Then $\widehat{\mathscr{G}}_{*,1}$ is the space of rooted graphs of radius 1.

Given an undirected graph $G = (V, E)$ with vertex set V and edge set E, for $U \subset V$, let $\mathcal{N}_U = \mathcal{N}_U(G) := \{v \in V : d_G(v, U) := \min_{u \in U} d_G(u, v) = 1\}$ denote the neighbors of U in G and let $\mathrm{cl}_U = \mathrm{cl}_U(G) := U \cup \mathcal{N}_U$ denote its closure. We also define the α-neighborhood $\mathcal{N}_U^\alpha := \mathcal{N}_U^\alpha(G) := \{v \in V \setminus U : d_G(v, U) \leq \alpha\}$. For any $v \in V$, set $\mathcal{N}_v := \mathcal{N}_{\{v\}}$, $\mathrm{cl}_v := \mathrm{cl}_{\{v\}}$ and $\mathcal{N}_v^\alpha := \mathcal{N}_{\{v\}}^\alpha$. We define $\Lambda_G := \{U \subseteq V : |U| < \infty\}$ to be the set of finite subsets of the vertices in G. Recall that the degree of a vertex v is equal to $|\mathcal{N}_v|$, where for any set A, $|A|$ denotes its cardinality. The graph G is said to be locally finite if each of its vertices has finite degree. We always assume graphs are simple (i.e., they do not have self-loops or multi-edges) and locally finite.

Let $G_i = (V_i, E_i), i = 1, 2$, be (unrooted) graphs. A mapping $\varphi : V_1 \to V_2$ is said to be an unrooted *isomorphism* from G_1 to G_2 if it is a bijection and $e = \{u, v\} \in E_1$ if and only if $\varphi(e) := \{\varphi(u), \varphi(v)\} \in E_2$. Given roots $\varnothing_i \in V_i, i = 1, 2$, φ is a rooted isomorphism from the rooted graph $(G_1, \varnothing_1)$ to the rooted graph $(G_2, \varnothing_2)$ if, in addition, $\varphi(\varnothing_1) = \varnothing_2$. Recall that when denoting the rooted graph, we often omit the explicit dependence on the root. Given *locally finite* rooted graphs G_1 and G_2, $I(G_1, G_2)$ denotes the collection of isomorphisms from G_1 to G_2. If $I(G_1, G_2)$ is non-empty, then G_1 and G_2 are said to be isomorphic, which is denoted $G_1 \cong G_2$. For any graph G, an automorphism of G is an isomorphism $\varphi \in I(G, G)$. An emphunrooted automorphism is an automorphism φ for which $\varphi(\varnothing_G)$ is not necessarily equal to $\varnothing_G$.

A rooted tree is a connected, rooted acyclic graph $\mathcal{T} = (V, E, \varnothing)$. Any pair of vertices $u, v \in V$

has a unique path connecting them. A vertex $v \in V$ is said to be in the kth generation of $\mathcal{T}$ if $d_{\mathcal{T}}(\varnothing, v) = k$. For $v \in V \setminus \{\varnothing\}$, the parent of v, denoted $\pi_v(\mathcal{T})$, is the unique neighbor u of v such that $d_{\mathcal{T}}(\varnothing, u) < d_{\mathcal{T}}(\varnothing, v)$. We also set $c_v(\mathcal{T}) := \mathcal{N}_v(\mathcal{T}) \setminus \pi_v(\mathcal{T})$ to be the children of v. We will write π_v and c_v when $\mathcal{T}$ is clear from context. For any vertex set $U \subseteq V$, we will let $c_U(\mathcal{T}) = \bigcup_{u \in U} c_u(\mathcal{T})$ In addition, given $v, w \in V$, w is said to be a descendant of v if there exists $n \in \mathbb{N}$ and a path $(v = u_0, u_1, \ldots, u_n = w)$ in $\mathcal{T}$ such that for every $i = 1, \ldots, n$, $u_i = c_{u_{i-1}}(\mathcal{T})$. The leaves of G are given by the set $\mathcal{L}(G) := \{v \in V : c_v(G) = \emptyset\}$. We can define a partial order on the vertices in G by setting $u < v$ for $u, v \in V$ if u is descended from v.

Path Space Notation: Given any countable index set U and Polish space $\mathcal{Z}$, let $\mathcal{Z}^U = \{(z_v)_{v \in U} : z_v \in \mathcal{Z} \text{ for all } v \in U\}$ denote the corresponding configuration space, equipped with the product topology. For any $z \in \mathcal{Z}^V$, $z_U \in \mathcal{Z}^U$ denotes the restriction of z to $\mathcal{Z}^U$, that is, $z_U = (z_v)_{v \in U}$. Given two vertex sets V_1 and V_2, a map $\varphi : V_1 \to V_2$ and configurations $x \in \mathcal{Z}^{V_1}$, $y \in \mathcal{Z}^{V_2}$, we write $y = \varphi_* x$ to mean $y_v = x_{\varphi^{-1}(v)}$ for all $v \in V_2$.

We consider IPS with a countable state space $\mathcal{X}$, which we identify with a subset of the integers $\mathbb{Z}$ equipped with the discrete topology. Given $U \subseteq V$ and a (closed, half-open or open) interval $I \subseteq [0, \infty)$, let $\mathcal{D}(I, \mathcal{X}^U)$ denote the space of càdlàg functions from I to $\mathcal{X}^U$. Given $0 < t < \infty$, and $I = [0, t]$, $[0, t)$, $(-\infty, t]$ and $(-\infty, t)$, for conciseness denote $\mathcal{D}(I, \mathcal{X}^U)$ by $\mathcal{D}_t^U$, $\mathcal{D}_{t-}^U$, $\overline{\mathcal{D}}_t^U$ and $\overline{\mathcal{D}}_{t-}^U$ respectively. Also, set $\mathcal{D}^U := \mathcal{D}([0, \infty), \mathcal{X}^U)$ and $\overline{\mathcal{D}}^U := \mathcal{D}(\mathbb{R}, \mathcal{X}^U)$ and omit the superscript U from the notation when $|U| = 1$. If $x \in \mathcal{D}^U$ (respy. $\overline{\mathcal{D}}^U$) and $v \in U$, then $x_v(t)$ denotes the value of the vth component of x at time $t \geq 0$ (respy. $t \in \mathbb{R}$). For any $t \geq 0$, the restrictions of x to $[0, t]$ and $[0, t)$ are respectively denoted by $x[t] \in \mathcal{D}_t^U$ and $x[t) \in \mathcal{D}_{t-}^U$. Also, set $\Delta x(t) := x(t) - x(t-)$. For $0 \leq s < \infty$, an interval $I \subseteq \mathbb{R}$, finite $U \subseteq V$ and $x \in \mathcal{D}(I, \mathcal{X}^U)$, define the set of jump times as follows:

$$\operatorname{Disc}_s(x) := \{s' \in I \cap (0, s] : x(s') \neq x(s'-)\}, \tag{1.4.1}$$

and for $x \in \mathcal{D}^U$, set $\operatorname{Disc}(x) := \cup_{s \in (0, \infty)} \operatorname{Disc}_s(x)$. For $t \in [0, \infty]$, $\mathcal{D}_t^U$ is equipped with the product J1 topology, under which $\mathcal{D}_t^U$ is a Polish space [13, Theorems 12.2 and 16.3]. Note that for $\{x_n\} \subseteq \mathcal{D}^U$, $x_n \to x$ if for every $t \notin \operatorname{Disc}_{,}(x)$ $x_n[t] \to x[t]$. Lastly, fix $t \in \mathbb{R}_+$. Given any strictly increasing locally Lipschitz function $\psi : [0, t) \to [0, \infty)$ with a locally Lipschitz inverse (e.g., $\psi(s) := \frac{1}{t-s} - \frac{1}{t}$), the bijection $\mathcal{D}^U_{t-} \ni x \to x \circ \psi^{-1} \in \mathcal{D}^U$ induces the J1 topology on $\mathcal{D}^U_{t-}$. This can be used to show that $\mathcal{D}^U_{t-}$ is also Polish under the J1 topology. By a similar argument, $\overline{\mathcal{D}}, \overline{\mathcal{D}}_t$ and $\overline{\mathcal{D}}_{t-}$ are all Polish under the J1 topology as well.

Measure Notation: For any Polish space $\mathcal{Z}$, let $\mathcal{B}(\mathcal{Z})$ be the Borel σ-algebra on $\mathcal{Z}$, and let $\mathcal{P}(\mathcal{Z})$ be the Polish space of probability measures on $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}))$ equipped with the topology of weak convergence [13, Theorem 6.8]. Given $U \subseteq V$ and $\eta \in \mathcal{P}(\mathcal{Z}^V)$, let $\eta[U]$ be the marginal distribution of η restricted to $\mathcal{Z}^U$. Given any $\eta \in \mathcal{P}(\mathcal{Z})$ and a $\mathcal{Z}$-valued random element Z, we say $Z \sim \eta$ if the distribution of Z is given by η in which case we write $\mathrm{Law}(Z) := \eta$. We write $Y \overset{(\mathrm{d})}{=} Z$ if Y and Z have the same distribution, and write $Y_1 \perp\!\!\!\perp Y_2$ (respy. $Y_1 \perp\!\!\!\perp Y_2 | Y_3$) if Y_1 and Y_2 are independent (respy. conditionally independent given Y_3). For each $k \in \mathbb{N}$, let $\mathcal{Z}_k$ be a Polish space equipped with the measure $\eta_k \in \mathcal{P}(\mathcal{Z}_k)$. Then we will denote the product measure by $\otimes_{k \in \mathbb{N}} \eta_k$. If $\eta \in \mathcal{P}(\mathcal{D}(\mathbb{R}_+, \mathcal{Z}))$ for some Polish space $\mathcal{Z}$, then $\eta_t \in \mathcal{P}(\mathcal{D}([0,t], \mathcal{Z}))$ and $\eta_{t-} \in \mathcal{P}(\mathcal{D}([0,t), \mathcal{Z}))$ denote the restrictions of η to the respective Borel σ-algebras $\mathcal{B}(\mathcal{D}([0,t], \mathcal{Z}))$ and $\mathcal{B}(\mathcal{D}([0,t), \mathcal{Z}))$. Given a graph G, a $\mathcal{Z}^{V_G}$-random variable Z and a measure $\mathrm{Law}(Z) = \eta \in \mathcal{P}(\mathcal{Z}^{V_G})$, $\varphi_* \eta = \mathrm{Law}(\varphi_* Z)$. Then η is automorphism invariant if and only if $\varphi_* \eta = \eta$ for all automorphisms $\varphi \in I(G)$.

A filtration is said to satisfy the usual conditions if it is complete and right-continuous. Unless otherwise stated, all filtrations are assumed to be augmented so as to satisfy the usual conditions. Filtrations will typically be represented by the letters $\mathbb{F}, \mathbb{G}$ and $\mathbb{H}$, indexed by $\mathbb{R}_+$ or $[0,T]$ and for each $t \in \mathbb{R}_+$, the corresponding σ-algebras will be denoted by $\mathcal{F}_t, \mathcal{G}_t, \mathcal{H}_t$ respectively. Given a filtration $\mathbb{G} := \{\mathcal{G}_t\}_{t \in \mathbb{R}_+}$, recall that a simple sufficient condition for a process Z to be $\mathbb{G}$-*predictable* is that $t \mapsto Z_t$ is almost surely left-continuous and $\mathbb{G}$-adapted. Given two filtrations $\mathbb{F}$ and $\mathbb{G}$, as usual $\mathbb{F} \vee \mathbb{G} := (\mathcal{F}_t \vee \mathcal{G}_t)_{t \in \mathbb{R}_+}$ denotes the smallest filtration containing both $\mathbb{F}$ and $\mathbb{G}$. Given a random element ζ, we use $\mathcal{H}^\zeta$ to denote the completion of the σ-algebra generated by ζ (with respect to a probability measure that will be expressed explicitly if not clear from the context), and for any càdlàg stochastic process Z, we define $\mathbb{H}^Z := \{\mathcal{H}^Z_t\}_{t \in \mathbb{R}_+}$ to be the smallest filtration satisfying the usual conditions such that Z is adapted to $\mathbb{H}^Z$. For all processes Z considered in this paper, $\mathbb{H}^Z$ will be equal to the completion of the natural filtration of Z so that for all $t \in \mathbb{R}_+$, $\mathcal{H}^Z_t = \mathcal{H}^{Z[t]}$; see the discussion in [28, page 357] for more details.

Poisson Point Processes: Let $\mathcal{Z}$ be a Polish space equipped with its Borel σ-algebra and a metric $d_{\mathcal{Z}}$ that induces the Polish topology. On intervals $I \subseteq \mathbb{R}_+$ and on countable spaces (which will be assumed to have an implicit embedding in $\mathbb{N}$), this will be the standard absolute difference metric and on càdlàg spaces it will be the J1 metric. Finally, if $\mathcal{Z} := \prod_{i \in I} \mathcal{Z}_i$ for some finite index set $I \subseteq \mathbb{N}$ with metrics $d_{\mathcal{Z}_i}$, $i \in I$, then we set $d_{\mathcal{Z}}(z, z') = \sum_{i \in I} d_{\mathcal{Z}_i}(z_i, z_i')$. Let $\mathscr{N}(\mathcal{Z})$ be the space of locally finite, nonnegative integer-valued measures, that is, for any $p \in \mathscr{N}(\mathcal{Z})$ and $A \in \mathcal{B}(\mathcal{Z})$, $p(A) \in \mathbb{N} \cup \{\infty\}$ and $p(A) < \infty$ for every A that is bounded with respect to $d_{\mathcal{Z}}$. We equip $\mathscr{N}(\mathcal{Z})$ with the weak-hash topology, which then makes it a Polish space [28, page 2 and Proposition 9.1.IV(iii)]. Also, note that the map $\mathscr{N}(\mathcal{Z}) \ni p \mapsto p(A)$ is Borel-measurable for any $A \in \mathcal{B}(\mathcal{Z})$. A point process on $\mathcal{Z}$ is any $\mathscr{N}(\mathcal{Z})$-random element.

Let η be any nonnegative, locally finite Borel measure on $\mathcal{Z}$, that is, $\eta(B) < \infty$ for every bounded set $B \in \mathcal{B}(\mathcal{Z})$. A *Poisson point process* $\mathbf{N}$ on $\mathcal{Z}$ with *intensity measure* η is a point process on $\mathcal{Z}$ such that for any disjoint $A, B \in \mathcal{B}(\mathcal{Z})$, $\mathbf{N}(A)$ is Poisson distributed with expectation $\eta(A)$ and $\mathbf{N}(A) \perp\!\!\!\perp \mathbf{N}(B)$. A point process P is said to be $\mathbb{G}$-adapted if for every $t \geq 0$ and $A \in \mathcal{B}([0, t] \times \mathcal{Z})$, $P(A)$ is $\mathcal{G}_t$-measurable. We let $\mathbb{H}^P$ denote the minimal filtration satisfying the usual conditions such that P is $\mathbb{H}^P$-adapted. A Poisson point process $\mathbf{N}$ is said to be a $\mathbb{G}$-Poisson process if it is $\mathbb{G}$-adapted and for every $t \geq 0$ and $A \in \mathcal{B}((t, \infty) \times \mathcal{Z})$, $\mathbf{N}(A) \perp\!\!\!\perp \mathcal{G}_t$.

CHAPTER 2

Local Convergence and a Space of Marked Graphs

2.1 Convergence of Isomorphism Classes of Graphs

This section presents a brief introduction to the space of isomorphism classes of marked graphs and standard notions of convergence for isomorphism classes. In Section 2.1.1 we start with some basic definitions of the space of isomorphism classes, the local topology and local weak convergence. In Section 2.1.2 define the stronger notion of convergence: convergence in probability in the local weak sense. For more detailed information, see the excellent lecture notes by Bordenave [17] and book by Hofstad [88]. See also [61, Appendix A] which introduces two different metrics for the space of isomorphism classes of vertex marked graphs as well as an alternative characterization of convergence in probability in the local weak sense.

2.1.1 Local Weak Convergence

Let $\mathscr{G}_*$ be the space of isomorphism classes of connected, locally finite, rooted graphs. Then for any connected, locally finite rooted graph G, we let $\langle G \rangle \in \mathscr{G}_*$ denote the isomorphism class of G, namely $\langle G \rangle$ is the collection of connected locally finite rooted graphs isomorphic to G. Conversely, we refer to G as a *representative graph* of $\langle G \rangle$. Clearly, if $H \cong G$ then $H \in \langle G \rangle$. For each $m \in \mathbb{N}$, let $B_m(G)$ be the induced subgraph of G consisting of all vertices within (graph) distance m of the root. We equip $\mathscr{G}_*$ with the topology of *local convergence* in which $\{\langle G_n\rangle\}_{n\in\mathbb{N}} \subset \mathscr{G}_*$ converges to $\langle G\rangle \in \mathscr{G}_*$ if for every $m \in \mathbb{N}$, there exists $n_m < \infty$ such that $B_m(G_n) \cong B_m(G)$ for all $n \geq n_m$, $G_n \in \langle G_n\rangle$ and $G \in \langle G\rangle$.

Next, fix any two Polish spaces $\overline{\mathcal{Z}}$ and $\mathcal{Z}$ that represent the edge and vertex mark spaces, respectively, and consider a (not necessarily connected) marked rooted graph $G = (V, E, \varnothing, \overline{\vartheta}, \vartheta)$, where $(V, E, \varnothing)$ is a rooted graph, $\overline{\vartheta} \in \overline{\mathcal{Z}}^E$ and $\vartheta \in \mathcal{Z}^V$. Such a graph is referred to as a $[\overline{\mathcal{Z}}, \mathcal{Z}]$*-graph*. When G is random, it is called a $[\overline{\mathcal{Z}}, \mathcal{Z}]$*-random graph*. Let $[G_*] = (V, E, \varnothing)$ denote the marked rooted graph G with its marks removed, and let $[G]$ denote G with its root and marks removed. For $m \in \mathbb{N}$, let $B_m(G)$ be the induced marked rooted subgraph of G consisting of all vertices within (graph) distance m of the root, equipped with the same marks and root. We slightly abuse notation at times by allowing $B_m(G)$ to also denote the set of vertices within graph distance m of the root. We say the marked graphs $G := (V, E, \varnothing, \overline{\vartheta}, \vartheta)$ and $G' := (V', E', \varnothing', \overline{\vartheta}', \vartheta')$ with edge and vertex marks in $\overline{\mathcal{Z}}$ and $\mathcal{Z}$, respectively, are isomorphic, and write $G \cong G'$ if there exists an isomorphism $\varphi \in I([G_*], [G'_*])$ such that $\vartheta_v = \vartheta'_{\varphi(v)}$ and $\overline{\vartheta}_e = \overline{\vartheta}'_{\varphi(e)}$ for all $v \in V$ and $e \in E$. We let $I(G, G')$ denote the set of isomorphisms between G and G', and let $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ denote the collection of isomorphism classes of connected $[\overline{\mathcal{Z}}, \mathcal{Z}]$-graphs. Once again, for any connected $[\overline{\mathcal{Z}}, \mathcal{Z}]$-graph G, $\langle G\rangle \in \mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ denotes the isomorphism class of G. Likewise, for any $\langle H\rangle \in \mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, the $[\overline{\mathcal{Z}}, \mathcal{Z}]$-graph H denotes an arbitrary representative of $\langle H\rangle$. In a slight abuse of notation, when H is disconnected, we let $\langle H\rangle := \langle \mathscr{C}_{\varnothing_H}(H)\rangle$ denote the isomorphism class of the connected component of H containing the root. This ensures that $\langle H\rangle$ is an element of $\mathscr{G}_*$ for any locally finite $[\overline{\mathcal{Z}}, \mathcal{Z}]$-graph. Also, given a (possibly marked or rooted) graph G, we let V_G and E_G, respectively, denote the vertex and edge sets of G. We also occasionally abuse notation by letting G denote its own vertex set.

We equip $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ with the topology of local convergence, defined as follows.

Definition 2.1 (Local convergence)**.** The sequence $\langle G_n\rangle \in \mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, $n \in \mathbb{N}$, is said to converge locally to $\langle G\rangle \in \mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ if for every sequence of representatives $G_n = (V_n, E_n, \varnothing_n, \overline{\vartheta}^n, \vartheta^n)$, $n \in \mathbb{N}$, and $G = (V, E, \varnothing, \overline{\vartheta}, \vartheta)$, and for every $m \in \mathbb{N}$ there exists $n_m < \infty$ and a sequence $\varphi_{n,m} \in I(B_m([G_*]), B_m([G_{n,*}])), n > n_m, n \in \mathbb{N}$, such that for every $v \in V_{B_m([G_*])}$ and $e \in E_{B_m([G_*])}$, $\vartheta^n_{\varphi_{n,m}(v)} \to \vartheta_v$ and $\overline{\vartheta}^n_{\varphi_{n,m}(e)} \to \overline{\vartheta}_e$ as $n \to \infty$.

We also let $\mathscr{G}_{*,1}[\overline{\mathcal{Z}}, \mathcal{Z}] \subset \mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ denote the (closed) space of isomorphism classes of graphs in $\widehat{\mathscr{G}}_{*,1}$ with edge and vertex marks in $\overline{\mathcal{Z}}$ and $\mathcal{Z}$, respectively, equipped with the topology induced by $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. Note that one can also view $\mathscr{G}_* = \mathscr{G}_*[\{1\}, \{1\}]$ as a space of (isomorphism classes of) marked graphs with trivial marks, that is, when both mark spaces are equal to the trivial Polish space $\{1\}$, in which case the local convergence defined above coincides with the notion defined earlier on $\mathscr{G}_*$.

It is well known that $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ and hence, $\mathscr{G}_{*,1}[\overline{\mathcal{Z}}, \mathcal{Z}]$ and $\mathscr{G}_*$ (equipped with the above local topology), are Polish spaces (see [17, Lemma 3.4]). If $\{\zeta_n\}_{n\in\mathbb{N}} \subset \mathcal{P}(\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}])$ converges in distribution to ζ, then we say that ζ_n converges to ζ *locally weakly*, and denote it by $\zeta_n \Rightarrow \zeta$. If $\langle G_n\rangle \sim \zeta_n$ for every $n \in \mathbb{N}$ and $\langle G\rangle \sim \zeta$, we may also write $\langle G_n\rangle \Rightarrow \langle G\rangle$ to denote the local weak convergence of $\langle G_n\rangle$ to $\langle G\rangle$. As mentioned in Section 1.4, we assume all $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$-random elements are measurable with respect to the Borel σ-algebra defined in terms of the local topology.

2.1.2 Convergence in Probability in the Local Weak Sense

The local topology is not well suited towards understanding many *global* functionals of marked graphs of interest. In this section we briefly introduce a few such functionals which are applied concretely in Section 3.2. The proofs of all lemmas in this section are deferred to Section 2.4.

Given a Polish space $\mathcal{Z}'$ and a finite, unrooted $[\overline{\mathcal{Z}}, \mathcal{Z}']$-graph G, define the (global) empirical measure of the finite, (possibly disconnected) unrooted $[\overline{\mathcal{Z}}, \mathcal{Z} \times \mathcal{Z}']$-graph (G, z) by

$$\pi^{G,z}(\cdot) := \frac{1}{|V_G|} \sum_{v\in V_G} \delta_{z_v}(\cdot),$$

where $\delta_{z_v}(\cdot)$ is the Dirac delta measure concentrated at z_v. Note that for $(G', z') \cong (G, z)$, $\pi^{G,z} = \pi^{G',z'}$, so the empirical measure is determined only by the isomorphism class $\langle G, z\rangle$ of (G, z). More generally, consider the so-called *neighborhood empirical distribution* given by

$$\overline{\pi}^{G,z}_{\varnothing} := \frac{1}{|V_G|} \sum_{v\in V_G} \delta_{\langle B_1(\mathscr{C}_v(G,z))\rangle},$$

where $\mathscr{C}_v(G, z)$ denotes the rooted $[\overline{\mathcal{Z}}, \mathcal{Z} \times \mathcal{Z}']$-graph obtained by restricting (G, z) to the connected component of v in $[G]$, equipped with v as its root. Note that $\overline{\pi}^{G,z}_{\varnothing}$ is a $\mathcal{P}(\mathscr{G}_{*,1})$-valued random element that describes the empirical distribution of the isomorphism class of a uniformly distributed root in (G, z) and its neighborhood. On sparse graphs, neighboring particles are typically strongly dependent. Due to this strong dependence between neighboring vertices, more complex empirical quantities such as $\overline{\pi}_{\varnothing}^{G,z}$ are also of interest (see also Example 3.15 for additional motivation).

In the context of this thesis, we are interested in functionals of IPS X, where for each $v \in V_G$, X_v is the trajectory of a single particle. In Corollary 5.5, we establish conditions under which a sequence of IPS X^n expressed as isomorphism classes $\langle G_n, X^n \rangle$ converge locally weakly to some IPS $\langle G, X \rangle$ given that its initial data $\langle G_n, \xi^n \rangle$ converges locally weakly to the initial data $\langle G, \xi \rangle$. However, we need more than local weak convergence to study the asymptotics of empirical quantities such as $\pi^{G,X}$ or $\overline{\pi}_{\varnothing}^{G,X}$. Theorem 5.7 and Corollary 5.8 of Section 5.2.2 show that under a slightly stronger convergence condition on the initial data than that imposed in Corollary 5.5, these empirical quantities do have a deterministic limit. A key ingredient of the proof is a certain asymptotic correlation decay property, which is first stated in Section 5.2.1 (see Theorem 5.6).

Definition 2.2 (Convergence in probability in the local weak sense)**.** Consider a sequence $\{G_n\}_{n\in\mathbb{N}}$ of finite, (possibly disconnected) unrooted $[\overline{\mathcal{Z}}, \mathcal{Z}]$-random graphs. Then G_n converges to a $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$-random element $\langle G \rangle$ in probability in the local weak sense if for every bounded, continuous mapping $f : \mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}] \to \mathbb{R}$ and as $n \to \infty$,

$$\frac{1}{|V_{G_n}|} \sum_{v \in V_{G_n}} f(\langle \mathscr{C}_v(G_n) \rangle) \to \mathbb{E}\left[f(\langle G \rangle)\right] \text{ in probability.} \tag{2.1.1}$$

Remark 2.3. This definition is the one provided in [61, Definition 2.6]. It differs slightly from the definition of "convergence locally in probability" provided by Hofstad in [88, Definition 2.12(b)], which allows for the possibility that the expectation in (2.1.1) is a conditional expectation. Hofstad gives an example in which for each n, G_n is an Erdös-Rènyi graph with n vertices and parameter λ/n, where λ is a random variable not depending on n. Then G_n converges locally in probability to the Galton-Watson tree whose offspring distribution is Poisson with (random) mean λ [88]. However, G_n does not converge in probability in the local weak sense to anything within our framework. When the limit in (2.1.1) is deterministic, the two definitions coincide.

Convergence in probability in the local weak sense of marked graphs is of a more global nature than local weak convergence and can be used to establish hydrodynamic limits.

Lemma 2.4. *If the unrooted $[\overline{\mathcal{Z}}, \mathcal{Z} \times \mathcal{Z}']$-random graphs (G_n, z^n) converge to $\langle G, z \rangle$ in probability in the local weak sense, then $\pi^{G_n,z^n} \to Law(z_\varnothing)$ and $\overline{\pi}_{\varnothing}^{G_n,z^n} \to Law(\langle B_1(G,z) \rangle)$ in probability with* *respect to* $\mathcal{P}(\mathcal{Z}')$ *and* $\mathcal{P}(\mathscr{G}_{*,1}[\overline{\mathcal{Z}}, \mathcal{Z} \times \mathcal{Z}'])$ *respectively.*

Lacker et al. showed that graphs converge in probability in the local weak sense if, formally speaking, they converge locally weakly and they exhibit a certain correlation decay property [61, Lemma 2.8]. A similar result can be found for unmarked graphs [88, Corollary 2.20(b)]. We make use of the following slight extension of this result, which is similar to the statement of [36, Lemma 7.1].

Lemma 2.5. *Consider a collection of finite, (possibly disconnected) unrooted* $[\overline{\mathcal{Z}}, \mathcal{Z}]$*-random graphs* $\{G_n\}_{n\in\mathbb{N}}$ *and a* $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$*-random element* $\langle G\rangle$ *defined on a common probability space such that* $|G_n| \to \infty$ *in probability. For each* $n \in \mathbb{N}$*, let* $\varnothing_n^i$*,* $i = 1, 2$*, be i.i.d. uniform vertices in* V_{G_n}*. Then,* G_n *converges in probability in the local weak sense to* $\langle G\rangle$ *if and only if* $(\langle \mathscr{C}_{\varnothing_n^i}(G_n)\rangle)_{i=1,2}$ *converges weakly to* $(\langle G^i\rangle)_{i=1,2}$ *in* $(\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}])^2$*, where* $\{\langle G^i\rangle\}_{i=1,2}$ *are i.i.d. copies of* $\langle G\rangle$*.*

2.1.3 Examples of Convergence

In this section we provide several short examples of common graphs and their local weak limits. Unless otherwise stated, each graph G_n is unrooted and the root $\varnothing_n$ is chosen uniformly at random from the vertices of G_n. We only discuss unmarked graphs here.

We begin with a few definitions.

Definition 2.6. A *regular* graph is a graph whose vertices all have degree d for some fixed $d \in \mathbb{N}$.

Definition 2.7. For any $d > 1$, the graph $\mathbb{T}_d$ is the rooted regular tree with degree d.

Example 2.8. Let $G_n = C_n$ be the cycle with n vertices. Then $\langle G_n, \varnothing_n\rangle \to \langle \mathbb{T}_2\rangle$, which is the infinite cycle or path, and G_n converges in probability in the local weak sense to $\mathbb{T}_2$.

Example 2.9. Let $(\mathcal{T}_n, \varnothing_n)$ be the n-generation tree whose non-leaf vertices all have degree $d > 1$ and whose root is the unique vertex of distance n from all leaf vertices. Then $\langle \mathcal{T}_n, \varnothing_n\rangle \to \langle \mathbb{T}_d\rangle$, where $\mathbb{T}_d$ is the rooted, infinite d-regular tree. For $d > 2$, there exists a constant $f_d \in (0, 1)$ such that a uniformly chosen vertex of $\mathcal{T}_n$ will be a leaf node with probability greater than f_d for every $n \in \mathbb{N}$, therefore $\mathcal{T}_n$ does not converge to $\langle \mathbb{T}_d\rangle$ in probability in the local weak sense (see [61, Proposition 7.7] and [61, Theorem 6.4] for more details).

We now describe the convergence of random graphs. We start with a simple graph model.

Definition 2.10. A random regular graph of degree d and with n vertices such that nd is even (denoted $\mathrm{RR}_n(d)$) is an unrooted graph chosen uniformly at random from the set of simple regular graphs with n vertices and uniform degree d.

Example 2.11. If G_n are sampled from $\mathrm{RR}_{2n}(d)$, then $\langle \mathscr{C}_{\varnothing_n}(G_n)\rangle \Rightarrow \langle \mathbb{T}_d\rangle$ and G_n converges in probability in the local weak sense to $\langle \mathbb{T}_d\rangle$ [88, Theorem 2.18].

Definition 2.12. The Erdös-Rènyi random graph with n vertices and edge probability p (denoted $\mathrm{ER}_n(p)$) is the random unrooted graph G of n vertices such that each pair of vertices forms an edge independently with probability p.

The Erdös-Rènyi graph is the first model in this section with a random limit. This limit is the well-known Galton-Watson tree.

Definition 2.13. Let $p \in \mathcal{P}(\mathbb{N}_0)$ be a probability distribution. For any rooted tree $\mathcal{T}$ and $v \in V_{\mathcal{T}}$ recall that $c_v(\mathcal{T})$ is the set of children of v in $\mathcal{T}$. Then the *Galton-Watson* (GW) tree with offspring distribution p (denoted $\mathrm{GW}(p)$) is a random tree $\mathcal{T}$ with root $\varnothing$ such that for any $n \in \mathbb{N}_0$ and conditioned on $B_n(\mathcal{T})$, $\{|c_v(\mathcal{T})|\}_{v:d_{\mathcal{T}}(v,\varnothing)=n}$ is an i.i.d. collection of random variables with law p.

Example 2.14. Let $p_n \in [0,1]$ be a sequence of positive numbers such that $\frac{p_n}{n} \to p$ for some $p \in \mathbb{R}_+$. If for each $n \in \mathbb{N}$, G_n is sampled from $\mathrm{ER}_n(p_n)$ and $\mathcal{T}$ is sampled from $\mathrm{GW}(p)$, then $\langle \mathscr{C}_{\varnothing_n}(G_n)\rangle \Rightarrow \langle \mathcal{T}\rangle$ and G_n converges in probability in the local weak sense to $\langle \mathcal{T}\rangle$ [88, Theorem 2.19].

Lastly, we introduce the configuration model.

Definition 2.15. The configuration with n vertices and degree distribution $\mathbf{d} = (d_0, d_1, \dots) \in \mathbb{N}_0^{\mathbb{N}_0}$ (denoted $\mathrm{CM}_n(\mathbf{d})$) is an unrooted graph G chosen uniformly at random from the set of graphs with n vertices of which d_k vertices have degree k for each $k \in \mathbb{N}_0$.

Note that for a graph with degree distribution $\mathbf{d}$ to exist, $\sum_{k\in\mathbb{N}_0} d_k$ must be even. Even when this is the case, there may not exist a simple graph with degree distribution $\mathbf{d}$. Therefore configuration models are multigraphs with self-loops. Although we exclusively work with simple graphs, it is possible to encode multi-edges and self-loops via vertex and edge marks, so we assume a configuration model is a random graph encoded in this way. See [88, Section 1.3.3 and Chapter 4] for a detailed discussion of configuration models, including other methods of reducing G_n to a simple graph and how these different methods impact the law of G_n and its limits.

The limit of the (multigraph) configuration model is described by a special type of GW tree.

Definition 2.16. For any probability distribution $p \in \mathcal{P}(\mathbb{N}_0)$, define

$$\widehat{p}(\{k\}) := \frac{(k+1)p(\{k+1\})}{\sum_{n\in\mathbb{N}} np(\{n\})}, k \in \mathbb{N}_0. \tag{2.1.2}$$

Then a *unimodular Galton-Watson* (UGW) tree or *size-biased* GW tree with offspring distribution p (denoted UGW(p)) is a $GW(\widehat{p})$ tree except that the offspring distribution of the root is given by p.

Example 2.17. For each $n \in \mathbb{N}$, let G_n be a configuration model sampled from $\mathrm{CM}_n(\mathbf{d}_n)$. Suppose also that if D_n is a random variable representing the number of degrees of a vertex chosen uniformly at random from G_n, then there exists a random variable D such that $D_n \Rightarrow D$ and $\mathbb{E}[D_n] \to \mathbb{E}[D] < \infty$. Let $\mathcal{T}$ be sampled from UGW(Law(D)). Then $\langle G_n, \varnothing_n\rangle \Rightarrow \langle\mathcal{T}\rangle$ and G_n converges in probability in the local weak sense to $\langle\mathcal{T}\rangle$ [88, Theorem 4.1].

Remark 2.18. Preferential attachment graphs (for example, graphs generated by the Barabàsi-Albert model) are known to converge locally and in probability in the local weak sense to a class of so-called "Pólya point trees". However, the construction of these graphs and the limit tree is somewhat involved and therefore we omit it from this thesis. See [88, Chapter 5] for details.

2.2 A Space of Marked Graphs and its Properties

The results stated in this and the next section are fundamental to this thesis. However, all results stated within the section are supplementary results intended to simplify the proofs of our main results. As many of the proofs involved are technical, we relegate the proofs of all results within this section to Appendix A.1.

For any Polish spaces $\overline{\mathcal{Z}}$ and $\mathcal{Z}$, consider the space

$$\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}] := \{[\overline{\mathcal{Z}}, \mathcal{Z}]\text{-graphs } G \text{ s.t. } V_G \subseteq \mathbb{N}\}. \tag{2.2.1}$$

It should be noted that most marked graphs of interest in applications can easily be represented as elements of $\widehat{\mathscr{G}}_*$. For example, finite random graphs such as Erdös-Rènyi graphs (with n vertices) are often described as graphs with vertex set $V_G := \{1, \ldots, n\} \subseteq \mathbb{N}$. Even spatial graphs (in which vertices are placed on some domain within a Euclidean space) can be represented in this manner by encoding the location of the vertex as a vertex mark.

The choice of setting $V_G \subseteq \mathbb{N}$ is somewhat arbitrary, and the arguments made in this thesis would hold with minimal changes if $\mathbb{N}$ were replaced with any other discrete, countably infinite space in (2.2.1). For example, [60] uses the Ulam-Harris-Neveu labeling scheme to identify the vertex sets of infinite, locally finite trees with subsets of the countable set

$$\mathbb{V} := \{\varnothing\} \cup \left(\bigcup_{n \in \mathbb{N}} \mathbb{N}^n \right). \tag{2.2.2}$$

We specifically use $\mathbb{N}$ because $\mathbb{N}$ comes equipped with a notion of ordering and addition which we make use of to more easily construct certain explicit isomorphisms between graphs in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. Such operations could be replicated on any discrete, countably infinite set $\mathcal{S}$, but doing so would complicate the construction of these isomorphisms.

It is now possible to discuss the topology of $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. In the ensuing definition, use the following standard notion of convergence of subsets of $\mathbb{N}$. Given $S_n \subseteq \mathbb{N}, n \in \mathbb{N}$, and $S \subseteq \mathbb{N}$, we write $S_n \to S$ if and only if

$$S = \bigcup_{n \in \mathbb{N}} \bigcap_{n' > n} S_{n'} = \bigcap_{n \in \mathbb{N}} \bigcup_{n \in \mathbb{N}} S_{n'}.$$

Equivalently, $S_n \to S$ if and only if for every $k \in S$, there exist only finitely many $n \in \mathbb{N}$ such that $k \notin S_n$ and for every $k' \notin S$ there exist only finitely $n' \in \mathbb{N}$ such that $k' \in S_{n'}$.

Definition 2.19. [36, Definition B.3] We equip the space $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ with the following notion of convergence: $G_n \to G$ in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ if and only if

1. $\lim_{n\to\infty} V_n = V$;

2. $\lim_{n\to\infty} E_n = E$;

3. $\lim_{n\to\infty} \varnothing_n = \varnothing$;

4. $\lim_{\substack{n\to\infty \\ e \in E_n}} \overline{\vartheta}_e^n = \overline{\vartheta}_e$ for all $e \in E$;

5. $\lim_{\substack{n\to\infty \\ v\in V_n}} \vartheta_v^n = \vartheta_v$ for all $v \in V$.

6. for each $v \in V$, $\lim_{\substack{n\to\infty \\ v\in V_n}} \max\{u \in \mathrm{cl}_v(G_n)\} = \max\{u \in \mathrm{cl}_v(G)\}$.

Remark 2.20. [36, Remark B.4] The least intuitive condition is perhaps condition 6 of Definition 2.19, but it is necessary for the topology on $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ to be compatible with the topology of $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ (in the sense made precise in Proposition 2.26). This is best illustrated via an example of how compatibility could fail without condition 6. Let $\overline{\mathcal{Z}} = \mathcal{Z} = \{1\}$ be trivial and suppose that $V \subseteq \{2k : k \in \mathbb{N}\}$, $\varnothing := 2$ and E is any set of distinct pairs of V such that $G = (V, E, \varnothing)$ is locally finite. If for each $n \in \mathbb{N}$, G_n is defined by $\varnothing_n := 2$, $V_n := V \cup \{2n+1\}$ and $E_n := E \cup \{\{2, 2n+1\}\}$, then conditions 1-5 of Definition 2.19 are all satisfied, and only condition 6 fails. However, note that $G_n \cong G_{n'} \not\cong G$ for all $n, n' \in \mathbb{N}$, so $\langle G_n \rangle \to \langle G_1 \rangle \neq \langle G \rangle$ in $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$.

Definition 2.19 can be tedious to verify. To aid with this, we provide a few simple lemmas that can be used to simplify the processes of checking whether a sequence of graphs $G_n \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, $n \in \mathbb{N}$ converges to $G \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. These lemmas should also illustrate that the topology of the space $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ is actually quite intuitive.

Lemma 2.21. *Let $\mathcal{Z}'$ be another Polish space. If $G_n \to G$ in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, then the marked graphs $(G_n, z^n) \to (G, z)$ in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z} \times \mathcal{Z}']$ if and only if for each $v \in V_G$,*

$$\lim_{\substack{n\to\infty \\ v\in V_{G_n}}} z_v^n = z_v.$$

By Lemma 2.21, if we wish to show that a sequence of IPS (G_n, X^n), $n \in \mathbb{N}$, defined on the graphs G_n, $n \in \mathbb{N}$ converges to (G, X) given the fact that $G_n \to G$, it suffices to prove componentwise convergence of X^n to X. That is, a sequence of IPS defined on convergent underlying graphs in $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K}]$ converges in $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{D}]$ if and only if the dynamics at each individual particle of the IPS converge.

Lemma 2.22. *If conditions 1,2,3 and 6 of Definition 2.19 hold for the sequence of graphs $G_n \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, $n \in \mathbb{N}$ and $G \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, then for all sufficiently large n, $B_m([G_{n,*}]) = B_m([G_*])$. The converse holds under the extra condition that $V_G = \{1, \ldots, |V_G|\}$ where $\{1, \ldots, \infty\} = \mathbb{N}$.*

We finish with one slightly unintuitive lemma. It is easily seen that if $\langle B_{m_n}(G_n) \rangle \to \langle G \rangle$ in $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ for a given sequence of integers $m_n \to \infty$, then $\langle G_n \rangle \to \langle G \rangle$. However, this is not quite true of convergence in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. Instead, we need the following slightly stronger condition.

Lemma 2.23. *Suppose there exist two sequence $m_n, N_n \to \infty$ such that $B_{m_n}(G_n) \to G$ and $G_n[\{1, \dots, N_n\}] \to G$ in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. Then $G_n \to G$ in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$.*

Remark 2.24. Some care needs to be taken with regard to the convergence of disconnected graphs $(G_n, \varnothing_n) \to (G, \varnothing)$. Even if $G_n \to G$, $\{\mathscr{C}_{\varnothing_n}(G_n)\}$ may fail to be a convergent sequence. For an example, suppose $\overline{\mathcal{Z}} = \mathcal{Z} = \{1\}$ is trivial and $\varnothing_n = \varnothing = 1$ for all $n \in \mathbb{N}$. For each even n, let G_n be the path $\{1, 3, \dots, 2n+1, 2n, 2n-2, \dots, 2\}$ and for each odd n let G_n be the disjoint union of the paths $\{1, 3, \dots, 2n+1\}$ and $\{2, 4, \dots, 2n\}$. Then if G is the disjoint union of the paths $\{1, 3, \dots\}$ and $\{2, 4, \dots\}$ rooted at 1, then $G_n \to G$ in $\widehat{\mathscr{G}}_*$ but $\mathscr{C}_{\varnothing_n}(G_n)$, $n \in \mathbb{N}$ is not even a convergent sequence in $\widehat{\mathscr{G}}_*$.

Lastly, the topology generated by convergence as in Definition 2.19 is Polish.

Proposition 2.25. *[36, Lemma B.5] The space $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ is Polish.*

2.3 The Relationship between $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ and $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$

In this section we describe the relationship between the spaces $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ and $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. We begin by establishing the continuity of the map $G \mapsto \langle G \rangle$ and the lower semicontinuity of its inverse (in the sense of [5, Definition 1.4.2]). These continuity properties of $G \mapsto \langle G \rangle$ and its inverse ensure that limit theorems in the space $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ and limit theorems in the space $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ are often interchangeable in some sense. This is then used to prove a number of selection theorems with a focus on interchanging the results of limit theorems in $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ and limit theorems in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. We finish with a technical lemma that, in combination with Lemma 2.5, turns statements about convergence in probability in the local weak sense to a statement about convergence in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. All proofs within this section are deferred to Appendix A.2.

We start by showing that the map $G \mapsto \langle G \rangle$ is continuous and its set-valued inverse map $\langle G \rangle \mapsto \{G \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}] : G \in \langle G \rangle\}$ is lower semicontinuous.

Proposition 2.26. *If $G_n \to G$ in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, then the isomorphism classes also converge locally, that is, $\langle G_n \rangle \to \langle G \rangle$ in $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. Moreover, given the limit $\langle G_n \rangle \to \langle G \rangle$ in $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ and any* *representative* $G \in \langle G \rangle$ *such that* $G \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, *there exists a sequence of representatives* $G_n \in \langle G_n \rangle$, $n \in \mathbb{N}$, *such that* $G_n \to G$ *in* $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. *In other words, the correspondence* $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}] \ni \langle G \rangle \mapsto \{G \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}] : G \in \langle G \rangle\}$ *is lower semicontinuous in the sense of [5, Definition 1.4.2].*

Remark 2.27. Recall that $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ supports disconnected graphs and for any $G \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, $\langle G \rangle = \langle \mathscr{C}_\varnothing(G) \rangle \in \mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. Thus, if G and H are connected and $G, H \in \langle G \rangle$ then $G \cong H$, but this is not necessarily true if either or both graphs are disconnected. It should be noted that the set of connected graphs in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ is not closed. Let $\varnothing_n = 1$ for all n and let G_n be the path $\{1, 3, \dots, 2n+1, 2n, 2n-2, \dots, 2\}$. Then G_n converges to G which is the disjoint union of the paths $\{1, 3, \dots\}$ and $\{2, 4, \dots\}$ with $\varnothing = 1$.

So far we have only discussed deterministic graphs and isomorphism classes. However, we are also interested in translating results about random graphs to results about random isomorphisms and vice versa. If $\langle G \rangle$ is a random isomorphism class, then we need a way to ensure that it is possible to describe a random graph G such that G is a.s. a representative of $\langle G \rangle$. Moreover, this selection needs to be measurable in the following sense.

Definition 2.28. Given a $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$-random element $\langle G \rangle$ and a σ-algebra $\mathcal{F} \supseteq \sigma(\langle G \rangle)$, a $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$-random element G is said to be a $\mathcal{F}$-*measurable representative* of $\langle G \rangle$ if $G \in \langle G \rangle$ a.s. and G is $\mathcal{F}$-measurable. If $\mathcal{F} = \sigma(\langle G \rangle)$, then G is said to be a *measurable representative* of $\langle G \rangle$.

This leads to the following measurable selection theorem, which is a consequence of the Kuratowski and Ryll-Nardzewski measurable selection theorem which we state as Theorem A.2.

Theorem 2.29. *Given a* $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$*-random element* $\langle G \rangle$, *there exists a connected measurable representative* G *of* $\langle G \rangle$. *Moreover, for any sequence of* $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$*-random elements* $\langle G_n \rangle$, $n \in \mathbb{N}$ *and* $\langle G \rangle$, *there exists for each* $n \in \mathbb{N}$ *a connected* $\sigma(\langle G \rangle) \vee \sigma(\langle G_n \rangle)$*-measurable representative* G_n *of* $\langle G_n \rangle$ *such that if* $\langle G_n \rangle \to \langle G \rangle$ *a.s. in* $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, *then* $G_n \to G$ *a.s. in* $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$.

Lastly, although we do not establish an equivalent notion of convergence in probability in the local weak sense for marked graphs (rather than isomorphism classes) in this thesis, we do introduce the following corollary of Theorem 2.29 that complements Lemma 2.5. For this theorem, we introduce the notion of a *doubly-rooted* graph, which is simply a (possibly marked) graph $(G, \varnothing^1, \varnothing^2)$

equipped with two (ordered) roots. An isomorphism between doubly rooted graphs $(G_i, \varnothing_i^1, \varnothing_i^2)$, $i \in 1,2$, is simply an isomorphism $\varphi : V_{G_1} \to V_{G_2}$ such that $\varphi(\varnothing_1^k) = \varnothing_2^k$ for $k = 1,2$.

Corollary 2.30. *Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a collection of finite, (possibly disconnected) unrooted $[\overline{\mathcal{Z}}, \mathcal{Z}]$-random graphs $\{G_n\}_{n\in\mathbb{N}}$ whose vertex sets are subsets of $\mathbb{N}$, and for each $n \in \mathbb{N}$ let $\varnothing_n^i$, $i = 1,2$, be $\mathcal{F}$-measurable vertices in V_{G_n} such that $d_{G_n}(\varnothing_n^1, \varnothing_n^2) \to \infty$ in probability. For each i, suppose that $\langle \mathscr{C}_{\varnothing_n^i}(G_n)\rangle$ converges a.s. to the $\mathcal{F}$-measurable $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$-random element $\langle G^i\rangle$. Then there exists an $\mathcal{F}$-measurable sequence of doubly rooted $[\overline{\mathcal{Z}}, \mathcal{Z}]$-random graphs $(\widetilde{G}_n, \widetilde{\varnothing}_n^1, \widetilde{\varnothing}_n^2) \cong (G_n, \varnothing_n^1, \varnothing_n^2)$, $n \in \mathbb{N}$, $(G, \varnothing^1, \varnothing^2)$, with vertices in $\mathbb{N}$ such that for $i = 1,2$, $(\widetilde{G}_n, \widetilde{\varnothing}_n^i)$ converges in probability to $(G, \varnothing^i)$ in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ and $\mathscr{C}_{\varnothing^i}(G)$ is a measurable representative of $\langle G^i\rangle$. Furthermore, G consists of two connected components, and $d_G(\varnothing_1, \varnothing_2) = \infty$ a.s..*

Remark 2.31. Note that a doubly rooted $[\overline{\mathcal{Z}}, \mathcal{Z}]$-random graph $(G, \varnothing^1, \varnothing^2)$ with vertices in $\mathbb{N}$ is $\mathcal{F}$-measurable if $(G, \varnothing^1)$ is a $\mathcal{F}$-measurable $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$-random element and $\varnothing^2$ is a $\mathcal{F}$-measurable random variable in $\mathbb{N}$.

Remark 2.32. In Section 5.3.2, Corollary 2.30 is used together with Lemma 2.5 to establish a given correlation decay property for a sequence of IPS X^n whose initial data converge in probability in the local weak sense to the initial data of another IPS X. Lemma 2.5 is used again to show that this correlation decay property combined with the local weak convergence of X^n to X (which is established in Corollary 5.5) implies that X^n converges in probability in the local weak sense to X.

2.4 Proofs of Lemmas in Section 2.1.2

We start with a quick proof that convergence in probability in the local weak sense implies convergence in probability of the empirical quantities $\pi^{G,z}$ and $\overline{\pi}_{\varnothing}^{G,z}$.

Proof of Lemma 2.4. $\pi^{G_n,z^n} \to \mathrm{Law}(z_\varnothing)$ by [61, Lemma 2.7]. The proof of the second statement is nearly identical. Let $f : \mathscr{G}_{*,1}[\overline{\mathcal{Z}}, \mathcal{Z} \times \mathcal{Z}'] \to \mathbb{R}$ be a bounded, continuous function. Then by (2.1.1),

$$\int f(H)\, \overline{\pi}^{G_n,z^n}(dH) = \frac{1}{|V_{G_n}|} \sum_{v\in V_{G_n}} f(\langle \mathscr{C}_v(B_1(G_n, z^n))\rangle) \to \mathbb{E}\left[f(\langle B_1(G,z)\rangle)\right],$$

in probability. Since this holds for all bounded, continuous functions f, $\overline{\pi}_{\varnothing}^{G_n,z^n} \to \mathrm{Law}(\langle B_1(G,z)\rangle)$ in probability. $\square$

We finish the section by establishing the alternate characterization of convergence in probability in the local weak sense.

Proof of Lemma 2.5. Although [61, Lemma 2.8] is stated and proved for graphs and isomorphism classes with vertex marks but no edge marks, the analogous result holds for graphs and isomorphism classes with edge marks by the exact same argument. Therefore, since $|G_n| \to \infty$ in probability, G_n converges to $\langle G\rangle$ in probability in the local weak sense if and only if for any bounded, continuous functions $f_1, f_2 : \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}] \to \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{E}\left[f_1(\langle \mathscr{C}_{\varnothing_n^1}(G_n)\rangle) f_2(\langle \mathscr{C}_{\varnothing_n^2}(G_n)\rangle)\right] = \mathbb{E}\left[f_1(\langle G\rangle)\right] \mathbb{E}\left[f_2(\langle G\rangle)\right]. \tag{2.4.1}$$

If $(\langle \mathscr{C}_{\varnothing_n^i}(G_n)\rangle)_{i=1,2}$ converges weakly to $(\langle G^i\rangle)_{i=1,2}$, then given that $\langle G^1\rangle \overset{(\mathrm{d})}{=} \langle G^2\rangle \overset{(\mathrm{d})}{=} \langle G\rangle$ are independent, (2.4.1) holds by definition of weak convergence. Now suppose (2.4.1) holds. Since $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ is a metric space, the algebra of separable bounded functions on $(\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}])^2$ strongly separates points and hence, is convergence determining [33, Theorem 3.4.5(b)]. Thus, (2.4.1) implies

$$\lim_{n\to\infty} \mathbb{E}\left[f(\langle \mathscr{C}_{\varnothing_n^1}(G_n)\rangle, \langle \mathscr{C}_{\varnothing_n^2}(G_n)\rangle)\right] = \mathbb{E}\left[f(\langle G^1\rangle, \langle G^2\rangle)\right],$$

for all bounded, continuous $f : (\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}])^2 \to \mathbb{R}$, which completes the proof. $\square$

CHAPTER 3

Framework, Assumptions and Examples

3.1 The Model

Consider an IPS in which each particle takes values in a countable state space $\mathcal{X}$, which is identified with a subset of $\mathbb{Z}$ and equipped with the discrete topology. Let $\mathcal{J} \subseteq \{i-j : i,j \in \mathcal{X}, i \neq j\}$ denote the set of allowable transitions or jumps of any particle, and assume $\mathcal{J}$ is finite. We equip $\mathcal{J}$ with the counting measure $\#_{\mathcal{J}}$. For convenience, we denote $\#_{\mathcal{J}}(\{j\})$ by $\#_{\mathcal{J}}(j)$ for each $j \in \mathcal{J}$. Let $\overline{\mathcal{K}}$ and $\mathcal{K}$ be two Polish spaces that specify static parameters of the model such as random environments (see Example 3.17), histories before time zero for non-Markovian processes (see Example 3.18), or heterogeneities and other graph attributes. Supposing that the underlying $[\overline{\mathcal{K}}, \mathcal{K}]$-graph G of the IPS specifying instantaneous interactions between particles is non-random, we may represent the initial state of the process by a $\mathcal{X}^{V_G}$-random element ξ. Then the model is specified by a family of jump rate functions $\mathbf{r} := \{r_j^{G,v} : \mathbb{R}_+ \times \mathcal{D}^{V_G} \to \mathbb{R}_+, j \in \mathcal{J}, G \text{ a } [\overline{\mathcal{K}}, \mathcal{K}]\text{-graph and } v \in V_G\}$, where $\mathcal{D}$ is the space of càdlàg functions from $\mathbb{R}_+$ to $\mathcal{X}$. In other words, we can define an IPS with jump rates $\mathbf{r}$ and initial data given by the $[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-graph (G, ξ) to be a $\mathcal{D}^{V_G}$-valued random element $X^{G,\xi}$ such that for each $v \in V_G$, $t \in \mathbb{R}_+$ and $j \in \mathcal{J}$, the transition rate from $X_v^{G,\xi}(t)$ to $X_v^{G,\xi}(t) + j$ is given by $r_j^{G,v}(t, X^{G,\xi})$. In Section 3.1.1 below, we establish a standing assumption on the jump rates that is held throughout the thesis. Then in Section 3.1.2 we define the dynamics of the IPS in terms of an SDE and discuss various notions of solutions to the SDE.

3.1.1 A Standing Assumption on the Jump Rates

We now define a standing assumption on the set of jump rate functions. Let $\widehat{\mathscr{G}}_{*,1}$ denote the set of finite rooted graphs $H = (V_H, E_H, \varnothing_H)$ of radius 1.

Standing Assumption. For every $j \in \mathcal{J}$, there exists a family of measurable local jump rate functions $\{\bar{r}_j^H : \mathbb{R}_+ \times \mathcal{D}^{V_H} \times \overline{\mathcal{K}}^{E_H} \times \mathcal{K}^{V_H} \to \mathbb{R}_+\}_{H=(V_H,E_H,\varnothing_H)\in\widehat{\mathscr{G}}_{*,1}}$ such that,

(1) for every $H \in \widehat{\mathscr{G}}_{*,1}$, $(\overline{\vartheta}, \vartheta) \in \overline{\mathcal{K}}^{E_H} \times \mathcal{K}^{V_H}$ $(t,x) \mapsto \bar{r}_j^H(t,x,\overline{\vartheta},\vartheta)$ is predictable in the sense that for every $t > 0$ and $x, y \in \mathcal{D}^{V_H}$,

$$x(s) = y(s) \quad \forall s \in [0,t) \quad \Rightarrow \quad \bar{r}_j^H(t,x,\overline{\vartheta},\vartheta) = \bar{r}_j^H(t,y,\overline{\vartheta},\vartheta);$$

(2) given any (possibly unrooted) $[\overline{\mathcal{K}}, \mathcal{K}]$-graph $G = (V, E, \overline{\vartheta}, \vartheta)$ or $G = (V, E, \varnothing, \overline{\vartheta}, \vartheta)$ and $v \in V$,

$$r_j^{G,v}(t,x) = \bar{r}_j^{H_v}(t, x_{H_v}, (\overline{\vartheta}_e)_{e\in E_{H_v}}, (\vartheta_u)_{u\in V_{H_v}}), \qquad \text{for every } (t,x) \in \mathbb{R}_+ \times \mathcal{D}^V, \tag{3.1.1}$$

where $H_v := (G[\mathrm{cl}_v], v)$ is the induced subgraph of G on the closure cl_v of v, with v as root;

(3) If G is a $[\overline{\mathcal{K}}, \mathcal{K}]$-random graph in some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then for any $\mathcal{H}^G$-measurable random vertex v in G, the map $\Omega \times \mathcal{J} \times \mathbb{R}_+ \times \mathcal{D} \ni (\omega, j, t, x) \mapsto r_j^{G(\omega),v(\omega)}(t,x)$ is $\mathcal{H}^G \otimes \mathcal{B}(\mathcal{J} \times \mathbb{R}_+ \times \mathcal{D})$-measurable.

Remark 3.1. Condition 3 of the standing assumption is a technical assumption that easily holds for all applications of interest. In particular, it holds automatically if there exists a discrete, countably infinite space $\mathcal{S}$ such that $V_G \subseteq \mathcal{S}$ almost surely. To see why, first note that $\widehat{\mathscr{G}}_{*,1}^{\mathcal{S}} := \{H \in \widehat{\mathscr{G}}_{*,1} : V_H \subseteq \mathcal{S}\}$ is a countable space that is naturally equipped with the discrete topology. Thus, if $V_G \subseteq \mathcal{S}$, then the map $\omega \mapsto H_v(\omega) := (G(\omega)[\mathrm{cl}_{v(\omega)}], v(\omega)) \in \widehat{\mathscr{G}}_{*,1}^{\mathcal{S}}$ is a measurable map with a countable image so $(\omega, j, t, x) \mapsto r_j^{G(\omega),v(\omega)}(t,x) = \sum_{H\in\widehat{\mathscr{G}}_{*,1}^{\mathcal{S}}} \mathbb{I}_{\{H=H_v(\omega)\}} \bar{r}_j^H(t,x,\overline{\vartheta}(\omega),\vartheta(\omega))$ is measurable by the measurability of the functions $\bar{r}_j^H$ for each $H \in \widehat{\mathscr{G}}_{*,1}^{\mathcal{S}}$ and $j \in \mathcal{J}$.

With a few exceptions, we also impose the following mild boundedness condition on the jump rates of the IPS model throughout the thesis.

Assumption 1. There exists a function $C : \mathbb{N} \times \mathbb{R}_+ \to \mathbb{R}_+$ that is non-decreasing in each of its arguments and such that for any $[\overline{\mathcal{K}}, \mathcal{K}]$-graph G, $v \in V_G$, $j \in \mathcal{J}$, $x \in \mathcal{D}^G$, $T \in \mathbb{R}_+$ and $t \in [0,T]$, $r_j^{G,v}(t,x) \leq C(|\mathrm{cl}_v(G)|, T)$.

For results in which the isomorphism class of the IPS is relevant, we introduce the following additional symmetry assumption.

Assumption 2. Let $\{\bar{r}_j^H : \mathbb{R}_+ \times \mathcal{D}^{V_H} \times \overline{\mathcal{K}}^{E_H} \times \mathcal{K}^{V_H} \to \mathbb{R}_+\}_{H=(V_H,E_H,\varnothing_H)\in\widehat{\mathscr{G}}_{*,1}}$ be the family of local jump rate functions defined in the standing assumption. Then for every $t > 0$, $H \mapsto \bar{r}_j^H(t,\cdot,\cdot,\cdot)$ is a class function in the sense that for any $(x, \overline{\vartheta}, \vartheta) \in \mathcal{D}^{V_H} \times \overline{\mathcal{K}}^{E_H} \times \mathcal{K}^{V_H}$, $\widehat{H} \in \widehat{\mathscr{G}}_{*,1}$ with $\widehat{H} \cong H$, and $\varphi \in I(H, \widehat{H})$,

$$\bar{r}_j^H(t, x, \overline{\vartheta}, \vartheta) = \bar{r}_j^{\widehat{H}}(t, \varphi_* x, \varphi_* \overline{\vartheta}, \varphi_* \vartheta). \tag{3.1.2}$$

Remark 3.2. [36, Remark 3.2] As a consequence of the class function property of $\bar{r}^H$ specified in Assumption 2, the jump rate function $r^{G,\cdot}$ satisfies a similar class property. Given any isomorphic graphs $G, G' \in \widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K}]$ and $\varphi \in I(G, G')$, for any $v \in V_G$, $t \in \mathbb{R}_+$ and $x \in \mathcal{D}^{V_G}$,

$$r_j^{G,v}(t,x) = r_j^{G',\varphi(v)}(t, \varphi_* x). \tag{3.1.3}$$

3.1.2 Dynamics and Notions of Solutions

Fix a family of jump rate functions $\mathbf{r}$ that satisfy the standing assumption introduced in the previous section. Also fix a random possibly unrooted $[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-graph (G, ξ), henceforth referred to as the *initial data*, which encodes both the random $[\overline{\mathcal{K}}, \mathcal{K}]$-graph describing the interaction structure of the IPS as well as the $\mathcal{X}$-valued initial conditions encoded by ξ. Then the dynamics of the associated IPS are described by the following SDE:

$$X_v^{G,\xi}(t) = \xi_v + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j \mathbb{I}_{\{r \leq r_j^{G,v}(s, X^{G,\xi})\}} \mathbf{N}_v^G(ds, dr, dj), \quad v \in V, t \in [0, \infty), \tag{3.1.4}$$

where $\mathbf{N}^G$ is the so-called driving noise, comprised of a collection of i.i.d Poisson point processes described in Definition 3.3 below. When (G, ξ) is deterministic, the driving noise $\mathbf{N}^G = \{\mathbf{N}_v^G\}_{v \in V_G}$ is simply a collection of i.i.d. adapted Poisson processes and (3.1.4) reduces to a standard Poisson-driven SDE. When (G, ξ) is random, analogous to what is done with initial conditions, it is more natural to think of the driving Poisson processes as marks on the graphs and their measurability and adaptedness properties have to be described with some care, as spelled out in Definition 3.3.

Definition 3.3 (Driving Noise)**.** [36, Definition 3.3] Given a complete, filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ such that $\mathbb{F}$ satisfies the usual conditions and a (possibly random) $\mathcal{F}_0$-measurable $[\overline{\mathcal{K}}, \mathcal{K}\times \mathcal{X}]$-graph (G, ξ), an $\mathbb{F}$-driving noise (compatible with G) is a $[\overline{\mathcal{K}}, \mathcal{K} \times \mathscr{N}(\mathbb{R}_+^2 \times \mathcal{J})]$-random graph $(G, \mathbf{N}^G)$ that satisfies the following properties:

1. conditioned on $\mathcal{F}_0$, $\mathbf{N}^G = \{\mathbf{N}_v^G\}_{v \in V_G}$ is a collection of i.i.d. Poisson processes on $\mathbb{R}_+^2 \times \mathcal{J}$ with intensity measure $\mathrm{Leb}^2 \otimes \#_{\mathcal{J}}$, indexed by the vertex set V_G of G;

2. for any $t > 0$ and $A \in \mathcal{B}((t, \infty) \times \mathbb{R}_+ \times \mathcal{J})$, the $[\overline{\mathcal{K}}, \mathcal{K} \times \mathbb{N}_0]$-random graph $(G, \mathbf{N}^G(A))$ is conditionally independent of $\mathcal{F}_t$ given $\mathcal{F}_0$;

3. for any $\mathcal{F}_0$-measurable $v \in V_G$, $\mathbf{N}_v^G$ is an $\mathbb{F}$-adapted point process in the sense described in Section 1.4.

With a slight abuse of notation, we often denote the driving noise $(G, \mathbf{N}^G)$ just by $\mathbf{N}^G$.

Remark 3.4. When V_G is a.s. a subset of a countable, deterministic set $\mathcal{S}$, Definition 3.3 can be simplified. Let $\mathbf{N} := \{\mathbf{N}_k\}_{k \in \mathcal{S}}$ be a collection of i.i.d. $\mathbb{F}$-Poisson processes on $\mathbb{R}_+^2 \times \mathcal{J}$ with intensity measure $\mathrm{Leb}^2 \otimes \#_{\mathcal{J}}$. Let $\mathbf{N}^G := \{\mathbf{N}_v\}_{v \in V_G}$. Then $(G, \mathbf{N}^G)$ is $\mathcal{F}$-driving noise compatible with G.

When (G, ξ) is random, analogous to what is done with initial conditions and the driving noise, it is natural to also encode the trajectories of the IPS, or equivalently any solution to (3.1.4), as additional marks on the random graph. This leads to the following definitions of weak and strong solutions to the SDE (3.1.4).

Given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ that supports a $[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X} \times \mathscr{N}(\mathbb{R}_+^2 \times \mathcal{J})]$-random graph $(G, \xi, \mathbf{N}^G)$, define $\mathcal{G}_t := \sigma((G, \xi, \mathbf{N}^G(A)) : A \in \mathcal{B}([0, t] \times \mathbb{R}_+ \times \mathcal{J}))$ for $t \geq 0$, and define $\mathbb{H}^{G,\xi,\mathbf{N}^G}$ to be the augmentation of the filtration $\mathbb{G} = \{\mathcal{G}_t\}_{t \in \mathbb{R}_+}$, that is, $\mathbb{H}^{G,\xi,\mathbf{N}^G}$ is the smallest complete, right-continuous filtration such that $\mathcal{H}_t^{G,\xi,\mathbf{N}^G} \supseteq \mathcal{G}_t$ for every $t \geq 0$, and $\mathcal{H}_0^{G,\xi,\mathbf{N}^G}$ contains all sets $N \subset A \in \mathcal{F}$ with $\mathbb{P}(A) = 0$. When (G, ξ) is deterministic, we denote $\mathbb{H}^{G,\xi,\mathbf{N}^G}$ simply by $\mathbb{H}^{\mathbf{N}^G}$.

Definition 3.5 (Weak and Strong Solutions)**.** [36, Definition 3.4] Given (possibly random) initial data (G, ξ), a weak solution to (3.1.4) is a tuple $((G, X^{G,\xi}, \mathbf{N}^G), (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}))$ such that

1. $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a complete, filtered probability space such that $\mathbb{F}$ satisfies the usual conditions;

2. $(G, \mathbf{N}^G)$ is an $\mathbb{F}$-driving noise compatible with G in the sense of Definition 3.3;

3. $(G, X^{G,\xi})$ is a random $[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{D}]$-graph with $\mathbb{F}$-adapted vertex marks (i.e., for any $\mathcal{F}_0$-measurable $v \in V_G$, $X_v^{G,\xi}$ is an $\mathbb{F}$-adapted process) such that $X^{G,\xi}$ satisfies (3.1.4) $\mathbb{P}$-a.s..

Given a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and an $\mathcal{H}^{G,\xi,\mathbf{N}^G}$-driving noise $(G, \mathbf{N}^G)$ compatible with G in the sense of Definition 3.3, a $\mathbf{N}^G$-strong solution $(G, X^{G,\xi})$ is a $[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{D}]$-random graph such that $((G, X^{G,\xi}, \mathbf{N}^G), (\Omega, \mathcal{F}, \mathbb{H}^{G,\xi,\mathbf{N}^G}, \mathbb{P}))$ is a weak solution to (3.1.4).

Remark 3.6. [36, Remark 3.5] For conciseness, we often omit mention of the whole probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and simply refer to $(\mathbb{F}, \mathbf{N}^G)$ as a filtration-Poisson process pair, and say $(G, X^{G,\xi})$ is a $(\mathbb{F}, \mathbf{N}^G)$-weak solution to (3.1.4) if $((G, X^{G,\xi}, \mathbf{N}^G), (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}))$ is a weak solution to (3.1.4). We also say that $(G, X^{G,\xi})$ is a weak solution to (3.1.4) if there exists a filtration-Poisson process pair $(\mathbb{F}, \mathbf{N}^G)$ such that $(G, X^{G,\xi})$ is a $(\mathbb{F}, \mathbf{N}^G)$-weak solution to (3.1.4). We would also like to emphasize that if X and X' are both $(\mathbb{F}, \mathbf{N}^G)$-weak solutions, then they are implicitly defined on the same filtered probability space with the same driving noise.

Remark 3.7. [36, Section 3.3] Under Assumption 2, weak solutions to (3.1.4) can be described by their isomorphism classes. Given $\langle G, \xi \rangle \in \mathscr{G}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$, let (G_i, ξ^i), $i = 1, 2$, be two different connected representatives of $\langle G, \xi \rangle$ and fix some isomorphism $\varphi \in I((G_1, \xi^1), (G_2, \xi^2))$. Let $(\mathbb{F}, \mathbf{N}^{G_1})$ be a filtration-Poisson process pair compatible with G_1 and let X^{G_1,ξ_1} be a $(\mathbb{F}, \mathbf{N}^{G_1})$-weak solution to (3.1.4) for the initial data (G_1, ξ_1). Define,

$$\mathbf{N}^{G_2} := \varphi_* \mathbf{N}^{G_1} \quad \text{and} \quad X^{G_2,\xi_2} := \varphi_* X^{G_1,\xi_1}.$$

Then $(\mathbb{F}, \mathbf{N}^{G_2})$ is a filtration-Poisson process pair compatible with G_2 and by Remark 3.2 and the form of the SDE (3.1.4), it follows that X^{G_2,ξ_2} is a $(\mathbb{F}, \mathbf{N}^{G_2})$-weak solution to (3.1.4).

3.1.3 Notions of Well-Posedness

Given the definitions of weak and strong solutions in the previous section, the definitions of uniqueness and well-posedness we employ are mostly standard, except for the notion of strong well-posedness which requires pathwise uniqueness of solutions instead of just uniqueness in law.

Definition 3.8 (Uniqueness notions)**.** [36, Definition 3.6] The SDE (3.1.4) is said to be *unique in law* for the initial data (G, ξ) if given any other initial data (G', ξ'), possibly defined on a different probability space, that satisfies $(G', \xi') \stackrel{(\mathrm{d})}{=} (G, \xi)$, any weak solutions (G, X) and (G', X') to the SDE for the respective initial data (G, ξ) and (G', ξ') satisfy $(G, X) \stackrel{(\mathrm{d})}{=} (G', X')$. The SDE (3.1.4) is said to be *pathwise unique* for (G, ξ) if for any filtration-Poisson process pair $(\mathbb{F}, \mathbf{N}^G)$ and any two $(\mathbb{F}, \mathbf{N}^G)$-weak solutions (G, X^1) and (G, X^2) to (3.1.4), $(G, X^1) = (G, X^2)$ a.s..

Solutions to the SDE (3.1.4) for the initial data (G, ξ) will typically be denoted by $(G, X^{G,\xi})$.

Definition 3.9 (Well-posedness)**.** [36, Definition 3.7] We say that the SDE (3.1.4) is *well-posed* for the initial data (G, ξ) if there exists at least one weak solution to (3.1.4) and the SDE (3.1.4) is unique in law for (G, ξ). We say the SDE (3.1.4) is *strongly well-posed* for the initial data (G, ξ) if there exists at least one weak solution to (3.1.4) and the SDE (3.1.4) is pathwise unique for (G, ξ).

The next lemma establishes an intuitive alternative formulation of strong well-posedness for random initial data. Its proof is deferred to Appendix D.2.

Lemma 3.10. *[36, Lemma 3.8] The SDE* (3.1.4) *is strongly well-posed for the random initial data* (G, ξ) *if it is strongly well-posed for a.s. every realization of* (G, ξ)*.*

In addition, we show that a standard Yamada-Watanabe type result holds below.

Lemma 3.11. *[36, Lemma 3.10] For the initial data* (G, ξ)*, if the SDE* (3.1.4) *is strongly well-posed, then it is also well-posed. Furthermore,* (3.1.4) *is strongly well-posed for* (G, ξ) *if and only if the set of weak solutions to* (3.1.4) *for* (G, ξ) *is non-empty and coincides with the set of strong solutions to* (3.1.4)*.*

Proof. The lemma can be deduced by showing that pathwise uniqueness as defined in Definition 3.8 matches the definition of pathwise uniqueness in [58]. Then by [58, Theorem 3.14], strong well-posedness is equivalent to all weak solutions being strong as desired. The same theorem also shows that strong well-posedness implies well-posedness. $\square$

Remark 3.12. [36, Remark 3.11] Definition 3.5 and Lemma 3.11 imply that under strong well-posedness of the SDE, $(G, X^{G,\xi})$ is characterized by the initial data (G, ξ) and a compatible driving noise $(G, \mathbf{N}^G)$. Hence, a coupling of the initial data and driving noises of a sequence of IPS immediately yields a coupling of the respective solutions to (3.1.4) (see Sections 5.3 and 5.3.3). It is worth emphasizing that *strong* well-posedness is key to facilitating the construction of such couplings.

Remark 3.13. [36, Section 3.3] It is an immediate consequence of Remark 3.7 that under Assumption 2, (strong) well-posedness of (3.1.4) is a class property in the sense that for any two $[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-graphs $(G_1, \xi^1) \cong (G_2, \xi^2)$, (3.1.4) is (strongly) well-posed for the initial data (G_1, ξ^1) if and only if it is (strongly) well-posed for the initial data (G_2, ξ^2).

Remarks 3.13 and 3.7 motivate the following definition.

Definition 3.14 (Strong well-posedness for isomorphism classes)**.** [36, Definition 3.9] If Assumption 2 holds, we say the SDE (3.1.4) is (strongly) well-posed for the (possibly random) initial data $\langle G, \xi \rangle$ taking values in $\mathscr{G}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$ if (3.1.4) is (strongly) well-posed for (G, ξ) for all (possibly random) connected $[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-graphs (G, ξ) with $(G, \xi) \in \langle G, \xi \rangle$ a.s.. Furthermore, we say that $\langle G, X^{G,\xi} \rangle$ is a strong (resp. weak) solution to (3.1.4) for $\langle G, \xi \rangle$ if there exists a connected (random) representative (G, ξ) that lies in $\langle G, \xi \rangle$ a.s. and a strong (resp. weak) solution $(G, X^{G,\xi})$ to (3.1.4) for (G, ξ) such that $(G, X^{G,\xi}) \in \langle G, X^{G,\xi} \rangle$ a.s..

3.2 Motivating Examples

We now present illustrative examples of IPS that fall within our framework. We assume throughout that these IPS are restricted to graphs whose vertex sets are contained in a deterministic, countable and discrete set $\mathcal{S}$ so that condition 3 of the standing assumption holds (by Remark 3.1). Under this condition, all examples satisfy the standing assumption.

Example 3.15 (Finite-state Markovian IPS). [36, Example 4.8] Consider the graph $G = (V, E, \varnothing)$ (i.e., $\overline{\mathcal{K}}$ and $\mathcal{K}$ are trivial) and suppose that for every $j \in \mathcal{J}$, there exist functions $\tilde{r}_j^H : \mathcal{X}^{V_H} \to \mathbb{R}_+, H \in \widehat{\mathscr{G}}_{*,1}$, such that the functions $\bar{r}_j^H, H = (V_H, E_H, \varnothing_H) \in \widehat{\mathscr{G}}_{*,1}$, from the standing assumption satisfy

$$\bar{r}_j^H(t, x) = \tilde{r}_j^H(x(t-)) \quad \forall (t, x) \in \mathbb{R}_+ \times \mathcal{D}^{V_H}. \tag{3.2.1}$$

If there exist non-decreasing constants $\{C_k\}_{k\in\mathbb{N}}$ such that $\sup_{x\in\mathcal{X}^{V_H}} |\widetilde{r}^H_j(x)| \leq C_{|H|}$ for all $j \in \mathcal{J}$ and $H \in \widehat{\mathscr{G}}_{*,1}$, then Assumption 1 holds with $C(k,T) := C_k$. Moreover, Assumption 3 holds trivially for all initial data because the mark spaces $\overline{\mathcal{K}} = \mathcal{K} = \{1\}$ are trivial, and the solution $X^{G,\xi}$ is a homogeneous Markov process (and Feller when $\mathcal{X}$ is finite and G has bounded maximal degree [65, Chapter I, Theorem 3.9]). Below we write out the jump rate functions of a variety of standard Markovian IPS. The contact process and SIR model are described with parameter $\lambda > 0$ and Glauber dynamics for the Ising model has parameter $\beta \in \mathbb{R}$ and no magnetic field with heat-bath transition rates:

Glauber Dynamics (Ising) [70, Definition 2]:

$$\widetilde{r}^H_{-2a_{\varnothing_H}}(a) := \frac{\exp\left(-\beta a_{\varnothing_H} \sum_{u\in\mathcal{N}_{\varnothing_H}} a_u\right)}{\exp\left(-\beta a_{\varnothing_H} \sum_{u\in\mathcal{N}_{\varnothing_H}} a_u\right) + \exp\left(\beta a_{\varnothing_H} \sum_{u\in\mathcal{N}_{\varnothing_H}} a_u\right)};$$

$$\text{SIR Model [54, Appendix B]:} \quad \widetilde{r}^H_1(a) := \mathbb{I}_{\{a_{\varnothing_H}=0\}} \lambda \sum_{u\in\mathcal{N}_{\varnothing_H}} a_u + \mathbb{I}_{\{a_{\varnothing_H}=1\}};$$

$$\text{Voter Model [23, page 2]:} \quad \widetilde{r}^H_1(a) := \frac{\mathbb{I}_{\{a_{\varnothing_H}=0\}}}{|\mathcal{N}_{\varnothing_H}|} \sum_{u\in\mathcal{N}_{\varnothing_H}} a_u,$$

$$\widetilde{r}^H_{-1}(a) := \frac{\mathbb{I}_{\{a_{\varnothing_H}=-1\}}}{|\mathcal{N}_{\varnothing_H}|} \sum_{u\in\mathcal{N}_{\varnothing_H}} (1-a_u);$$

$$\text{Contact Process [11, page 248]:} \quad \widetilde{r}^H_1(a) := \mathbb{I}_{\{a_{\varnothing_H}=0\}} \lambda \sum_{u\in\mathcal{N}_{\varnothing_H}} a_u,$$

$$\widetilde{r}^H_{-1}(a) := \mathbb{I}_{\{a_{\varnothing_H}=1\}}.$$

The above equations hold for all $a \in \mathcal{X}^{V_H}$, where recall that $\mathcal{X}$ is the state space of the IPS. The SIR model has state space $\mathcal{X} = \{0,1,2\}$. In the contact process and voter model, $\mathcal{X} = \{0,1\}$ while in the majority process and Glauber dynamics, $\mathcal{X} = \{\pm 1\}$.

It is easily verified that the models above all satisfy the standing assumption, Assumption 1 and Assumption 2. Note that the processes above have numerous variants that also satisfy these assumptions. For example, the SIR model can be replaced by the SEIR model, the voter model by the noisy voter model [44] and the contact process by the reversible contact process [80, page 218].

We now use the contact process introduced by Harris [47] to illustrate some interesting functionals of epidemic models that can be estimated using the empirical measure and neighborhood empirical distribution of Section 2.1.2. We show how framework can be used to describe simple extensions of the contact process to the contact process in (possibly random) heterogeneous environments and semi-Markovian IPS by once again using the contact process as an example.

Example 3.16 (**Functionals of the Contact Process).** Fix $0 < s < t < \infty$, $k \in \mathbb{N}$ and consider the following bounded, a.s. continuous functions $f : \mathcal{D}^{V_H} \to \mathbb{R}$:

(3.2.2)

$$f_1(x) = x_{\varnothing_H}(t), \quad f_2(x) = \int_s^t x_{\varnothing_H}(s')\,ds', \quad f_3(x) = \frac{1}{|V_H|-1}\sum_{u\in\mathcal{N}_{\varnothing_H}} x_u(t), \quad f_4(x) = \mathbb{I}_{\left\{\sum_{u\in\mathcal{N}_{\varnothing_H}}\geq k\right\}}.$$

Then if (G, X) is the trajectory of the contact process (assuming G and λ are such that (3.1.4) is strongly well-posed), then $\int f_1\pi^{G,X}$ represents the proportion of the population that is infected at time t; $\int f_2\pi^{G,X}$ represents the mean time spent infected during the interval $[s, t]$ by members of the population; $\int f_3\overline{\pi}^{G,X}_{\varnothing}$ represents the mean proportion of infected neighbors of each member of the population at time t; $\int f_4\overline{\pi}^{G,X}_{\varnothing}$ represents the proportion of the population possessing at least k infected neighbors. All of these functionals and more can be understood if we understand $\pi^{G,X}$ and $\overline{\pi}^{G,X}_{\varnothing}$.

Example 3.17 (**The Contact Process in a Heterogeneous Environment).** [36, Example 4.9] Suppose each individual v has an associated recovery rate $\kappa_v \geq 0$ and the level of transmission between two neighboring individuals u and v is proportional to $\overline{\kappa}_{\{u,v\}} \geq 0$. The model is then described by the $[\overline{\mathcal{K}}, \mathcal{K}] := [[0, M], [0, M]]$-graph $G = (V, E, \varnothing, \overline{\kappa}, \kappa)$ and, as before, $\mathcal{X} = \{0, 1\}$, $\mathcal{J} = \{-1, 1\}$, but now the standing assumption holds with $\bar{r}^H_1(t, x, \overline{\vartheta}, \vartheta) = \tilde{r}^H(x(t-), \overline{\vartheta}, \vartheta)$, for $H = (V_H, E_H, \varnothing_H) \in \widehat{\mathscr{G}}_{*,1}$, where $\tilde{r}^H_j : \mathcal{X}^{V_H} \times \overline{\mathcal{K}}^{E_H} \times \mathcal{K}^{V_H} \to \mathbb{R}$ is given by

$$\tilde{r}^H_1(z, \overline{\vartheta}, \vartheta) = \mathbb{I}_{\{z_{\varnothing_H}=0\}} \sum_{u\in\mathcal{N}_{\varnothing_H}(H)} \overline{\vartheta}_{\{\varnothing_H,u\}} z_u \quad \text{and} \quad \tilde{r}^H_{-1}(z) = \vartheta_{\varnothing_H}\mathbb{I}_{\{z_{\varnothing_H}=1\}} \qquad \text{for } z \in \mathcal{X}^{V_H}.$$

This model clearly satisfies Assumptions 1 and 3.

Example 3.18 (**Non-Markovian Contact Process).** [36, Example 4.10] We now introduce a class of non-Markovian contact processes that fall within our framework. (Such processes have recently been considered on the lattice in [34] under the rubric of renewal contact processes.) As before, let $\mathcal{X} = \{0, 1\}$ but now consider $G = (V, E, \varnothing, \{1\}, \kappa) \in \mathscr{G}_*[\{1\}, \mathcal{K}]$, where $\mathcal{K} := \mathcal{D}([-T, 0], \mathcal{X})$ for some fixed $T \in [0, \infty]$. The mark κ represents initial data that encodes the history of the process from time $-T$ up to time 0. Let f_I and f_R be the probability density functions of two positive continuous random variables that represent the respective times required for a particle to infect another particle or to recover from infection, and let h_I and h_R be the respective hazard rates: $h_I(t) = f_I(t)/\int_t^\infty f_I(s)\,ds$ and $h_R(t) = f_R(t)/\int_t^\infty f_R(s)\,ds$. In the Markov case, f_I and f_R are both exponential densities with respective means $1/\lambda$ and 1. Also, recall that $\Delta g(s) = g(s) - g(s-)$ for any càdlàg function g, and for any $U \subseteq V$ and $(\vartheta_U, x_U) \in (\mathcal{K} \times \mathcal{D})^U$ where we define $\Delta x_v(s) = 0$ for $s \in [-T, 0]$ and $\Delta\vartheta_v(s) = 0$ for $s > 0$, consider the functional $\tau_U : [-T, \infty) \times (\mathcal{K} \times \mathcal{D})^U \to [-T, \infty)$ given by

$$\tau(t, \vartheta_U, x_U) := \sup\{s \in (-T, t), v \in U : \Delta\vartheta_v(s) \neq 0 \text{ or } \Delta x_v(s) \neq 0\}, \quad t \geq -T, (\vartheta_U, x_U) \in (\mathcal{K}\times\mathcal{D})^U,$$

where we apply the convention that if the above set is empty then the corresponding supremum is $-T$. For any $U \subseteq V$, and at any time $t > 0$, when x is a realization of the actual process, $\tau_U(t, \vartheta_U, x_U)$ represents the last time that a particle in U experienced an event (infection or cure), and it takes the value $-T$ if no such events have occurred prior to time t. Then, the jump rate functions $\bar{r}^H_{\pm 1} : \mathbb{R}_+ \times \mathcal{D}^{V_H} \times \overline{\mathcal{K}}^{E_H} \times \mathcal{K}^{E_H} \to \mathbb{R}_+$ are given by

$$\begin{aligned} \bar{r}^H_1(t, x, \overline{\vartheta}, \vartheta) &:= \mathbb{I}_{\{x_{\varnothing_H}(t-)=0\}} \sum_{\substack{u \in \mathcal{N}_{\varnothing_H}(H): \\ x_u(t-)=1}} h_I((t - \tau(t, x_{\{\varnothing_H, u\}}, \vartheta_{\{\varnothing_H, u\}}))-), \\ \bar{r}^H_{-1}(t, x, \overline{\vartheta}, \vartheta) &:= \mathbb{I}_{\{x_{\varnothing_H}(t-)=1\}} h_R((t - \tau(t, x_{\varnothing_H}, \vartheta_{\varnothing_H}))-). \end{aligned}$$

Then Assumption 2 automatically holds. If $T < \infty$, then Assumption 1 holds, for example, when the distributions have infinite support and the hazard rate functions are locally bounded, i.e., $\sup_{s\in[0,t]}[h_I(s) + h_R(s)] < \infty$ for all $t \in \mathbb{R}_+$. When $T = \infty$, Assumption 1 holds under the more stringent conditions that the distributions have infinite support, the hazard rate functions are uniformly bounded on $\mathbb{R}_+$ and $\lim_{t\to\infty} h_I(t)$ and $\lim_{t\to\infty} h_R(t)$ both exist. Again note that in the Markov case, $h_I \equiv \lambda$ and $h_R \equiv 1$ are both constant.

Part II

Well-Posedness and Limits of Interacting Particle Systems

CHAPTER 4

Locality and Well-Posedness

4.1 Well-Posedness Results

In the case of *finite initial data*, that is, when the initial data consists of a finite (possibly unrooted) graph, Assumption 1 implies the rates are uniformly bounded and strong well-posedness is easily established via a simple recursive construction. For completeness, this result is summarized in Proposition 4.1 and its proof provided in Appendix D.1.

Proposition 4.1. *[36, Proposition 4.1] Suppose Assumption 1 holds. Then the SDE* (3.1.4) *is strongly well-posed for all finite initial data, and all solutions to the SDE for finite initial data are a.s. proper.*

We now turn to the main case of infinite graphs, with possibly unbounded maximal degree. On such graphs, since Assumption 1 allows the jump rates to grow with the degree of the vertex, which is typical of most IPS of interest as the examples in Section 3.2 demonstrate, jump rates can be unbounded. In this case well-posedness of even Markovian SDEs of the form (3.3) is subtle, and may in fact fail to hold. Indeed, in Section 4.3 we construct a simple Markovian IPS satisfying Assumption 1 that admits multiple strong solutions (with different laws) on certain graphs with super-exponential growth. Nevertheless, the following main result of this section shows that strong well-posedness does hold for IPS on a large class of so-called a.s. *finitely dissociable* graphs, whose precise definition is deferred to Section 4.2.3.

Theorem 4.2. *[36, Theorem 4.2] Suppose Assumption 1 holds and fix an a.s. finitely dissociable* $[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$*-random graph* (G, ξ) *(as specified in Definition 4.15). Then the SDE* (3.1.4) *is strongly well-posed for* (G, ξ) *in the sense of Definition 3.9.*

As shown in Proposition 4.18, Corollary 4.19 and Proposition 4.20, the class of a.s. finitely dissociable graphs includes GW trees with offspring distributions that have finite first moment, and hence, UGW trees whose offspring distributions have finite second moments, in addition to graphs with bounded maximal degree. Thus, in particular, when applied to the contact process, Theorem 4.2 provides rigorous justification that the contact process on GW trees, which was for example studied in [71, 11, 78, 48], is well defined.

Theorem 4.2 is a direct consequence of more general results established in Section 4.2, which are broad enough to cover applications for which Assumption 1 may fail, but the conclusion of Proposition 4.1 nevertheless holds, that is, for which the following condition is satisfied.

Assumption 1′. The family of jump rate functions $\mathbf{r} = \{r_j^{G,v}, j \in \mathcal{J}, G$ a $[\overline{\mathcal{K}}, \mathcal{K}]$-graph and $v \in V_G\}$ is such that the SDE (3.1.4) is strongly well-posed for all finite initial data.

Under Assumption 1′ (and hence, under Assumption 1), Proposition 4.10 establishes well-posedness of the SDE (3.1.4) for all initial data that "spatially localize" the SDE in the sense of Definition 4.4. Together with Proposition 4.22, which shows that (under Assumption 1) any finitely dissociable graph spatially localizes the SDE, this proves Theorem 4.2.

Lastly, we show that under Assumption 2, (strong) well-posedness of (3.1.4) for isomorphism class initial data is equivalent to (strong) well-posedness for marked graph initial data.

Remark 4.3. Remark 3.13 immediately implies that under Assumption 2, the (strong) well-posedness of (3.1.4) for (possibly random) isomorphism class initial data $\langle G, \xi \rangle$ is directly implied by the (strong) well-posedness of (3.1.4) for any (measurable) connected representative graph $(G, \xi) \in \langle G, \xi \rangle$. Therefore, most results in this section are stated for marked graphs rather than isomorphism classes. However, they can be directly extended to the isomorphism class case if we additionally assume that Assumption 2 holds.

4.2 Spatial Localization and the Proof of Well-Posedness

Throughout this section we continue to assume, without explicit mention, that the standing assumption holds. In Section 4.2.1 we introduce the notions of spatial localization and consistent spatial localization (see Definitions 4.4 and 4.7) and in Section 4.2.2 show that the SDE (3.1.4) is strongly well-posed for any marked graph that "spatially localizes" the IPS. In Section 4.2.3, we define the class of finitely dissociable graphs and provide examples of graphs in this class. In Section 4.2.4 we show that sequences of such graphs consistently spatially localize the SDE (3.1.4).

4.2.1 Spatial Localization and Consistent Spatial Localization

We now introduce the notion of spatial localization for a graph, which roughly speaking says that on any finite time interval the marginal of the associated IPS dynamics on any finite (random) subset U of the graph coincides with the marginal dynamics (on U) of an IPS on the induced (random) subgraph of G on a finite subset that contains U. When combined with Assumption $1'$, this will allow us in Section 4.2.2 to establish well-posedness for the class of graphs that spatially localize the SDE (3.1.4). The measurability and consistency properties required to state this property precisely are carefully spelled out in the following definition. Recall the notion of filtration-Poisson process pairs from Remark 3.6 and recall from Section 1.4 that $\Lambda_G := \{U \subset V_G : |U| < \infty\}$. Now, supposing that Assumption $1'$ holds, we introduce the notion of spatial localization. In what follows, we will deal with deterministic initial data (G, ξ), and recall the notion of a filtration-Poisson process pair $(\mathbb{F}, \mathbf{N}^G)$ from Remark 3.6 and its associated augmented filtration $\mathbb{H}^{\mathbf{N}^G} = \{\mathcal{H}_t^{\mathbf{N}^G}\}_{t \geq 0}$ introduced prior to Definition 3.5. Also recall from Section 1.4 that given $f \in \mathcal{D}$ and $t > 0$, $f[t]$ represents the truncated path that lies in $\mathcal{D}_t$.

Definition 4.4 (Spatial localization)**.** [36, Definition 5.1] A connected, deterministic $[\overline{\mathcal{K}}, \mathcal{K}]$-graph G is said to spatially localize the SDE (3.1.4) if for any filtration-Poisson process pair $(\mathbb{F}, \mathbf{N}^G)$ defined on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and $T \in \mathbb{R}_+$, there exists a mapping $\mathcal{S}_T(\cdot; G, \mathbf{N}^G) : \Lambda_G \times \Omega \to 2^{V_G}$ such that for every $U, U' \in \Lambda_G$,

$$\{\mathcal{S}_T(U; G, \mathbf{N}^G) \subseteq U'\} := \{\omega \in \Omega : \mathcal{S}_T(U; G, \mathbf{N}^G)(\omega) \subseteq U'\} \in \mathcal{F}_T^{\mathbf{N}^G}, \tag{4.2.1}$$

$$\mathbb{P}(\mathcal{S}_T(U; G, \mathbf{N}^G) \in \Lambda_G) = 1, \tag{4.2.2}$$

and the following properties hold:

1. for each $\mathcal{O} \in \Lambda_G$, $\mathcal{S}_T(\mathcal{O}; G, \mathbf{N}^G)(\omega) \supseteq \mathcal{O}$ for every $\omega \in \Omega$;

2. given any $\ell \in \mathbb{N}, \xi \in \mathcal{X}^G$ and $U, W \subset V$ such that $U \subseteq B_\ell(G) \subseteq W$, every $(\mathbb{F}, \mathbf{N}^G)$-weak solution $X^{G[W],\xi_W}$ to the SDE (3.1.4) for the initial data $(G[W], \xi_W)$ satisfies

$$X_U^{G[W],\xi_W}[T] = X_U^{B_\ell(G),\xi_{B_\ell(G)}}[T] \quad \text{a.s. on the event } \{\mathcal{S}_T(U; G, \mathbf{N}^G) \subseteq B_\ell(G)\}, \tag{4.2.3}$$

where $X^{B_\ell(G),\xi_{B_\ell(G)}}$ is the unique $\mathbf{N}^G_{B_\ell(G)}$-strong solution to (3.1.4) for $(B_\ell(G), \xi_{B_\ell(G)})$. In this case $\mathcal{S}_T(\cdot; G, \mathbf{N}^G)$ is said to be a localizing map for the SDE (3.1.4) on $(G, \mathbf{N}^G)$.

Remark 4.5. [36, Remark 5.2] Note that Assumption 1′ ensures that $X^{B_\ell(G),\xi_{B_\ell(G)}}$ in Definition 4.4 is well defined. Also, recall from Proposition 4.1 that Assumption 1′ is implied by Assumption 1, but is strictly more general and may hold even when Assumption 1 fails.

Remark 4.6. In Definition 4.4, we may omit one or both of the last two arguments of $\mathcal{S}_T$ when the graph and/or Poisson processes are clear from the context.

Given a fixed model (equivalently, a family of local jump rate functions $\bar{\mathbf{r}} = \{\bar{r}_j^H\}_{H \in \widehat{\mathcal{G}}_{*,1}, j \in \mathcal{J}}$ as specified in the standing assumption), we now introduce the concept of consistent spatial localization of a sequence of graphs, which will be used in the proofs of local convergence and the hydrodynamic limit in Sections 5.3 and 5.3.3, respectively.

Definition 4.7 (Consistent Spatial Localization)**.** A sequence of connected, deterministic $[\overline{\mathcal{K}}, \mathcal{K}]$-graphs $\{G_n\}_{n \in \mathbb{N}}$ defined on some common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to consistently spatially localize the SDE (3.1.4) if for every $n \in \mathbb{N}$ and corresponding filtration-Poisson process pair $(\mathbb{F}^n, \mathbf{N}^{G_n})$ on $(\Omega, \mathcal{F}, \mathbb{P})$, there exists a mapping $\mathcal{S}_T(\cdot; G_n, \mathbf{N}^{G_n}) : \Lambda_{G_n} \times \Omega \to 2^{V_{G_n}}$ that satisfies the following properties:

1. for each $n \in \mathbb{N}$, $\mathcal{S}_T(\cdot; G_n, \mathbf{N}^{G_n})$ is a localizing map for the SDE (3.1.4) on $(G_n, \mathbf{N}^{G_n})$ in the sense of Definition 4.4;

2. for every $n, n', \ell \in \mathbb{N}$ such that $B_\ell([G_{n,*}]) = B_\ell([G_{n',*}])$, the following property holds: for each set $U \subseteq B_\ell(G_n)$,

$$\mathcal{S}_T(U; G_n, \mathbf{N}^{G_n}) = \mathcal{S}_T(U; G_{n'}, \mathbf{N}^{G_{n'}}) \text{ a.s. on the event } \{\mathcal{S}_T(U; G_n, \mathbf{N}^{G_n}) \subseteq B_\ell(G_n)\} \cap \mathcal{I},$$

where $\mathcal{I}$ is the $\mathcal{H}_T^{\mathbf{N}^{G_{n'}}} \vee \mathcal{H}_T^{\mathbf{N}^{G_n}}$ measurable event given by

$$\mathcal{I} := \{B_\ell([G_{n',*}], \mathbf{N}^{G_{n'}}) = B_\ell([G_{n,*}], \mathbf{N}^{G_n})\}.$$

In this case $\{\mathcal{S}_T(\cdot; G_n, \mathbf{N}^{G_n})\}_{n\in\mathbb{N}}$ is said to be a consistent sequence of localizing maps for the SDE (3.1.4) on $\{(G_n, \mathbf{N}^{G_n})\}_{n\in\mathbb{N}}$.

Note that this definition is similar to the definition of consistent spatial localization from [36], but with a few differences. Notably, we now set $\mathcal{I}$ to be the event on which the truncations are equal rather than isomorphic. This impacts the equality in the first display of property 2 of consistent spatial localization. We will nevertheless continue to cite [36] for results which invoke consistent spatial localization as the modifications needed to obtain our results from the results of [36] are trivial.

In abstract, consistent spatial localization may not seem much more restrictive than spatial localization. Indeed, if G_n, $n \in \mathbb{N}$ each have distinct roots, then consistent spatial localization is equivalent to the spatial localization of G_n for each $n \in \mathbb{N}$. However, if $\{G_n\}_{n\in\mathbb{N}}$ is a convergent sequence of graphs in $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K}]$, then by Lemma 2.22, for any ℓ and sufficiently large n, n', $B_\ell([G_{n,*}]) = B_\ell([G_{n',*}])$ so condition 2 of Definition 4.7 becomes relevant.

It is not hard to show that even under Assumption 2, consistent spatial localization is not a property of isomorphism classes. However, even though consistent spatial localization is not a class property, it is possible to define a notion of consistent spatial localization for isomorphism classes. It should be noted that these definitions apply even when Assumption 2 is violated.

Definition 4.8. An isomorphism class $\langle G\rangle \in \mathscr{G}_*[\overline{\mathcal{K}}, \mathcal{K}]$ is said to spatially localize the SDE (3.1.4) whenever each connected representative $[\overline{\mathcal{K}}, \mathcal{K}]$-graph G does the same. A sequence of isomorphism classes $\{\langle G_n\rangle\}_{n\in\mathbb{N}}$ in $\mathscr{G}_*[\overline{\mathcal{K}}, \mathcal{K}]$ is said to consistently spatially localize the SDE (3.1.4) if every sequence of connected representative graphs $G_n \in \langle G_n\rangle, n \in \mathbb{N}$, does the same.

In Proposition 4.22 we show that any sequence of isomorphism classes that is a subset of the class of so-called finitely dissociative isomorphism classes (see Definition 4.15) consistently spatially localizes the SDE (3.1.4). Note that while spatial localization is a class property, the notion of a localizing map given in Definition 4.4 applies to a given graph, and not its isomorphism class.

Remark 4.9. [36, Remark 5.6] Definitions 4.4 and 4.7 are abstract properties that easily extend to a more general class of graph-indexed jump processes $X^{G,\xi}$ that satisfy a different Poisson-driven SDE from (3.1.4), as long as Assumption 1′ still holds for that SDE. For instance, they could apply to IPS such as the exclusion process in which particles may experience simultaneous jumps, where it would be more natural to index $\mathbf{N}^G$ by the edges, rather than the vertices, of G, or processes with both types of jumps.

4.2.2 Well-posedness on spatially localizing graphs

Proposition 4.10. *[36, Proposition 5.7] Suppose that Assumption 1′ holds, and that (G, ξ) is a connected $[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-random graph such that G a.s. spatially localizes the SDE* (3.1.4)*. Then the SDE* (3.1.4) *is strongly well-posed for the initial data (G, ξ).*

Proof. By Lemma 3.10, it suffices to prove that (3.1.4) is strongly well-posed for a.s. every realization of (G, ξ). Therefore, for the remainder of the proof we assume without loss of generality that (G, ξ) is a connected, deterministic $[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-graph where G spatially localizes the SDE (3.1.4).

We now explicitly construct a strong solution to the SDE (3.1.4) for (G, ξ). Let $\mathbf{N}^G = \{\mathbf{N}^G_v\}_{v \in V_G}$ be a collection of i.i.d. Poisson processes on $\mathbb{R}^2_+ \times \mathcal{J}$ with intensity $\mathrm{Leb}^2 \otimes \#_{\mathcal{J}}$. Let $\mathbb{F} := \{\mathcal{F}_t\} := \mathbb{H}^{\mathbf{N}^G}$ be the associated filtration (as defined prior to Definition 3.5). It is clear from Definition 3.3 that $(G, \mathbf{N}^G)$ is an $\mathbb{F}$-driving noise. So by Remark 3.6, $(\mathbb{F}, \mathbf{N}^G)$ is a filtration-Poisson process pair. Fix $T \in \mathbb{R}_+$, and let $\mathcal{S}_T(\cdot) := \mathcal{S}_T(\cdot; G, \mathbf{N}^G)$ be a corresponding localizing map, which exists due to our assumption that G spatially localizes the SDE. For notational conciseness, let $B_m := B_m(G)$ for each $m \in \mathbb{N}$, and additionally fix ξ and omit the dependence on ξ from the superscript, with the understanding that the mark is always the restriction of ξ to the corresponding graph in the superscript. Furthermore, for any $\ell \in \mathbb{N}$, recalling that by Assumption 1′ the SDE (3.1.4) is strongly well-posed for the finite data (B_ℓ, ξ_{B_ℓ}), we let X^{B_ℓ} denote the corresponding $\mathbf{N}^G_{B_\ell}$-strong solution to (3.1.4). By Definition 4.4, for each $m \in \mathbb{N}$, there exists a random finite set $\mathcal{O}_m := \mathcal{S}_T(B_m)$ such that for any $\ell, \ell' \in \mathbb{N}$, $\ell' \geq \ell \geq m$, by (4.2.1) and applying (4.2.3) with $W = B_{\ell'}$ and $U = B_m$,

$$X^{B_\ell}_{B_m}[T] = X^{B_{\ell'}}_{B_m}[T] \quad \text{on the event } \{\mathcal{O}_m \subseteq B_\ell\} \in \mathcal{F}_T.$$

For $m \in \mathbb{N}$, we define the $\mathcal{D}^{B_m}_T$-valued random element X^m and the random integer M_m as follows:

$$X^m[T] := \lim_{\ell\to\infty} X^{B_\ell}_{B_m}[T] \quad \text{and} \quad M_m := \inf\{\ell \geq m : \mathcal{O}_m \subseteq B_\ell\}. \tag{4.2.4}$$

Since $\bigcup_{\ell\in\mathbb{N}} B_\ell = V_G$ and $\mathcal{O}_m$ is a.s. finite, X^m is well defined on a set of full measure and M_m is a.s. finite. Moreover, since X^{B_ℓ} is $\mathbb{F}$-adapted for every $\ell \in \mathbb{N}$ and $\mathbb{F}$ is complete, it follows that $X^m[T]$ is also $\mathbb{F}$-adapted. The last two displays together then imply that a.s.,

$$X^m[T] = \lim_{\ell\to\infty} X^{B_\ell}_{B_m}[T] = X^{B_{M_m}}_{B_m}[T]. \tag{4.2.5}$$

Furthermore, clearly the sequence $(X^m[T])_{m\in\mathbb{N}}$ is consistent: for any $m' > m$, a.s.,

$$X^{m'}_{B_m}[T] = \lim_{\ell\to\infty} (X^{B_\ell}_{B_{m'}})_{B_m}[T] = \lim_{\ell\to\infty} X^{B_\ell}_{B_m}[T] = X^m[T] = X^m_{B_m}[T], \tag{4.2.6}$$

where the first and third equalities invoke (4.2.5) and the remaining equalities hold trivially. Now, for every $v \in V_G$, there exists an integer $m_v \in \mathbb{N}$ such that $v \in B_{m_v}$, and the last display shows that $X^{m_v}_v[T] = X^{m'}_v[T]$ a.s. when $m' \geq m_v$. Because V_G is countable and $\mathbb{F}$ is complete, we can define the $\mathbb{F}$-adapted $\mathcal{D}^G_T$-random element $X[T]$ by setting

$$X_v[T] := \lim_{m\to\infty} X^m_v[T] = X^{m_v}_v[T], \quad v \in V_G, \tag{4.2.7}$$

on the set of measure one where the latter limits exist, and setting $X_v[T] \equiv \xi_v, v \in V_G$, on the complement.

To show that the $X[T]$ thus constructed is a $\mathbf{N}^G$-strong solution to the SDE (3.1.4) on $[0,T]$, fix $v \in V_G$ and define $\bar{m}_v := \max\{m_u : u \in \mathrm{cl}_v\}$. Then $\mathrm{cl}_v \subset B_{\bar{m}_v}$ and from (4.2.5)–(4.2.7), it follows that for $\ell \in \mathbb{N}$,

$$X_{\mathrm{cl}_v}[T] = X^{\bar{m}_v}_{\mathrm{cl}_v}[T] = X^{B_\ell}_{\mathrm{cl}_v}[T] \quad \text{on} \quad A_\ell(v) := \{\ell \geq M_{\bar{m}_v}\}. \tag{4.2.8}$$

Since X^{B_ℓ} is a $\mathbf{N}^G_{B_\ell}$-strong solution to the SDE (3.1.4), we obtain a.s., for $t \in [0,T]$,

$$X_v(t)\mathbb{I}_{A_\ell(v)} = X^{B_\ell}_v(t)\mathbb{I}_{A_\ell(v)}$$

$$= \left(\xi_v + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r\le r_j^{G[B_\ell],v}(s,X^{B_\ell})\right\}} \mathbf{N}_v^G(ds,dr,dj) \right) \mathbb{I}_{A_\ell(v)}.$$

By (3.1.1) of the standing assumption and (4.2.8), it follows that with $H := G[\mathrm{cl}_v]$,

$$r_j^{G[B_\ell],v}(s,X^{B_\ell}) = \bar{r}_j^H\left(s, X^{B_\ell}_{\mathrm{cl}_v}, (\overline{\kappa}_e)_{e\in E_H}, (\kappa_u)_{u\in\mathrm{cl}_v}\right) = \bar{r}_j^H\left(s, X_{\mathrm{cl}_v}, (\overline{\kappa}_e)_{e\in E_H}, (\kappa_u)_{u\in\mathrm{cl}_v}\right) = r_j^{G,v}(s,X),$$

for every $j \in \mathcal{J}$ and $s \in \mathbb{R}_+$ on the event $A_\ell(v)$. Hence, we have

$$X_v(t)\mathbb{I}_{A_\ell(v)} = \left(\xi_v + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r\le r_j^{G,v}(s,X)\right\}} \mathbf{N}_v^G(ds,dr,dj) \right) \mathbb{I}_{A_\ell(v)}.$$

Since $M_{\bar{m}_v}$ is finite, taking the limit as $\ell \to \infty$ shows that a.s.,

$$X_v(t) = \xi_v + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r\le r_j^{G,v}(s,X)\right\}} \mathbf{N}_v^G(ds,dr,dj).$$

Thus, we have proved the existence of a $\mathbf{N}^G$-strong solution to (3.1.4) on any interval $[0,T]$.

We now turn to the proof of pathwise uniqueness. Suppose that $\widetilde{X}$ and $\widetilde{X}'$ are any two $(\mathbb{F}, \mathbf{N}^G)$-weak solutions to the SDE (3.1.4) for (G,ξ). Since G spatially localizes the SDE, for any $\ell > m \in \mathbb{N}$, and M_m defined as in (4.2.4), invoking (4.2.1) and applying (4.2.3), with $W = V$ and $U = B_m$, to both weak solutions $\widetilde{X}$ and $\widetilde{X}'$, we obtain a.s. on the event $\{\ell \ge M_m\} = \{\mathcal{S}_T(U; G, \mathbf{N}^G) \subseteq B_\ell(G)\} \in \mathcal{H}_T^{\mathbf{N}^G}$,

$$\widetilde{X}_{B_m}[T] = X^{B_\ell(G)}_{B_m}[T] = \widetilde{X}'_{B_m}[T],$$

where recall from Definition 4.4 that $X^{B_\ell(G)}$ is the unique $\mathbf{N}^G_{B_\ell(G)}$-strong solution to (3.1.4) for the initial data $B_\ell(G,\xi)$. Taking the limit as $\ell \to \infty$ and using the almost sure finiteness of M_m, it follows that $\widetilde{X}_{B_m}[T] = X_{B_m}[T]$ a.s. for every $m \in \mathbb{N}$, which in turn shows that $X[T] = \widetilde{X}[T]$ a.s.

Since T is arbitrary for both existence and pathwise uniqueness, $X[T]$, $T > 0$, are consistent. So there exists a.s. a unique pathwise extension X of the strong solution to all of $[0,\infty)$. This concludes the proof. □

Remark 4.11. Most of the proof of Proposition 4.10 also extends to more general IPS. For instance, if an IPS has simultaneous jumps, then the proof will hold given Assumption 1′ and spatial localization (see Remark 4.9) except for the verification that $X := \lim_{m\to\infty} X^m$ solves (3.1.4). Instead of the latter, one would have to prove that if X^{B_ℓ} satisfies the SDE defining the new model on the finite graph $B_\ell(G)$, and X^m is as defined in (4.2.5), then the limit $X = \lim_{m\to\infty} X_m$ also satisfies that SDE on the infinite graph G.

Remark 4.12. [36, Remark 5.8] Note that if G is a disconnected graph and if $\mathscr{C}_\varnothing(G)$ spatially localizes the SDE (3.1.4), then Proposition 4.10 implies that the SDE is strongly well-posed for the initial data $\mathscr{C}_\varnothing(G,\xi)$. Then by the form of the equation (3.1.4) and the standing assumption, for any weak solution $X^{G,\xi}$ to (3.1.4) for the initial data (G,ξ), $X^{G,\xi}_{V_{\mathscr{C}_\varnothing(G)}} = X^{\mathscr{C}_\varnothing(G,\xi)}$ a.s. where $X^{\mathscr{C}_\varnothing(G,\xi)}$ is the unique strong solution to (3.1.4) for the initial data $\mathscr{C}_\varnothing(G,\xi)$ and driven by the same driving noise as $X^{G,\xi}$.

4.2.3 Finitely Dissociable Graphs

Definition of Finitely Dissociable Graphs

We now introduce the class of finitely dissociable graphs, which are defined in terms of an inhomogeneous site percolation on the graph. Recall the definition of the measure space $\mathscr{N}(0,T]$ from Section 1.4.

Definition 4.13. [36, Definition 5.9] For any $0 < T < \infty$, let (G,ζ) be a (possibly disconnected) $[\{1\},\mathscr{N}(0,T]]$-random graph. Fix $0 \le t_1 < t_2 \le T$, and set $R_v = R_v(t_1,t_2) := \mathbb{I}_{\{\zeta_v(t_1,t_2]>0\}}$ for $v \in V_G$. Then the percolated graph $\mathrm{perc}_{t_1,t_2}(G,\zeta)$ is defined to be the (possibly disconnected and random) subgraph of G induced by the vertex set $\{v \in V_G : R_v = 1\}$. When $t_1 = 0$, we write $\mathrm{perc}_{t_2}(G,\zeta) := \mathrm{perc}_{0,t_2}(G,\zeta)$.

In the percolation we will refer to vertices $v \in V_G$ with $R_v = 1$ as active, and those with $R_v = 0$ as inactive. In our application of Definition 4.13, the vertex marks ζ of the graph G will be realizations of (Poisson) point processes.

Definition 4.14. [36, Definition 5.10] Given $0 < \Delta \le T < \infty$, we say a $[\{1\},\mathscr{N}(0,T]]$-graph (G,ζ) Δ-*dissociates* if all connected components of $\mathrm{perc}_\Delta(G,\zeta)$ are finite.

For any graph G, let $(G,\mathbf{N}^G)$ be a driving noise in the sense of Definition 3.3. Also, for $T \in (0,\infty)$, let $\mathbf{N}^{G,T} = (\mathbf{N}^{G,T}_v)_{v\in V_G}$ be the collection of point processes on $[0,T]$ defined by

$$\mathbf{N}^{G,T}_v(t_1,t_2] := \mathbf{N}^G_v\left((t_1,t_2] \times (0, C(|\mathrm{cl}_v|,T)] \times \mathcal{J}\right), \quad \text{for } 0 < t_1 < t_2 \le T < \infty, v \in V_G, \tag{4.2.9}$$

where C is the function from Assumption 1. Note that the union of the events of $\mathbf{N}_v^{G,T}, v \in V_G$, almost surely contains the set of discontinuities of any weak solution $X^{G,\xi}[T]$ of the SDE (3.1.4).

Definition 4.15 (Finite dissociability)**.** [36, Definition 5.11] A deterministic graph G is said to be *finitely dissociable* if for any $T \in (0, \infty)$, there exists $\Delta > 0$ such that $(G, \mathbf{N}^{G,T})$ Δ-dissociates a.s.. We call Δ a *G-dissociation number*. If G is a marked graph, we will say G is finitely dissociable if and only if the corresponding unmarked graph $[G]$ is finitely dissociable.

Remark 4.16. [36, Remark 5.12] In the percolated graph $\mathrm{perc}_\Delta(G, \mathbf{N}^{G,T})$ described in Definition 4.15, setting

$$\bar{C}(k,T) := |\mathcal{J}| C(k,T), \quad \forall k \in \mathbb{N}, T \in \mathbb{R}_+, \tag{4.2.10}$$

each vertex v is removed from G independently with a probability $\exp\Big(-\Delta \bar{C}(|\mathrm{cl}_v|, T)\Big)$, which is decreasing in the degree of the vertex v.

Remark 4.17. [36, Remark 5.13] Finite dissociability is a "class property" in that it depends only on the isomorphism class $\langle G \rangle \in \mathscr{G}_*$ of the graph G, and not on the particular representative or choice of the driving noise. Indeed, if $G_1 \cong G_2$, are connected graphs, then for any fixed $T \in (0, \infty)$ and Poisson processes $\mathbf{N}^{G_i,T}$, $i = 1, 2$, constructed as in (4.2.9), it is easy to construct a coupling $(\widetilde{\mathbf{N}}^{G_1,T}, \widetilde{\mathbf{N}}^{G_2,T})$ of $(\mathbf{N}^{G_1,T}, \mathbf{N}^{G_2,T})$ in which $\widetilde{\mathbf{N}}^{G_i,T} \overset{(\mathrm{d})}{=} \mathbf{N}^{G_i,T}$ for $i = 1, 2$, such that $(G_1, \widetilde{\mathbf{N}}^{G_1,T}) \cong (G_2, \widetilde{\mathbf{N}}^{G_2,T})$ a.s.. Hence, for $\Delta \in [0, T]$,

$$\mathrm{perc}_\Delta(G_1, \mathbf{N}^{G_1,T}) \overset{(\mathrm{d})}{=} \mathrm{perc}_\Delta(G_1, \widetilde{\mathbf{N}}^{G_1,T}) \cong \mathrm{perc}_\Delta(G_2, \widetilde{\mathbf{N}}^{G_2,T}) \overset{(\mathrm{d})}{=} \mathrm{perc}_\Delta(G_2, \mathbf{N}^{G_2,T}),$$

with the equivalence holding a.s.. Since the finite dissociability of G_i only depends on the probability that $\mathrm{perc}_\Delta(G_i, \mathbf{N}^{G,T})$ has an infinite component for some $\Delta > 0$, and the existence of an infinite component is isomorphism invariant, this shows that the finite dissociability property is also invariant with respect to graph isomorphisms. Thus, the statement "$\langle G \rangle \in \mathscr{G}_*$ is (or is not) finitely dissociable" is well defined.

Examples of Finitely Dissociable Graphs

We now show that the class of (almost surely) finitely dissociable graphs encompasses several interesting classes of graphs of interest in applications, including lattices and regular, GW and

UGW trees (see Definitions 2.13 and 2.16).

Proposition 4.18. *[36, Proposition 5.15] If $\rho \in \mathcal{P}(\mathbb{N}_0)$ has a finite first moment, that is, $\sum_{k\in\mathbb{N}} k\rho_k < \infty$, then the GW tree $\mathcal{T} := GW(\rho)$ is a.s. finitely dissociable.*

Proof. Fix $T < \infty$, $\Delta \in (0,T)$, and let $\mathcal{T}^\Delta := \mathrm{perc}_\Delta(\mathcal{T}, \mathbf{N}^{\mathcal{T},T})$ and $R_v = R_v^\Delta := \mathbb{I}_{\{\mathbf{N}_v^{\mathcal{T},T}(0,\Delta]>0\}}$, $v \in V$, be as in Definition 4.13. Recall that $\mathcal{T}^\Delta$ is precisely the subgraph of $\mathcal{T}$ induced by active vertices in V. Also, recall that for $v \in V_{\mathcal{T}^\Delta}$, $\mathscr{C}_v(\mathcal{T}^\Delta)$ denotes the connected component of $\mathcal{T}^\Delta$ containing v, with v as its root. With a small abuse of notation, we extend the definition of $\mathscr{C}_v(\mathcal{T}^\Delta)$ to all $v \in V$ by setting $\mathscr{C}_v(\mathcal{T}^\Delta)$ to be the 1-vertex graph $\{v\}$ for $v \in V \setminus V_{\mathcal{T}^\Delta}$. In addition, for $v \in V$, let $\mathcal{T}_v$ and $\mathcal{T}_v^\Delta$ be the restrictions of $\mathcal{T}$ and $\mathcal{T}^\Delta$, respectively, to the set containing v and its descendants in $\mathcal{T}$. By the self-similarity of the GW tree, for each $v \in V$, $\mathrm{Law}(\mathcal{T}, (R_w)_{w\in V}) = \mathrm{Law}(\mathcal{T}_v, (R_w)_{w\in V_{\mathcal{T}_v}})$, and hence, $\mathrm{Law}(\mathcal{T}_\varnothing^\Delta | R_\varnothing = 1) = \mathrm{Law}(\mathcal{T}_v^\Delta | R_v = 1)$. For $v \in V$, $R_v = 0$ implies $\mathscr{C}_v(\mathcal{T}_v^\Delta)$ consists of a single isolated vertex. Hence, $|\mathscr{C}_\varnothing(\mathcal{T}^\Delta)| < \infty$ a.s. if and only if $|\mathscr{C}_v(\mathcal{T}_v^\Delta)| < \infty$ a.s. for all $v \in V$, in which case all connected components of $\mathcal{T}^\Delta$ must be a.s. finite. Thus, it suffices to prove that $|\mathscr{C}_\varnothing(\mathcal{T}^\Delta)| < \infty$ a.s..

Since the percolation probability at a site or vertex depends on its degree via the dependence on $C(k,T)$ in (4.2.9), to bound the size of $\mathscr{C}_\varnothing(\mathcal{T}^\Delta)$, we couple $\mathscr{C}_\varnothing(\mathcal{T}^\Delta)$ to a larger set obtained from a simpler percolation that only removes vertices from the odd generations of $\mathcal{T}$. To this end, for any rooted tree $\mathcal{T}$ and $n \in \mathbb{N}$, let $L_n(\mathcal{T}) := \{v \in V : d_{\mathcal{T}}(\varnothing, v) = n\}$ denote the set of vertices in the nth generation. Define the half-percolated forest $\widehat{\mathcal{T}}^\Delta := \mathrm{hperc}_\Delta(\mathcal{T}, \mathbf{N}^{\mathcal{T},T})$ to be the subgraph of $\mathcal{T}$ induced by the vertex set $\{v \in V : R_v = 1 \text{ or } d_{\mathcal{T}}(\varnothing, v) \text{ is even}\}$. Then let $\widehat{\mathcal{T}}_\varnothing^\Delta := \mathscr{C}_\varnothing(\widehat{\mathcal{T}}^\Delta)$ be the subtree of $\widehat{\mathcal{T}}^\Delta$ that contains the root (note that the root always belongs to $\widehat{\mathcal{T}}^\Delta$). Clearly, $|\mathscr{C}_\varnothing(\mathcal{T}^\Delta)| \le |\widehat{\mathcal{T}}_\varnothing^\Delta|$. Thus, to prove the proposition it clearly suffices to show that for all sufficiently small $\Delta > 0$,

$$\lim_{n\to\infty} \mathbb{E}\left[\left|L_{2n}(\widehat{\mathcal{T}}_\varnothing^\Delta)\right|\right] = 0, \tag{4.2.11}$$

since then $\mathbb{P}(|\widehat{\mathcal{T}}_\varnothing^\Delta| = \infty) = \mathbb{P}\left(\left|L_{2n}(\widehat{\mathcal{T}}_\varnothing^\Delta)\right| > 0 \text{ for all } n \in \mathbb{N}\right) \le \inf_{n\in\mathbb{N}} \mathbb{P}\left(\left|L_{2n}(\widehat{\mathcal{T}}_\varnothing^\Delta)\right| > 0\right) = 0$.

To prove (4.2.11), choose $n \in \mathbb{N}_0$ with $L_{2n}(\widehat{\mathcal{T}}_\varnothing^\Delta) \neq \emptyset$, and fix $v \in L_{2n}(\widehat{\mathcal{T}}_\varnothing^\Delta)$. Recall that for any tree $\mathcal{T}$ and $v \in V$, $c_v(\mathcal{T})$ denotes the collection of children of v in $\mathcal{T}$ and $\pi_v(\mathcal{T})$ is the parent of v

in $\mathcal{T}$ whenever $v \neq \varnothing$, and note that $c_v(\widehat{\mathcal{T}}^{\Delta}_{\varnothing}) = c_v(\widehat{\mathcal{T}}^{\Delta}) = \{u \in c_v(\mathcal{T}) : R_u = 1\}$. Also, observe that

(4.2.12) $$c_{c_v(\widehat{\mathcal{T}}^{\Delta}_{\varnothing})}(\widehat{\mathcal{T}}^{\Delta}_{\varnothing}) := \bigcup_{u \in c_v(\widehat{\mathcal{T}}^{\Delta}_{\varnothing})} c_u(\widehat{\mathcal{T}}^{\Delta}_{\varnothing}) = \{w \in c_u(\mathcal{T}) : u \in c_v(\mathcal{T}) \text{ and } R_u = 1\},$$

where the equality uses the fact that the half-percolation does not remove any vertices from even generations of the tree. Now, for each $w \in L_{2n}(\mathcal{T}) = L_{2n}(\widehat{\mathcal{T}}^{\Delta})$, let $Z_{\Delta,w}$ be the number of grandchildren of w in $\widehat{\mathcal{T}}^{\Delta}$. Then, since $v \in L_{2n}(\widehat{\mathcal{T}}^{\Delta}_{\varnothing})$, $Z_{\Delta,v}$ is also the number of grandchildren of v in $\widehat{\mathcal{T}}^{\Delta}_{\varnothing}$ and hence, by (4.2.12),

(4.2.13)
$$Z_{\Delta,v} := \left|c_{c_v(\widehat{\mathcal{T}}^{\Delta}_{\varnothing})}(\widehat{\mathcal{T}}^{\Delta}_{\varnothing})\right| = \left|c_{c_v(\widehat{\mathcal{T}}^{\Delta}_{\varnothing})}(\mathcal{T})\right| \text{ is } \sigma\left(\{c_u(\mathcal{T})\}_{u \in \{v\} \cup c_v(\mathcal{T})}, \{R_u\}_{u \in c_v(\mathcal{T})}\right)\text{-measurable.}$$

Furthermore, $L_{2n}(\mathcal{T})$ and $L_{2n}(\widehat{\mathcal{T}}^{\Delta}_{\varnothing})$ are both measurable with respect to $\mathcal{H}_n := \sigma(B_{2n}(\mathcal{T}), \{R_w\}_{w \in B_{2n-1}(\mathcal{T})})$. Let $A \in \mathcal{H}_n$ be an atomic event, that is, $A = \{B_{2n}(\mathcal{T}) = \widetilde{\mathcal{T}}, \{R_w\}_{w \in B_{2n-1}(\mathcal{T})} = \{r_w\}_{w \in B_{2n-1}(\mathcal{T})}\}$ for some k-generation tree $\widetilde{\mathcal{T}}$, with $k \leq 2n$ and some $(r_w)_{w \in B_{2n-1}(\mathcal{T})} \in \{0,1\}^{B_{2n-1}(\mathcal{T})}$. Then, conditioned on A, the collection of random variables

$$\left(\{c_u(\mathcal{T})\}_{u \in \{v\} \cup c_v(\mathcal{T})}, \{R_u\}_{u \in c_v(\mathcal{T})}\right)_{v \in L_{2n}(\mathcal{T})}$$

is equal in distribution to $|L_{2n}(\mathcal{T})|$ independent copies of $\left(\{c_u(\mathcal{T})\}_{u \in \{\varnothing\} \cup c_{\varnothing}(\mathcal{T})}, \{R_u\}_{u \in c_{\varnothing}(\mathcal{T})}\right)$. Together with (4.2.12), this implies that $\gamma_{\Delta} = \gamma_{\Delta,v} := \operatorname{Law}(Z_{\Delta,v}|A)$ does not depend on A or the specific choice of v in $L_{2n}(\mathcal{T})$. Moreover, conditioned on A, (4.2.12) implies

$$Z_{\Delta,v} = \sum_{u \in c_v(\mathcal{T})} R_u |c_u(\mathcal{T})| \overset{(d)}{=} \sum_{u \in c_{\varnothing}(\mathcal{T})} R_u |c_u(\mathcal{T})| = Z_{\Delta,\varnothing}.$$

Thus, conditioned on $\mathcal{H}_n$, $\{Z_{\Delta,v}\}_{v \in L_{2n}(\mathcal{T})}$ is equal in distribution to $L_{2n}(\mathcal{T})$ i.i.d. copies of $Z_{\Delta,\varnothing}$. Also, for $v \in L_{2n}(\mathcal{T})$, by the assumption that ρ has finite mean

$$\mathbb{E}\left[Z_{\Delta,v}|\mathcal{H}_n\right] = \mathbb{E}\left[Z_{\Delta,\varnothing}\right] \leq \mathbb{E}\left[\sum_{u \in c_{\varnothing}(\mathcal{T})} |c_u(\mathcal{T})|\right] = \left(\sum_{k=0}^{\infty} k\rho(k)\right)^2 < \infty.$$

Since $L_{2n}(\widehat{\mathcal{T}}^{\Delta}_{\varnothing})$ is $\mathcal{H}_n$-measurable and $\gamma_{\Delta,v} = \gamma_{\Delta}$ for all $v \in L_{2n}(\widehat{\mathcal{T}}^{\Delta}_{\varnothing})$, a recursive calculation shows

$$\mathbb{E}\left[\left|L_{2n+2}(\widehat{\mathcal{T}}^{\Delta}_{\varnothing})\right|\right] = \mathbb{E}\left[\sum_{v \in L_{2n}(\widehat{\mathcal{T}}^{\Delta}_{\varnothing})} \mathbb{E}\left[Z_{\Delta,v}|\mathcal{H}_n\right]\right] = \mathbb{E}\left[|L_{2n}(\widehat{\mathcal{T}}^{\Delta}_{\varnothing})|\right]\mathbb{E}\left[Z_{\Delta,\varnothing}\right] = \mathbb{E}\left[Z_{\Delta,\varnothing}\right]^{n+1}.$$

Thus, to show (4.2.11), it suffices to prove that for sufficiently small $\Delta > 0$, $\mathbb{E}\left[Z_{\Delta,\varnothing}\right] < 1$. However, note that for all $\Delta > 0$, $\mathbb{E}\left[Z_{\Delta,\varnothing}\right] < \infty$. Furthermore, the definition of Δ-dissociation clearly implies $\lim_{\Delta\to 0} Z_{\Delta,\varnothing} = 0$ a.s., so the Lebesgue dominated convergence theorem implies $\lim_{\Delta\to 0}\mathbb{E}\left[Z_{\Delta,\varnothing}\right] = 0$, which concludes the proof. □

In light of Definition 2.16, this immediately implies the corresponding result for UGW trees (see Definition 2.16).

Corollary 4.19. *[36, Corollary 5.16] If $\mathcal{T} \sim UGW(\rho)$, where $\rho \in \mathcal{P}(\mathbb{N}_0)$ has a finite second moment, then $\mathcal{T}$ is a.s. finitely dissociable.*

Proof. The fact that ρ has a finite second moment implies that $\widehat{\rho}$ in Definition 2.16 has a finite first moment. Let $\mathcal{T}_v$ denote the descendant tree of T rooted at v. Then for each $v \neq \varnothing$, $\mathcal{T}_v$ is a GW($\widehat{\rho}$)-tree, so by Proposition 4.18, for sufficiently small $\Delta > 0$, $\mathscr{C}_v(\mathrm{perc}_\Delta(\mathcal{T}_v, \mathbf{N}^{\mathcal{T}_v,T}))$ is a.s. finite. Since the Δ-percolated descendant tree of every non-root vertex is finite, it follows that $\mathcal{T}$ a.s. Δ-dissociates. □

For completeness, we also show that graphs with bounded maximum degree (such as infinite lattices) are also finitely dissociable.

Proposition 4.20. *[36, Proposition 5.17] Let $G = (V, E)$ be a graph with finite maximum degree: $d^* := \sup_{v\in V} |\mathcal{N}_v| < \infty$. Then G is finitely dissociable.*

Proof. Fix $0 < T < \infty$ and $\Delta \in (0, T]$. Then by Definitions 4.13 and 4.2.9, each vertex in G is inactive independently with probability

$$p_v = \exp\left(-\Delta\bar{C}(|\mathrm{cl}_v|, T)\right) \geq \exp\left(-\Delta\bar{C}_{d^*+1,T}\right) := p_{\Delta,T},$$

where the inequality invokes (4.2.10) and the definition of d^*. Therefore, the probability that $(G, \mathbf{N}^{G,T})$ Δ-dissociates is greater than or equal to the probability that G fails to percolate with respect to a standard site percolation in which each vertex is independently removed with probability $p_{\Delta,T}$. For any $T < \infty$, $\lim_{\Delta\to 0} p_{\Delta,T} = 1$, but as is well known, the critical probability (i.e., the probability p_c such that G fails to percolate a.s. when vertices are independently removed with probability $p > p_c$) is strictly less than 1 (see [45, Equation (0.3)]). Thus, for all T, there exists a sufficiently small Δ such that $(G, \mathbf{N}^{G,T})$ Δ-dissociates a.s. By Definition 4.15, this shows that G is finitely dissociable. □

As demonstrated in Section 4.3, even Markovian IPS with very regular jump rate functions, finite dissociability and well-posedness can fail on some graphs.

4.2.4 Consistent Spatial Localization on Finitely Dissociative Graph Sequences

The main result of this section is Proposition 4.22, which states that sequences of finitely dissociable graphs consistently spatially localize the SDE (3.1.4). A key challenge is to find an explicit and consistent representation of the localizing map $\mathcal{S}_T(\cdot; G, \mathbf{N}^G)$ on finitely dissociative graphs. To do this, we start by introducing the notion of a *causal chain*. Fix G and let

$$\mathcal{E}_{v,T} := \{t \in (0,T] : \mathbf{N}_v^{G,T}(\{t\}) = 1\}, \qquad v \in V_G, T \in \mathbb{R}_+, \tag{4.2.14}$$

where $\mathbf{N}_v^{G,T}$ is defined as in (4.2.9).

Definition 4.21 (Causal chains)**.** [36, Definition 5.18] Given $T \in (0, \infty)$, an interval $I := [t_1, t_2] \subseteq [0, T]$ and vertices $u, v \in V_G$, a $(G, \mathbf{N}^{G,T})$-causal chain from v to u during I is a path $\Gamma := (v = u_0, u_1, \dots, u_n = u)$ in G for some $n \in \mathbb{N}_0$ such that if $n \neq 0$, there exists an increasing sequence $t_1 := s_0 < s_1 < \cdots < s_n \leq t_2$ for which $s_i \in \mathcal{E}_{u_i,T}$, $i = 1, \dots, n$. We write $v \rightsquigarrow_{t_1,t_2} u$ if there exists a $(G, \mathbf{N}^{G,T})$-causal chain from v to u during $I = [t_1, t_2]$, and for any $U \subset V_G$, we write $v \rightsquigarrow_{t_1,t_2} U$ if $v \rightsquigarrow_{t_1,t_2} u$ for some $u \in U$.

Intuitively, causal chains describe long-range interactions over the graph that can develop over an interval I, even though the instantaneous evolution of the state of a vertex is only influenced by the states of neighboring vertices. Specifically, given a graph G, $\mathcal{O} \in \Lambda_G$ and $0 \leq t_1 < t_2 < \infty$, define

$$\mathcal{A}^G_{t_1,t_2}(\mathcal{O}) := \{v \in V_G : v \rightsquigarrow_{t_1,t_2} \mathcal{O}\} \quad \text{and} \quad \mathcal{A}^G_{t_1}(\mathcal{O}) := \mathcal{A}_{0,t_1}(\mathcal{O}). \tag{4.2.15}$$

Then $\mathcal{A}^G_{t_1,t_2}(\mathcal{O})$ represents the set of vertices in G that are "seen" by vertices in $\mathcal{O}$ through a causal chain during some time interval $[t_1, t_2]$. We now state the main result of this section.

Proposition 4.22. *[36, Proposition 5.19] Suppose Assumption 1 holds, and suppose the deterministic, (possibly disconnected) graph* $(G, \varnothing)$ *is finitely dissociable. Then* $\mathscr{C}_\varnothing(G)$ *spatially localizes the SDE* (3.1.4) *in the sense of Definition 4.4 for all* $v \in V_G$*. Moreover, any sequence of connected, deterministic finitely dissociable* $[\overline{\mathcal{K}}, \mathcal{K}]$*-graphs* $\{G_n\}_{n\in\mathbb{N}}$ *consistently spatially localizes the*

SDE (3.1.4).

We first show why the proposition directly implies Theorem 4.2, and subsequently present its proof, which proceeds by showing that the (random) map that takes $\mathcal{O} \in \Lambda_G$ to the closure of $\mathcal{A}_T^G(\mathcal{O})$ for finitely dissociable graphs G defines a consistent family of localizing maps.

Proof of Theorem 4.2 given Proposition 4.22: By Lemma 3.10, it suffices to prove the strong well-posedness of (3.1.4) under the additional condition that (G, ξ) is deterministic. Let $N \in \mathbb{N} \cup \{\infty\}$ be the number of connected components of G and let $\{\varnothing_i\}_{i<N+1}$ be a collection of vertices chosen from each of the connected components of G.

Note that (3.1.4) is strongly well-posed for the initial data (G, ξ) if and only if it is strongly well-posed for the initial data $\mathscr{C}_{\varnothing_i}(G, \xi)$ for each $i < N + 1$. Fix any such i. Then given Assumption 1, Proposition 4.22 and Definition 4.8 show that $\mathscr{C}_{\varnothing_i}(G)$ a.s. spatially localizes (3.1.4), and Proposition 4.1 shows that Assumption $1'$ holds. Therefore, the theorem follows from Proposition 4.10. $\square$

Proof of Proposition 4.22: We start with the proof of the first statement of the proposition. Fix an arbitrary finitely dissociable, deterministic $[\overline{\mathcal{K}}, \mathcal{K}]$-graph G and let $(\mathbb{F}, \mathbf{N}^G)$ be a filtration-Poisson process pair. If G is finitely dissociable, then so is $\mathscr{C}_\varnothing(G)$. As we only need to show that $\mathscr{C}_\varnothing(G)$ spatially localizes the SDE (3.1.4), we may assume without loss of generality that G is a connected graph, in which case we show that G spatially localizes the SDE (3.1.4). For each (not necessarily finite) $W \subseteq V_G$, let $X^{G[W],\xi_W}$ be an arbitrary $(\mathbb{F}, \mathbf{N}_W^G)$-weak solution to (3.1.4) for $(G[W], \xi_W)$ (assuming one exists, which is always the case when W is finite by Proposition 4.1). Let $\mathcal{O} \in \Lambda_G$, and for $T \in \mathbb{R}_+$, set $\mathcal{S}_T(\mathcal{O}; G, \mathbf{N}^G) := \mathcal{A}_T(\mathcal{O}) := \mathcal{A}_T^G(\mathcal{O})$, with the latter defined as in (4.2.15). Note that by the definition of causal chains, for any $U' \in \Lambda_G$, $\{\mathcal{A}_T(\mathcal{O}) \subset U'\}$ lies in $\mathcal{H}_T^{\mathbf{N}^G}$, and hence $\mathcal{S}_T(\cdot; G, \mathbf{N}^G) = \mathcal{A}_T$ satisfies (4.2.1). Moreover, since $u \rightsquigarrow_{0,T} u$ for all $u \in \mathcal{O}$, $\mathcal{A}_T(\mathcal{O}) \supseteq \mathcal{O}$, thus verifying that $\mathcal{S}_T(\cdot; G, \mathbf{N}^G)$ satisfies condition 1 of Definition 4.4. To prove that finitely dissociable graphs spatially localize (3.1.4), we argue below that it suffices to establish the following claims for any $T \in (0, \infty)$:

Claim 1: If T is a G-dissociation number, then $|\mathcal{A}_T(\mathcal{O})| < \infty$ a.s. for every $\mathcal{O} \in \Lambda_G$.

Claim 2: If G is finitely dissociable, then $|\mathcal{A}_T(\mathcal{O})| < \infty$ a.s. for every $\mathcal{O} \in \Lambda_G$.

Claim 3: Fix any $\ell \in \mathbb{N}$ and (not necessarily finite) $W \subseteq V_G$ such that $B_\ell(G) \subseteq W$ and a weak solution $X^{G[W],\xi_W}$ exists. Then for each $\mathcal{O} \in \Lambda_{G[W]}$, $X_{\mathcal{O}}^{B_\ell(G),\xi_{B_\ell(G)}}[T] = X_{\mathcal{O}}^{G[W],\xi_W}[T]$ a.s. on the event $\{\mathcal{A}_T(\mathcal{O}) \subseteq B_\ell(G)\}$.

Claim 2 shows that $\mathcal{S}_T(\cdot; G, \mathbf{N}^G)$ satisfies (4.2.2) when G is finitely dissociable. Suppose Claims 1-3 hold, and let $W \supseteq B_\ell(G)$ be any (not necessarily finite) vertex set for which there exists a weak solution $X^{G[W],\xi_W}$ to the SDE (3.1.4). Then clearly $B_\ell(G[W]) = B_\ell(G)$ for $\mathcal{O} \in \Lambda_{G[W]}$ and on the event $\{B_\ell(G) \supseteq \mathcal{A}_T(\mathcal{O})\}$, the $(G[W], \mathbf{N}_W^{G,T})$-causal chains ending at $\mathcal{O}$ at time T are the same as the $(G, \mathbf{N}^{G,T})$-causal chains ending at $\mathcal{O}$ at time T, that is, $\mathcal{A}_T^{G[W]}(\mathcal{O}) = \mathcal{A}_T^G(\mathcal{O}) = \mathcal{A}_T(\mathcal{O})$. Thus, we can apply Claim 3 twice (first directly and then when W is replaced with V) to conclude that if there exists a weak solution $X^{G,\xi}$ to the SDE (3.1.4), then

$$X_{\mathcal{O}}^{G[W],\xi_W}[T] = X_{\mathcal{O}}^{B_\ell(G),\xi_{B_\ell(G)}}[T] = X_{\mathcal{O}}^{G,\xi}[T] \text{ a.s. on } \{\mathcal{A}_T(\mathcal{O}) \subseteq B_\ell(G)\}.$$

This implies that $\mathcal{S}_T(\cdot; G, \mathbf{N}^G)$ satisfies (4.2.3) and thus, condition 2 of Definition 4.4 holds. Since we have verified all properties of Definition 4.4 when G is dissociable, the first assertion of the proposition follows.

We now turn to the proofs of the claims.

Proof of Claim 1. Fix $\mathcal{O} \in \Lambda_G$. Define $\widehat{G} := \mathrm{perc}_T(G, \mathbf{N}^{G,T})$ as in Definition 4.13, where the point process $\mathbf{N}^{G,T}$ is given by (4.2.9). If T is a G-dissociation number, then each of the connected components of $\widehat{G}$ is a.s. finite, and $\mathbf{N}_w^{G,T}(0,T] = 0$ for all $w \in \mathcal{N}_{\widehat{\mathcal{O}}}$, where $\widehat{\mathcal{O}} := \mathcal{O} \cup \bigcup_{v\in\mathcal{O}} \mathscr{C}_v(\widehat{G})$. Since $\mathcal{O}$ is finite, $\widehat{\mathcal{O}}$ is a.s. finite. Now suppose $u \in \mathcal{O}$ and $v \in \mathcal{A}_T(u)$. Let $(v = u_0, u_1, \ldots, u_n = u)$ be a $(G, \mathbf{N}^{G,T})$-causal chain with respect to $[0,T]$. Then, for any $i = 1, \ldots, n$, u_i must be active, that is, $\mathbf{N}_{u_i}^{G,T}(0,T] > 0$. Thus, for each $i = 1, \ldots, n$, $u_i \in \widehat{\mathcal{O}}$, and so $v \in \mathrm{cl}_{\widehat{\mathcal{O}}}(G)$. Hence, $\mathcal{A}_T(\mathcal{O}) = \bigcup_{u\in\mathcal{O}} \mathcal{A}_T(u) \subseteq \mathrm{cl}_{\widehat{\mathcal{O}}}(G)$, which is a.s. finite since $\widehat{\mathcal{O}}$ is a.s. finite and G is locally finite.

Proof of Claim 2. Fix $u \in U$ and $v \in \mathcal{A}_T(U)$. Because G is finitely dissociable, there must exist a G-dissociation number $\Delta > 0$. If $\Delta \geq T$, the result follows from Claim 1. If $\Delta \in (0,T)$, then Claim 1 implies that a.s.,

$$|\mathcal{A}_\Delta(U)| < \infty \quad \text{for every } U \in \Lambda_G. \tag{4.2.16}$$

To complete the proof, for every $t \in [0, T-\Delta]$, we will show that if $|\mathcal{A}_t(U)| < \infty$ a.s. for every $U \in \Lambda_G$, then we also have $|\mathcal{A}_{t+\Delta}(\mathcal{O})| < \infty$ a.s. for every $\mathcal{O} \in \Lambda_G$.

To this end, fix $t \in [0, T-\Delta]$, and suppose $|\mathcal{A}_t(\mathcal{O})| < \infty$ a.s. for all $\mathcal{O} \in \Lambda_G$. Fix $\mathcal{O} \in \Lambda_G$, $u \in \mathcal{O}$ and $v \in \mathcal{A}_{t+\Delta}(u)$. Then there exists a $(G, \mathbf{N}^{G,T})$-causal chain $\Gamma = (v = u_0, u_1, \dots, u_n = u)$ during the interval $[0, t+\Delta]$, with the corresponding sequence of times $0 = s_0 < s_1 < \cdots < s_n \leq t + \Delta$. Let i_* be the largest integer $i \in \{0, \dots, n\}$ such that $s_i \leq \Delta$. Then $v \in \mathcal{A}_\Delta(u_{i_*})$. Furthermore, by considering the path $(u_0' = u_{i_*}, \dots, u_i' = u_{i+i_*}, \dots, u_{n-i^*}' = u)$ with times $s_0' = \Delta$ and $s_i' = s_{i+i^*}$, $i = 1, \dots, n - i_*$, it follows that $u_{i_*} \rightsquigarrow_{\Delta, t+\Delta} u$. Thus, $v \in \mathcal{A}_\Delta(\mathcal{A}_{\Delta,t+\Delta}(u))$, which implies $\mathcal{A}_{t+\Delta}(\mathcal{O}) \subseteq \mathcal{A}_\Delta(U)$, where $U := \mathcal{A}_{\Delta, t+\Delta}(\mathcal{O})$. In view of (4.2.16), it suffices to show that U is a.s. finite, but this holds because by time homogeneity of Poisson processes, $U \overset{(\mathrm{d})}{=} \mathcal{A}_t(\mathcal{O})$, which is a.s. finite by assumption.

Proof of Claim 3. For notational conciseness, let $H := G[W]$, where we recall that W is a deterministic vertex set such that $B_\ell(G) \subseteq W$. Additionally, let $H_\ell := B_\ell(G)$, and note that $H_\ell = B_\ell(H)$. Fix $\mathcal{O} \in \Lambda_H$ with $|\mathcal{A}_T^H(\mathcal{O})| < \infty$ and note that $\mathcal{A}_T^H(\mathcal{O}) = \mathcal{A}_T^G(\mathcal{O}) = \mathcal{A}_T(\mathcal{O})$ on the event $\{H_\ell \supseteq \mathcal{A}_T^H(\mathcal{O})\}$. To prove Claim 3, we need to show that

$$X_{\mathcal{O}}^{H,\xi_W}[T] = X_{\mathcal{O}}^{H_\ell,\xi_{H_\ell}}[T] \text{ a.s. on } \{\mathcal{A}_T(\mathcal{O}) \subseteq V_{H_\ell}\}. \tag{4.2.17}$$

Assume $\mathbf{N}_W^{G,T}(\{T\}) = 0$, which holds a.s. because $\mathbf{N}_W^{G,T}$ is a countable collection of homogeneous Poisson processes. We prove this claim in a recursive fashion by iterating over events in the driving noise $\mathbf{N}^G$, and relating them to the dynamics of the SDE (3.1.4). For each $v \in W$, let $\{t_i^v, i \in \mathbb{N}\}$ be an enumeration of the (a.s. finite) set $\mathcal{E}_{v,T}$ from (4.2.14) of events of $\mathbf{N}_v^{G,T}$ in $[0, T]$, arranged in increasing order. Below, we use the conventions that $\max \emptyset = 0$, $\inf \emptyset = \infty$ and $[0) = \{0\}$. Define $U_0 := \mathcal{O}$, $\tau_0 := T$, choose an arbitrary vertex $v_0 \in \mathcal{O}$ and for $k \in \mathbb{N}$, recursively define

$$\tau_k := \begin{cases} \max\{t_i^v : v \in U_{k-1}, t_i^v < \tau_{k-1}\} & \text{if } \tau_{k-1} > 0, \\ 0 & \text{if } \tau_{k-1} = 0, \end{cases} \tag{4.2.18}$$

$$v_k := \begin{cases} v \in U_{k-1} \text{ s.t. } \mathbf{N}_v^{G,T}(\{\tau_k\}) = 1 & \text{if } \tau_k > 0, \\ v_{k-1} & \text{if } \tau_k = 0, \end{cases} \tag{4.2.19}$$

$$U_k := \begin{cases} U_{k-1} \cup \mathrm{cl}_{v_k} & \text{if } \tau_k > 0, \\ U_{k-1} & \text{if } \tau_k = 0. \end{cases} \tag{4.2.20}$$

Also, set $K := \inf\{k \in \mathbb{N} : \tau_k = 0\}$. Note that the above construction is well defined even if $K = \infty$, the sequence $\{\tau_k\}_{k\in\mathbb{N}}$ is strictly decreasing, and the sequence $\{U_k\}_{k\in\mathbb{N}}$ is non-decreasing, but with possible repetitions (for example, if v_k lies in the interior of U_{k-1}).

The set U_K is specifically constructed so that $U_K \subseteq \mathcal{A}_T(\mathcal{O})$ and so that $X^{H,\xi_W}_{\mathcal{O}}[T] = X^{H_\ell,\xi_\ell}_{\mathcal{O}}[T]$ on the event $\{H_\ell \supseteq U_K\}$, which immediately implies Claim 3. In fact, we will show that the recursive construction (4.2.18)-(4.2.20) is such that the following two claims are true.

Claim 3A: For every $k \in \mathbb{N}_0$, $U_k \subseteq \mathcal{A}_T(\mathcal{O})$.

Claim 3B: For every $k \in \mathbb{N}_0$,

$$X^{H,\xi_W}_{U_k}[\tau_k) = X^{H_\ell,\xi_{H_\ell}}_{U_k}[\tau_k) \Rightarrow X^{H,\xi_W}_{\mathcal{O}}[T] = X^{H_\ell,\xi_{H_\ell}}_{\mathcal{O}}[T] \text{ a.s. on } \{U_k \subseteq V_{H_\ell}\}. \tag{4.2.21}$$

We first show how Claim 3 follows from the auxiliary claims. On the event $\{\mathcal{A}_T(\mathcal{O}) \subseteq V_{H_\ell}\}$, Claim 3A and (4.2.18) together show that $\{\tau_k\}_{k\in\mathbb{N}} \setminus \{0\}$ is contained in the events of $\mathbf{N}^{G,T}_{H_\ell}(0,T]$; since H_ℓ is finite, this implies $K \leq 1 + \sum_{v\in V_{H_\ell}} \mathbf{N}^{G,T}_v(0,T] < \infty$ a.s.. By the definition of K, (4.2.18) and (4.2.20) this implies that a.s. on the event $\{\mathcal{A}_T(\mathcal{O}) \subseteq V_{H_\ell}\}$, $\tau_K = 0$, $U_K = \bigcup_{k\in\mathbb{N}_0} U_k$ and (applying Claim 3A with $k = K$) $U_K \subseteq \mathcal{A}_T(\mathcal{O}) \subseteq V_{H_\ell} \subseteq W$. Together, these properties imply, $X^{H,\xi_W}_{U_K}[\tau_K) = \xi_{U_K} = X^{H_\ell,\xi_{H_\ell}}_{U_K}[\tau_K)$. Invoking Claim 3B with $k = K$, we see that a.s. on $\{U_K \subseteq V_{H_\ell}\} \supseteq \{\mathcal{A}_T(\mathcal{O}) \subseteq V_{H_\ell}\}$, we have $X^{H,\xi_W}_{\mathcal{O}}[T] = X^{H_\ell,\xi_{H_\ell}}_{\mathcal{O}}[T]$. This proves Claim 3.

We first provide a rough idea of the proof of the auxiliary claims. When $k = 0$, $U_k = \mathcal{O}$ and it is easy to see that claims 3A and 3B will hold trivially (due to the stipulation that $\mathcal{O} \subseteq \mathcal{A}_T(\mathcal{O})$ and the assumption that there is no jump at T). If there is any jump in the driving processes $\mathbf{N}^{G,T}_{\mathcal{O}}$ in $[0,T)$, then by the construction (4.2.18)-(4.2.20), the most recent event before $\tau_0 = T$ that could have influenced the value of the process X^{H,ξ_W} at τ_0 occurs at τ_1 corresponding to a transition at the vertex v_1. The local nature of the dynamics implies that this transition is influenced by the particles in cl_{v_1}, so that the trajectory of X^{H,ξ_W} up to time τ_1 is influenced by the trajectories of the particles in $\mathcal{O} \cup \mathrm{cl}_{v_1} = U_0 \cup \mathrm{cl}_{v_1} = U_1$ before time τ_1. Any vertex u in $U_1 \setminus \mathcal{O}$ belongs to $\mathcal{N}_{v_1}$

and so (u, v_1) forms a $(G, \mathbf{N}^{G,T})$ causal chain, showing that $U_1 \subset \mathcal{A}_T(0)$. Claims 3A and 3B follow by proceeding inductively in this manner, tracing backwards in time the $(G, \mathbf{N}^{G,T})$-causal chains that end in $\mathcal{O}$. We now provide fully rigorous proofs of the auxiliary claims.

Proof of Claim 3A. We prove the following assertion using induction: For any $k \in \mathbb{N}_0$ and $v \in U_k$, $v \rightsquigarrow_{\tau_{k+1},T} \mathcal{O}$. This is obviously true in the case $k = 0$, in which case $U_0 = \mathcal{O}$ and for any $v \in \mathcal{O}$, $v \rightsquigarrow_{\tau_1,T} v \in \mathcal{O}$ by definition. Now suppose that for all $v \in U_{k-1}$, $v \rightsquigarrow_{\tau_k,T} \mathcal{O}$. Then because $\tau_{k+1} \leq \tau_k$, it follows that $v \rightsquigarrow_{\tau_{k+1},T} \mathcal{O}$. There are now two cases to consider. In the first case, $U_k = U_{k-1}$ in which case the conclusion $v \rightsquigarrow_{\tau_{k+1},T} \mathcal{O}$ for all $v \in U_k$ holds trivially. In the second case, by (4.2.19) and (4.2.20), $U_k \setminus U_{k-1} \subset \mathcal{N}_{v_k}$, where $v_k \in U_{k-1}$, and furthermore, $\tau_k > 0$ and $\mathbf{N}^{G,T}_{v_k}(\{\tau_k\}) = 1$. Since $v_k \in U_{k-1}$, the induction assumption implies there exists a $(G, \mathbf{N}^{G,T})$-causal chain $\Gamma = (v_k := u_0, u_1, \ldots, u_n \in \mathcal{O})$ from v_k to $\mathcal{O}$ contained in the time interval $[\tau_k, T]$ with corresponding times $\tau_k := \theta_0 < \theta_1 < \cdots < \theta_n \leq T$. Then because $\mathbf{N}^{G,T}_{v_k}(\{\tau_k\}) = 1$, it follows that $(\tau_{k+1} = \theta < \theta_0 < \cdots < \theta_n)$ is a sequence of times such that for all $i = 0, \ldots, n$, $\mathbf{N}^{G,T}_{u_i}(\{\theta_i\}) = 1$, and for any $w \in U_k \setminus U_{k-1}$, $(w, v_k, u_1, \ldots, u_n)$ is a $(G, \mathbf{N}^{G,T})$-causal chain ending in $\mathcal{O}$. Thus, $w \rightsquigarrow_{\tau_{k+1},T} \mathcal{O}$ for all $w \in U_k$. By induction, we may conclude that for any $k \in \mathbb{N}_0$ and $v \in U_k$, $v \rightsquigarrow_{\tau_{k+1},T} \mathcal{O}$, which proves the assertion. Now, by Definition 4.21, the assertion in turn implies $v \rightsquigarrow_{0,T} \mathcal{O}$. By (4.2.15), this implies that $U_k \subseteq \mathcal{A}_T(\mathcal{O})$ for all $k \in \mathbb{N}_0$, concluding the proof of claim 3A. $\square$

Proof of Claim 3B. We will again use an argument by induction. First, note that the base case $k = 0$ in (4.2.21) is true because $U_0 = \mathcal{O}$, $\tau_0 = T$ and both $X^{H,\xi_W}_{\mathcal{O}}$ and $X^{H_\ell,\xi_{H_\ell}}_{\mathcal{O}}$ a.s. do not have a jump at T because $\mathbf{N}^{G,T}_{\mathcal{O}}(\{T\}) = 0$ a.s.. Now, suppose (4.2.21) holds for $k = m - 1$, for some $m \in \mathbb{N}$. To show it holds for $k = m$, we argue that it suffices to show that

$$X^{H,\xi_W}_{U_m}[\tau_m] = X^{H_\ell,\xi_{H_\ell}}_{U_m}[\tau_m] \quad \Rightarrow \quad X^{H,\xi_W}_{U_{m-1}}[\tau_{m-1}] = X^{H_\ell,\xi_{H_\ell}}_{U_{m-1}}[\tau_{m-1}] \text{ a.s. on } \{U_m \subseteq V_{H_\ell}\}. \tag{4.2.22}$$

Indeed, $U_{m-1} \subseteq U_m$ implies $U_{m-1} \subseteq V_{H_\ell}$ on the event $\{U_m \subseteq V_{H_\ell}\}$, and so (4.2.22) and (4.2.21), with $k = m - 1$, shows that (4.2.21) holds for $k = m$. Claim 3B then follows by induction.

To establish (4.2.22), assume $U_m \subseteq V_{H_\ell}$ and $X^{H,\xi_W}_{U_m}[\tau_m] = X^{H_\ell,\xi_{H_\ell}}_{U_m}[\tau_m]$. Since $U_{m-1} \subset U_m$, this implies $X^{H,\xi_W}_{U_{m-1}}[\tau_m] = X^{H_\ell,\xi_{H_\ell}}_{U_{m-1}}[\tau_m]$. Moreover, note from (4.2.18) that τ_m is the largest time prior to τ_{m-1} that there is an event for any of the Poisson processes in U_{m-1} and hence,

$\sum_{v\in U_{m-1}} \mathbf{N}^{G,T}_v(\tau_m, \tau_{m-1}) = 0$. The form of the SDE (3.1.4) then implies that both $X^{H,\xi_W}_{U_{m-1}}$ and $X^{H_\ell,\xi_{H_\ell}}_{U_{m-1}}$ are constant on (τ_m, τ_{m-1}). Thus, to establish (4.2.22), it suffices to show that $X^{H,\xi_W}_{U_{m-1}}(\tau_m) = X^{H_\ell,\xi_{H_\ell}}_{U_{m-1}}(\tau_m)$ a.s.. Now, by (4.2.18)-(4.2.19), v_m is the only vertex in U_{m-1} such that $\mathbf{N}^{G,T}_{v_m}(\{\tau_m\}) = 1$. Thus,

$$X^{H,\xi_W}_{U_{m-1}\setminus\{v_m\}}(\tau_m) = X^{H,\xi_W}_{U_{m-1}\setminus\{v_m\}}(\tau_m-) = X^{H_\ell,\xi_{H_\ell}}_{U_{m-1}\setminus\{v_m\}}(\tau_m-) = X^{H_\ell,\xi_{H_\ell}}_{U_{m-1}\setminus\{v_m\}}(\tau_m), \tag{4.2.23}$$

and so it only remains to show that $X^{H,\xi_W}_{v_m}(\tau_m) = X^{H_\ell,\xi_{H_\ell}}_{v_m}(\tau_m)$ a.s.. Since for $m \in \mathbb{N}$, $\mathrm{cl}_{v_m}(G) \subset U_m$ by (4.2.20), and $U_m \subseteq V_{H_\ell} \subseteq W$ by assumption, $\mathrm{cl}_{v_m}(H_\ell) = \mathrm{cl}_{v_m}(G) =: \mathrm{cl}_{v_m}$. Then the assumption $X^{H,\xi_W}_{U_m}[\tau_m) = X^{H_\ell,\xi_{H_\ell}}_{U_m}[\tau_m)$ implies $X^{H,\xi_W}_{\mathrm{cl}_{v_m}}[\tau_m) = X^{H_\ell,\xi_{H_\ell}}_{\mathrm{cl}_{v_m}}[\tau_m)$ a.s. The locality and predictability of the jump rates stated in the standing assumption then imply that for $j \in \mathcal{J}$, denoting by $(\overline{\kappa}, \kappa)$ the marks of $G[\mathrm{cl}_{v_m}]$ we have

$$r^{H,v_m}_j(s, X^{H,\xi_W}) = \bar{r}^{H[\mathrm{cl}_{v_m}]}_j(s, X^{H,\xi_W}_{\mathrm{cl}_{v_m}}, \overline{\kappa}, \kappa) = r^{H_\ell,v_m}_j(s, X^{H_\ell,\xi_{H_\ell}}), \quad s \in [0, \tau_m].$$

Due to the form of the SDE (3.1.4), this shows $X^{H,\xi_W}_{v_m}(\tau_m) = X^{H_\ell,\xi_{H_\ell}}_{v_m}(\tau_m)$, as desired. □

We now turn to the proof of the second assertion of the proposition. Let $\{G_n\}$ be a sequence of deterministic finitely dissociable $[\overline{\mathcal{K}}, \mathcal{K}]$-graphs, and for each $n \in \mathbb{N}$, let $\mathbf{N}^{G_n}$ be a driving noise compatible with G_n. Then define $\mathcal{A}^n_{t_1,t_2}$ analogously to $\mathcal{A}_{t_1,t_2}$, but with G and $\mathbf{N}^G$ replaced by G_n and $\mathbf{N}^{G_n}$, respectively. By the first assertion of the proposition just established above, for each n, $\mathcal{A}^n_T$ is a localizing map of the SDE (3.1.4) on $(G_n, \mathbf{N}^{G_n})$. It follows that $\mathcal{S}_T(\cdot; G_n, \mathbf{N}^{G_n}) := \mathrm{cl}_{\mathcal{A}^n_T(\cdot)}(G_n)$ is likewise a localizing map of the SDE (3.1.4) on $(G_n, \mathbf{N}^{G_n})$. Note that $\mathcal{S}_T(\cdot; G_n, \mathbf{N}^{G_n})$ satisfies (4.2.1) as a consequence of the fact that $\{\mathcal{A}^n_T(U) \subset U'\}$ lies in $\mathcal{H}^{\mathbf{N}^{G_n}}_T$ for every $U, U' \in \Lambda_{G_n}$. With this choice of localizing maps $\mathcal{S}_T(\cdot; G_n, \mathbf{N}^{G_n})$, clearly condition 1 of Definition 4.7 holds.

It only remains to show that condition 2 of Definition 4.7 is also satisfied. Fix $n, n', \ell \in \mathbb{N}$ such that $B_\ell([G_{n',*}]) = B_\ell([G_{n,*}])$. Define the $\mathcal{F}_0$-measurable event

$$\mathcal{I} := \{B_\ell([G_{n',*}], \mathbf{N}^{G_{n'}}) = B_\ell([G_{n,*}], \mathbf{N}^{G_n})\}.$$

and for notational conciseness, set $W := V_{B_{\ell-1}(G_n)} = V_{B_{\ell-1}(G_{n'})}$. Let $U \subset W$ be any set. Then it suffices to prove that on the event $\mathcal{I} \cap \{W \supseteq \mathcal{A}^n_T(U)\} = \mathcal{I} \cap \{B_\ell(G_n) \supseteq \mathcal{S}_T(U; G_n, \mathbf{N}^{G_n})\}$, $\mathcal{A}^n_T(U) = \mathcal{A}^{n'}_T(U)$ a.s..

Let $\Gamma = (v = u_0, u_1, \ldots, u_m \in U)$ be a $(G_n, \mathbf{N}^{G_n,T})$-causal chain up to time T which we will refer to as a G_n-causal chain for now. Let $(0 = t_0 < t_1 < \cdots < t_n \le T)$ be the corresponding times associated with Γ. By definition of $\mathcal{A}_T^n(U)$, $\Gamma \subseteq W$ on the event $\mathcal{I} \cap \{W \supseteq \mathcal{A}_T^n(U)\}$. Moreover, $([G_{n,*}[W]], \mathbf{N}_W^{G_n}) = ([G_{n',*}[W]], \mathbf{N}_W^{G_{n'}}))$, so Γ is a path in $G_{n'}$. Furthermore, for every $k \in \{1, \ldots, m\}$, $\mathbf{N}_{u_k}^{G_n}(\{t_k\}) = 1$, so $\mathbf{N}_{u_k}^{G_{n'}}(\{t_k\}) = \mathbf{N}_{u_k}^{G_n}(\{t_k\}) = 1$. It follows that Γ is also a $(G_{n'}, \mathbf{N}^{G_{n'}})$-causal chain which we will refer to as a $G_{n'}$-causal chain.

We now claim that any $G_{n'}$-causal chain Γ' is also a G_n-causal chain. To see why this claim holds, let $\Gamma' = (v' = u_0', u_1', \ldots, u_m' \in U)$. By an identical argument to what we applied in the previous paragraph, if $\Gamma' \subseteq W$, then Γ' is a G_n-causal chain. We now argue by contradiction to justify that this is the only possible case. Indeed, suppose there exists $k \in \{0, 1, \ldots, m\}$ such that $u_k' \notin W$. Note that $u_m' \in U \subseteq W$, so $k < m$. Choose the maximal such k, which implies that $(u_{k+1}', \ldots, u_m') \subseteq W$. Then because subpaths of causal chains are causal chains, $\overline{\Gamma}' := (u_k', \ldots, u_m')$ is also a $G_{n'}$-causal chain. Furthermore, $u_k' \in \mathcal{N}_W(G_{n'})$, so on the event $\mathcal{I}$, $\overline{\Gamma}'$ is a path and a causal chain in G_n. Moreover, $\overline{\Gamma}' \nsubseteq W$ which contradicts our assumption that $\mathcal{A}_T^n(U) \subseteq W$. This proves the claim.

Because the causal chains in G_n and $G_{n'}$ that end in U coincide, it follows that $\mathcal{A}_T^n(U) = \mathcal{A}_T^{n'}(U)$ on the event $\mathcal{I} \cap \{\mathcal{A}_T^n(U) \subseteq W\}$. Equivalently, $\mathcal{S}_T(U; G_n, \mathbf{N}^{G_n}) = \mathcal{S}_T(U; G_{n'}, \mathbf{N}^{G_{n'}}))$ on the event $\{\mathcal{S}_T(U; G_n, \mathbf{N}^{G_n}) \subseteq B_\ell(G_n)\} \cap \mathcal{I}$. It follows that condition 2 of Definition 4.7 also holds for the sequence $\{G_n\}_{n\in\mathbb{N}}$.

This proves the second assertion and hence, concludes the proof of the proposition. $\square$

Corollary 4.23. *Suppose Assumptions 1 and 2 hold. Suppose the deterministic sequence of isomorphism classes* $\{\langle G_n\rangle\}_{n\in\mathbb{N}}$ *in* $\mathscr{G}_*[\overline{\mathcal{K}}, \mathcal{K}]$ *are finitely dissociable. Then* $\{\langle G_n\rangle\}_{n\in\mathbb{N}}$ *consistently spatially localizes the SDE* (3.1.4) *in the sense of Definition 4.8.*

Proof. Let $\{G_n\}_{n\in\mathbb{N}}$ be any sequence of connected $[\overline{\mathcal{K}}, \mathcal{K}]$-graphs such that $G_n \in \langle G_n\rangle$ for all $n \in \mathbb{N}$. Then by Remark 4.17, G_n is finitely dissociable for each $n \in \mathbb{N}$. Then by Proposition 4.22, $\{G_n\}_{n\in\mathbb{N}}$ consistently spatially localizes the SDE (3.1.4). Since this sequence was chosen arbitrarily, $\{\langle G_n\rangle\}_{n\in\mathbb{N}}$ consistently spatially localizes the SDE (3.1.4). $\square$

4.3 Counterexample: When the SDE is not Well-Posed

There are two ways the SDE (3.1.4) can fail to be well-posed. Either there exist no solutions, or there exist multiple solutions. To construct an IPS with no solutions, it is a simple matter of constructing unbounded jump rate functions that increase fast enough that all stochastic processes X for which (3.1.4) holds a.s. with $X^{G,\xi} := X$ explode (have infinitely many discontinuities on some compact interval in $\mathbb{R}_+$) with positive probability. In this section, we show that well-posedness is non-trivial even if we restrict ourselves to Markov processes with bounded jump rate functions by constructing a graph and a collection of jump rate functions with respect to which the SDE (3.1.4) has multiple strong solutions with different laws.

Proposition 4.24. *[36, Proposition A.1] Let $\mathcal{T}$ be the tree such that for every $k \in \mathbb{N}$, all vertices v in the kth generation of $\mathcal{T}$ have 4^k children. Using the notation of Example 3.15, suppose $\mathcal{X} = \{0,1\}$, $\mathcal{J} = \{1\}$ and for $z \in \mathcal{X}^V$ and $v \in V$, let $\tilde{r}_1^{\mathcal{T}[cl_v],v}(z) = 1$ if both $z_v = 0$ and $\sum_{u \in \mathcal{N}_v} z_u > 0$, and let $\tilde{r}_1^{\mathcal{T}[cl_v],v}(z) = 0$ otherwise. Then the corresponding SDE* (3.1.4) *has multiple strong solutions for the initial state $\xi_v = X_v(0) = 0, v \in V$ and these solutions also have different laws.*

The proof of this proposition requires us to introduce the notion of infinite causal chains.

Definition 4.25 (Infinite causal chains)**.** [36, Definition A.2] For any locally finite graph G, $T \in (0,\infty)$ and associated Poisson processes $\mathbf{N}^{G,T}$ defined as in (4.2.9), an infinite $(G, \mathbf{N}^{G,T})$-causal chain ending at a vertex u is a path $\Gamma = (u = u_0, u_1, u_2, \dots)$ such that for every $n \in \mathbb{N}$, $(u_n, u_{n-1}, \dots, u_1, u_0)$ is a $(G, \mathbf{N}^{G,T})$-causal chain (see Definition 4.21).

To prove Proposition 4.24, we first show that the tree $\mathcal{T}$ has an infinite causal chain almost surely.

Lemma 4.26. *[36, Lemma A.3] Let $\mathcal{T} = (V, E, \varnothing)$ be the tree defined in Proposition 4.24 and let $\mathbf{N}^{\mathcal{T}}$ be the corresponding driving noise. Fix any $T \in (0,\infty)$ and a function $C : \mathbb{N} \times \mathbb{R}_+ \to \mathbb{R}_+$ that is non-decreasing in its arguments, and let $\mathbf{N}^{\mathcal{T},T}$ be the associated Poisson processes defined in* (4.2.9)*. Then there exists an infinite $(\mathcal{T}, \mathbf{N}^{\mathcal{T},T})$-causal chain with probability 1.*

Proof. Define the path $\Gamma := \{\varnothing := v_0, v_1, \dots\}$ using the following recursive construction. For $n \geq 1$, first check if there exists $u \in c_{v_n}(\mathcal{T})$ such that $\mathbf{N}_u^{\mathcal{T},T}(2^{-n-2}, 2^{-n-1}) > 0$. If this condition is satisfied, then set $v_{n+1} = u$ (if multiple children satisfy this then we choose one arbitrarily from amongst them). If not, then we choose any child of v_n arbitrarily and set it to be v_{n+1}.

Define $U := \{v_i, i \in \mathcal{I}\} \subseteq \Gamma$, where $\mathcal{I} := \{n \in \mathbb{N}_0 : \mathbf{N}_{v_n}^{\mathcal{T},T}(2^{-n-1}, 2^{-n}) > 0\}$. We now claim that U contains an infinite path a.s.. To see why, we prove the equivalent claim that U a.s. contains all but finitely many vertices in Γ. Let $p_n := \mathbb{P}(v_{n+1} \notin U)$. Note that $\overline{C}(k,T) \geq C := \overline{C}(1,T) > 0$ for all k. Thus, for each $u \in U$,

$$\mathbb{P}\left(\mathbf{N}_u^{\mathcal{T},T}(2^{-n-1}, 2^{-n}) = 0\right) = e^{-2^{-n-1}\overline{C}(|\mathrm{cl}_u|,T)} \leq e^{-2^{-n-1}C}.$$

The independence of the Poisson processes $(\mathbf{N}_u^{\mathcal{T},T})_{u \in V}$ then implies that

$$p_n = \prod_{u \in c_{v_n}(\mathcal{T})} \mathbb{P}\left(\mathbf{N}_u^{\mathcal{T},T}(2^{-n-1}, 2^{-n}) = 0\right) \leq \left(e^{-2^{-n-1}C}\right)^{4^n} = e^{-2^{n-1}C}.$$

Because p_n decreases super-exponentially fast, $\sum_{n=1}^{\infty} p_n < \infty$. By the Borel-Cantelli lemma, it follows that with probability 1, all but finitely vertices in $\{v_n\}_{n \in \mathbb{N}_0}$ are contained in U, thus proving the claim. Note that for any path $\Gamma' = (u_n, u_{n-1}, \ldots, u_0) \subseteq U$ where u_0 is closest to the root, there exists a decreasing sequence of times $\{t_k \in (2^{-(d_{\mathcal{T}}(u_0,\varnothing)+k)}, 2^{-(d_{\mathcal{T}}(u_0,\varnothing)+k)+1})\}$ such that $\mathbf{N}_{u_k}^{\mathcal{T},T}(\{t_k\}) = 1$. Thus, Γ' is a $(\mathcal{T}, \mathbf{N}^{\mathcal{T},T})$-causal chain, and we have shown that U contains an infinite $(\mathcal{T}, \mathbf{N}^{\mathcal{T},T})$-causal chain a.s.. $\square$

We can now prove Proposition 4.24.

Proof of Proposition 4.24: Although there are infinitely many strong solutions for the initial state $\xi \equiv 0$, we explicitly construct two of them. First, the process $X_v^{\mathcal{T},\xi}(t) = 0$ for all $v \in V$ and $t \in \mathbb{R}_+$ is clearly a strong solution to the SDE (3.1.4) for the initial data $(\mathcal{T}, \xi)$ with $r_1^{\mathcal{T},v}(t,x) := \tilde{r}_1^{\mathcal{T}[cl_v]}(x(t-))$, with $\tilde{r}$ as given in the proposition.

Fix $T < \infty$. Note that for each $v \in V$, by (4.2.9), $\mathbf{N}_v^{\mathcal{T},T}$ is a unit rate Poisson process. Then for $(v,t) \in V \times (0,T)$, we say (v,t) has an infinite causal chain if there exists an infinite $(\mathcal{T}, \mathbf{N}^{\mathcal{T},T})$-causal chain ending at v on the interval $[0,t]$. Then for $v \in V$ and $t \in [0,T]$, define

$$\widetilde{X}_v^{\mathcal{T},\xi}(t) := \begin{cases} 1 & \text{if } (v,t) \text{ has an infinite causal chain,} \\ 0 & \text{otherwise.} \end{cases} \tag{4.3.1}$$

We claim this is also a solution to the same SDE. Fix any $v \in V$ and define

$$\widehat{X}_v(t) := \xi_v + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} \mathbb{I}_{\left\{r \le r_1^{\mathcal{T},v}(s,\widehat{X}_v,\widetilde{X}^{\mathcal{T},\xi}_{\mathcal{N}_v(\mathcal{T})})\right\}} \mathbf{N}_v^{\mathcal{T}}(ds,dr,dj), \quad t \in \mathbb{R}_+. \tag{4.3.2}$$

If $(v,t) \in V \times \mathbb{R}_+$ has no infinite causal chain, then by Definition 4.25, for any $s \le t$, either $\mathbf{N}_v^{\mathcal{T},T}([s,t]) = \mathbf{N}_v^{\mathcal{T}}([s,t]\times(0,1]\times\{1\}) = 0$, or there does not exist a $u \in \mathcal{N}_v(\mathcal{T})$ for which (u,s) has an infinite causal chain. It immediately follows that for every event τ in $\mathbf{N}_v^{\mathcal{T},T}$ in the interval $(0,t]$, $r_1^{\mathcal{T},v}(\tau,\widehat{X}_v,\widetilde{X}^{\mathcal{T},\xi}_{\mathcal{N}_v(\mathcal{T})}) = 0$ so that $\widehat{X}_v(t) = \widetilde{X}_v^{\mathcal{T},\xi}(t) = 0$. If (v,t) has an infinite causal chain, then there exists a minimal event $\tau \le t$ in $\mathbf{N}_v^{\mathcal{T},T}$ such that there exists a neighbor u of v for which (u,τ) has an infinite causal chain. By minimality of τ, (v,s) does not have an infinite causal chain on the event $\{s \le \tau\}$, and hence, $\widehat{X}_v(\tau-) = 0$ by the previous case. Thus, $r_1^{\mathcal{T},v}(\tau,\widehat{X}_v,\widetilde{X}^{\mathcal{T},\xi}_{\mathcal{N}_v(\mathcal{T})}) = 1$, so $\widehat{X}_v(\tau) = 1$. It follows that $\widehat{X}_v(t) = \widetilde{X}_v(t)$.

Thus, $\widehat{X}_v = \widetilde{X}_v^{\mathcal{T},\xi}$. Because v and t were chosen arbitrarily, it follows that $\widetilde{X}^{\mathcal{T},\xi}$ satisfies (3.1.4). Finally, because $(\mathcal{T},\mathbf{N}^{\mathcal{T},T})$-causal chains are $\mathbb{H}^{\mathcal{T},\xi,\mathbf{N}^{\mathcal{T}}}$-adapted, $\widetilde{X}^{\mathcal{T},\xi}$ must also be $\mathbb{H}^{\mathcal{T},\xi,\mathbf{N}^{\mathcal{T}}}$-adapted by (4.3.1). Thus, $\widetilde{X}^{\mathcal{T},\xi}$ is also a strong solution to (3.1.4), and by Lemma 4.26, $X^{\mathcal{T},\xi} \neq \widetilde{X}^{\mathcal{T},\xi}$ a.s.. $\square$

CHAPTER 5

Almost Sure Convergence, Weak Convergence and Correlation Decay

5.1 Almost Sure Convergence of the Dynamics

We now address convergence in the space of mark graphs (in the sense of Definition 2.19) and, consequently, local weak convergence of processes. Given well-posedness, this is equivalent to establishing continuity of the graph G marked with the trajectory of the unique strong solution $X^{G,\xi}$ to the SDE (3.1.4) with respect to the $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K}\times\mathcal{X}]$-valued initial data (G, ξ). This requires the following additional continuity assumption on the local jump rates with respect to the "environment" marks, which holds trivially when the mark spaces $\overline{\mathcal{K}}$ and $\mathcal{K}$ are discrete.

Assumption 3. Let $(G, X^{G,\xi}) = (V, E, \varnothing, \overline{\kappa}, \kappa, X^{G,\xi})$ be a strong solution to (3.1.4) for the $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-initial data (G, ξ). Then a.s. for every $(j, v) \in \mathcal{J} \times V_G$ and Lebesgue-a.e. $t \in \mathbb{R}_+$, $(\overline{\kappa}, \kappa)$ is a continuity point of the map:

$$\overline{\mathcal{K}}^{E_G} \times \mathcal{K}^{V_G} \ni (\overline{\vartheta}, \vartheta) \mapsto r_j^{(V_G, E_G, \varnothing, \overline{\vartheta}, \vartheta), v}(t, X^{G,\xi}).$$

Remark 5.1. Under Assumption 2, Assumption 3 is a class property in the sense that it holds for any measurable random graph $(G', \xi') := (V', E', \varnothing', \overline{\kappa}', \kappa', \xi') \cong (G, \xi)$ if and only if it holds for (G, ξ). To see why, first note that we may assume without loss of generality that (G, ξ) and

(G', ξ') are deterministic (Assumption 3 holds for (G, ξ) if and only if it holds for a.s. all realizations of (G, ξ)). Fix any $\varphi \in I((G', \xi'), (G, \xi))$. Recall from the explanation above Definition 3.14 that for any strong solution $(G', X^{G',\xi'})$ to (3.1.4) for the initial data (G', ξ'), there exists a strong solution $(G, X^{G,\xi})$ to (3.1.4) (with the appropriately shifted driving noise) for the initial data (G', ξ') such that $\varphi \in I((G, X^{G,\xi}), (G', X^{G',\xi'}))$. Then by Remark 3.2, for each $(j, v) \in \mathcal{J} \times V$ and Lebesgue-a.e. $t \in \mathbb{R}_+$, $(\overline{\vartheta}, \vartheta) \mapsto r_j^{(V,E,\varnothing,\overline{\vartheta},\vartheta),v}(t, X^{G,\xi})$ is a.s. continuous at $(\overline{\kappa}, \kappa)$ if and only if $(\varphi_*\overline{\vartheta}, \varphi_*\vartheta) := (\overline{\vartheta}', \vartheta') \mapsto r_j^{(V',E',\varnothing',\overline{\vartheta}',\vartheta'),\varphi(v)}(t, X^{G',\xi'}) = r_j^{(V,E,\varnothing,\overline{\vartheta},\vartheta),v}(t, X^{G,\xi})$ is a.s. continuous at $(\overline{\vartheta}', \vartheta') = (\varphi_*\overline{\kappa}, \varphi_*\kappa) = (\overline{\kappa}', \kappa')$.

Remark 5.2. Assumption 3 is a very mild assumption as it allows the jump rate function to have a large number of discontinuities so long as the set of times at which the jump rate function is discontinuous with respect to the initial data has Lebesgue measure 0. Markovian IPS of the form 3.15 trivially satisfy Assumption 3 as the jump rate functions $\tilde{r}_j^H$, $H \in \widehat{\mathscr{G}}_{*,1}$, $j \in \mathcal{J}$, do not depend on initial data for $t > 0$. The contact process in a heterogeneous environment of Example 3.17 likewise satisfies Assumption 3 as the dependence of $\tilde{r}_j^H$ on ϑ and $\overline{\vartheta}$ is continuous. Even the non-Markovian contact process of Example 3.18 will satisfy the assumption whenever h_I and h_R are only discontinuous on a subset of $\mathbb{R}_+$ with positive Lebesgue measure.

It is worth noting that for the remaining results in this section, we only apply Assumption 3 under conditions that also imply strong well-posedness of (3.1.4), in which case the strong solution $(G, X^{G,\xi})$ in the statement of the assumption is well defined.

For our main result, we show that a collection of IPS defined as random elements in $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{D}]$ will converge a.s. under certain conditions if they are defined as solutions to (3.1.4) for a.s. convergent initial data. Naturally, this is not generally true if different IPS are driven by different Poisson processes. Therefore we introduce the following concept.

Definition 5.3. Let (G_n, X^{G_n,ξ^n}), $n \in \mathbb{N}$ be a sequence of $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-valued IPS defined on the same filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and driven by the $\mathbb{F}$-driving noise $(G_n, \mathbf{N}^{G_n})$, $n \in \mathbb{N}$. Then $\{(G_n, X^{G_n,\xi^n})\}_{n\in\mathbb{N}}$ is said to be *communally driven* if there exists an i.i.d. collection of $\mathbb{F}$-Poisson processes $\{\mathbf{N}_k\}_{k\in\mathbb{N}}$ such that for each $n, k \in \mathbb{N}$, $\mathbf{N}_k = \mathbf{N}_k^{G_n}$ a.s. on the event $\{k \in V_{G_n}\}$.

Theorem 5.4. *Suppose Assumption 1 holds, and $\{(G_n, \xi^n)\}_{n\in\mathbb{N}} \cup \{(G, \xi)\}$ is a sequence of a.s. finitely dissociable $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-random elements (in the sense of Definition 4.15) defined in a* *common probability space such that (G,ξ) satisfies Assumption 3. If $(G_n,\xi^n) \to (G,\xi)$ a.s. in $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}},\mathcal{K}\times\mathcal{X}]$ and $\{(G_n,X^{G_n,\xi^n})\}_{n\in\mathbb{N}} \cup \{(G,X^{G,\xi})\}$ is communally driven, then $(G_n,X^{G_n,\xi^n}) \to (G,X^{G,\xi})$ a.s. in $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}},\mathcal{K}\times\mathcal{D}]$.*

Corollary 5.5 follows from a more general version of this statement proved in Proposition 5.10 under weaker assumptions that only require Assumption 1′, a consistent spatial localization condition introduced in Section 4.2.1 and a weaker finite convergence condition, Assumption 3′, in place of Assumption 3, which is useful for several applications.

An immediate consequence of Theorem 5.4 is that the law of the trajectory of the IPS is continuous in the local weak topology with respect to the law of the initial data of the IPS.

Corollary 5.5. *[36, Theorem 4.3] Suppose Assumptions 2 and 1 hold and $\{\langle G_n,\xi^n\rangle\}_{n\in\mathbb{N}}\cup\{\langle G,\xi\rangle\}$ is a sequence of a.s. finitely dissociable $\mathscr{G}_*[\overline{\mathcal{K}},\mathcal{K}\times\mathcal{X}]$-random elements (in the sense of Remark 4.17) such that $\langle G,\xi\rangle$ satisfies Assumption 3. If $\langle G_n,\xi^n\rangle \Rightarrow \langle G,\xi\rangle$ in $\mathscr{G}_*[\overline{\mathcal{K}},\mathcal{K}\times\mathcal{X}]$, then $\langle G_n,X^{G_n,\xi^n}\rangle \Rightarrow \langle G,X^{G,\xi}\rangle$ in $\mathscr{G}_*[\overline{\mathcal{K}},\mathcal{K}\times\mathcal{D}]$.*

5.2 Hydrodynamic Limit and Correlation Decay

5.2.1 An Annealed Correlation Decay Property

Recall the notion of convergence of probability in the local weak sense from Definition 2.2. As explained in the introduction and Section 2.1.2, the neighboring particles of an IPS on a sparse graph will generally be strongly correlated. In order to establish the hydrodynamic limit regardless, we require the IPS to satisfy a correlation decay property. The following correlation decay property holds for sequences of unrooted $[\overline{\mathcal{K}},\mathcal{K}\times\mathcal{X}]$-random graphs $\{(G_n,\xi^n)\}_{n\in\mathbb{N}}$ with finite vertex sets $\{V_n\}_{n\in\mathbb{N}}$.

Theorem 5.6. *[36, Theorem 4.5] Suppose Assumptions 1 and 2 hold, and let $\langle G,\xi\rangle$ be a $[\overline{\mathcal{K}},\mathcal{K}\times\mathcal{X}]$-random element that satisfies Assumption 3 and is a.s. finitely dissociable in the sense of Remark 4.17. Suppose there exists a sequence $\{(G_n,\xi^n)\}_{n\in\mathbb{N}}$ of finite, (possibly disconnected) unrooted $[\overline{\mathcal{K}},\mathcal{K}\times\mathcal{X}]$-random graphs such that $|V_{G_n}|\to\infty$ in probability and (G_n,ξ^n) converges in probability in the local weak sense to $\langle G,\xi\rangle$. For each $n\in\mathbb{N}$, let o_n^i, $i=1,2$, be independent, uniformly* *distributed vertices of G_n. Then for any bounded continuous functions $f_i : \mathscr{G}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{D}] \to \mathbb{R}$, $i = 1, 2$,*

$$\lim_{n\to\infty} Cov\Big(f_1\Big(\langle \mathscr{C}_{o_n^1}(G_n, X^{G_n,\xi^n})\rangle\Big), f_2\Big(\langle \mathscr{C}_{o_n^2}(G_n, X^{G_n,\xi^n})\rangle\Big)\Big) = 0, \tag{5.2.1}$$

where Cov represents the covariance functional and for each $n \in \mathbb{N}$, (G_n, X^{G_n,ξ^n}) is the strong solution to (3.1.4) *for the initial data (G_n, ξ^n).*

When the jump rate functions of the SDE (3.1.4) are strongly Lipschitz continuous in the sense described in the introduction, then arguments similar to those used for diffusions in [61, Lemma 5.2] can be applied to obtain stronger quantitative quenched (i.e., conditioned on the graph) bounds on the decay of correlation of IPS that are uniform with respect to graphs, and only depend on the cardinality of the sets of particles being compared and the graph distance between the sets. Under Assumption 1, such a strong Lipschitz condition holds for Markov IPS on graphs with bounded maximal degree, but fails to hold for many interesting IPS on graphs with unbounded maximal degree. In such situations, one does not expect there to be a similar quenched correlation bound that is uniform over all graphs (or even all finitely dissociable graphs), and thus the arguments we use to prove Theorem 5.6 are crucially different from those used in [61]. Specifically, to establish the *annealed asymptotic* correlation decay property (5.2.1), which is averaged over the randomness of the initial data, we first use the fact that the initial data satisfies an analogous asymptotic correlation decay property (due to the assumed convergence in probability in the local weak sense) and then carefully construct an appropriate coupling and the stronger almost sure local convergence result of Proposition 5.10 mentioned in Section 5.1 to extend the correlation decay property to the solution process.

5.2.2 Hydrodynamic Limits

The existence of the hydrodynamic limit follows as a simple consequence of well-posedness of the limit, local convergence and asymptotic correlation decay.

Theorem 5.7. *[36, Theorem 4.6] Suppose Assumptions 1 and 2 hold and the sequence of finite, unrooted $[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-random graphs $\{(G_n, \xi^n)\}_{n\in\mathbb{N}}$ and the $\mathscr{G}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-random element $\langle G, \xi\rangle$ satisfy the conditions of Theorem 5.6. Let (G_n, X^{G_n,ξ_n}) and $(G, X^{G,\xi})$ be weak solutions to* (3.1.4) *for (G_n, ξ_n) and (G, ξ), respectively. Then the sequence $\{(G, X^{G_n,\xi^n})\}_{n\in\mathbb{N}}$ converges in probability in the local weak sense to $\langle G, X^{G,\xi}\rangle$.*

Proof. First note that, given any driving noise, $\langle G, X^{G,\xi}\rangle$ and (G_n, X^{G_n,ξ^n}), $n \in \mathbb{N}$, are well defined unique strong solutions to (3.1.4) by Theorem 4.2, Proposition 4.1 and Lemma 3.11. For each $n \in \mathbb{N}$, let (o_n^1, o_n^2) be two i.i.d. uniformly distributed vertices in G_n. Then, by Theorem 5.6 and Corollary 5.5, for any bounded, continuous functions $f_1, f_2 : \mathscr{G}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{D}] \to \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{E}\left[\prod_{i=1}^{2} f_i(\langle \mathscr{C}_{o_n^i}(G_n, X^{G_n,\xi^n})\rangle)\right] = \lim_{n\to\infty} \prod_{i=1}^{2} \mathbb{E}\left[f_i(\langle \mathscr{C}_{o_n^i}(G_n, X^{G_n,\xi^n})\rangle)\right] = \prod_{i=1}^{2} \mathbb{E}\left[f_i(\langle G, X^{G,\xi}\rangle)\right].$$

That this implies the desired result follows from Lemma 2.5. □

We now show that Theorem 5.7 implies the convergence in probability of both the empirical measure and the empirical neighborhood distribution to corresponding deterministic limits.

Corollary 5.8. *[36, Corollary 4.7] Under the conditions of Theorem 5.7, the $\mathcal{P}(\mathcal{D})$-valued random empirical measure sequence $\{\pi^{G_n, X^{G_n,\xi^n}}\}_{n\in\mathbb{N}}$ converges in probability to Law$(X_\varnothing^{G,\xi}) \in \mathcal{P}(\mathcal{D})$, and the $\mathcal{P}(\mathscr{G}_{*,1}[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{D}])$-valued random empirical neighborhood distribution sequence $\{\overline{\pi}_\varnothing^{G_n, X^{G_n,\xi^n}}\}_{n\in\mathbb{N}}$ converges in probability to Law$(\langle B_1(G, X^{G,\xi})\rangle) \in \mathcal{P}(\mathscr{G}_{*,1}[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{D}])$.*

Proof. By Theorem 5.7, the sequence $\{(G_n, X^{G_n,\xi^n})\}_{n\in\mathbb{N}}$ converges in probability in the local weak sense to $\langle G, X^{G,\xi}\rangle$. Then the result holds by Lemma 2.4 for $\overline{Z} = \overline{\mathcal{K}}$, $\mathcal{Z} = \mathcal{K}$, $\mathcal{Z}' = \mathcal{D}$, $(G_n, z^n) = (G_n, X^{G_n,\xi^n})$ and $\langle G, z\rangle = \langle G, X^{G,\xi}\rangle$. □

As was already observed in [61, Theorem 6.4 and Proposition 7.7] in the context of interacting diffusions, the stronger convergence in probability in the local weak sense of the initial data is in general necessary to obtain deterministic hydrodynamic limits that coincide with Law$(X_\varnothing^{G,\xi})$ because if one only has convergence in the local weak sense of the initial data as in Corollary 5.5, the hydrodynamic limit can fail to be deterministic, or fail to coincide with Law$(X_\varnothing^{G,\xi})$ even when deterministic.

5.3 Proofs

The proof of all of the results stated in this chapter (with the exception of Corollary 5.8 are located in this section. We begin in Section 5.3.1 by stating and proving Proposition 5.10, which is a restatement of Theorem 5.4 under weaker assumptions. Then we show that Proposition 5.10 implies Theorem 5.4. In Section 5.3.2 we use Theorem 5.4 to prove Corollary 5.5, that is, we prove that the law of the IPS is continuous with respect to local weak convergence. Finally, in Section 5.3.3 we prove correlation decay (Theorem 5.6).

5.3.1 Proof of Almost Sure Convergence

In a broad sense, the proof of Theorem 5.4 proceeds by using spatial locality to reduce our analysis of the problem to the case in which all graphs are a.s. finite. Then through a careful analysis of (3.1.4) (utilizing Assumption 1 to limit the number of events of the driving Poisson processes that need to be considered and Assumption 3 to ensure that convergence of the initial data implies convergence of the jump rate functions), we show that a.s. convergence of the initial data in the space $\widehat{\mathscr{G}}_*[\mathcal{K}, \mathcal{K} \times \mathcal{X}]$ implies a.s. convergence of the trajectories of the IPS in $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{D}]$. For this reason, we can weaken the conditions needed for the theorem to hold by introducing the following "finite convergence" condition on IPS defined on finite truncations of graphs (equipped with the associated marks). It is applied to a family of jump rate functions $\mathbf{r}$ and a sequence of $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K}]$-initial data $\{(G_n, \xi^n)\}_{n\in\mathbb{N}} \cup \{(G, \xi)\}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.

Assumption 3′. Assumption 1′ holds. Furthermore, if $\{X^{m,n}\}_{n,m\in\mathbb{N}} \cup \{X^{m,\infty}\}_{m\in\mathbb{N}}$ is a communally driven collection of IPS defined as strong solutions to (3.1.4) for the respective initial data $\{B_m(G_n, \xi^n)\}_{n,m\in\mathbb{N}} \cup \{B_m(G, \xi)\}_{m\in\mathbb{N}}$, then for each $m \in \mathbb{N}$ and $T \in \mathbb{R}_+$, there exists a $\mathcal{F}_T$-measurable random variable $\overline{N}_m := \overline{N}_{m,T}$ such that for every $n \in \mathbb{N}$,

$$(5.3.1) \qquad (B_m([G_{n,*}]), X^{m,n}[T]) = (B_m([G_*]), X^{m,\infty}[T]) \text{ a.s. on the event } \{n \geq \overline{N}_m\} \cap \widehat{\mathcal{I}},$$

where $\widehat{\mathcal{I}} = \{B_m([G_{n,*}], \xi^n) = B_m([G_*], \xi)\}$.

Note that Assumption 3 is similar in concept to [36, Assumption 2']. We believe that the two assumptions are equivalent, however we do not rigorously prove this as our use of Assumption 3 is merely intended to simplify the arguments of this chapter which are analogous to arguments presented in [36, Sections 6-7].

The following lemma shows that the finite convergence property of Assumption 3′ holds under our basic Assumptions 1 and 3, together with consistent spatial localization. We defer the proof of the lemma to the end of the section.

Lemma 5.9. *Fix a filtered probability space* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ *such that* $\mathbb{F}$ *satisfies the usual conditions. Suppose the family of jump rate functions* $\mathbf{r}$ *satisfies Assumption 1, and* (G, ξ) *satisfies Assumption 3. If the* $\mathcal{F}_0$*-measurable sequence* $\{G_n\}_{n\in\mathbb{N}} \cup \{G\}$ *of connected graphs in* $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K}]$*-random elements a.s. consistently spatially localizes the SDE* (3.1.4)*, then Assumption 3′ holds.*

The weaker Assumption 3′ has been introduced in place of Assumptions 1 and Assumption 3 with a view to future extensions of our results to more general IPS for which the latter assumptions may not hold but the former assumptions and consistent spatial localization can nevertheless be established. In the following proposition, let $G_n = (V_n, E_n, \varnothing_n, \overline{\kappa}^n, \kappa^n)$, $n \in \mathbb{N}$ and $G = (V, E, \varnothing, \overline{\kappa}, \kappa)$.

Proposition 5.10. *Suppose that Assumption 3′ holds and that* $\{\mathscr{C}_{\varnothing_n}(G_n)\}_{n\in\mathbb{N}} \cup \{\mathscr{C}_{\varnothing}(G)\}$ *a.s. consistently spatially localizes the SDE* (3.1.4)*. Also, let* $\{(G_n, X^{G_n,\xi^n})\}_{n\in\mathbb{N}} \cup \{(G, X^{G,\xi})\}$ *be a collection of IPS communally driven by* $\mathbf{N}$*. If* $(G_n, \xi^n) \to (G, \xi)$ *a.s. in* $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$*, then*

$$\lim_{\substack{n\to\infty \\ v\in V_n}} X_v^{G_n,\xi^n} = X_v^{G,\xi} \text{ a.s.,} \tag{5.3.2}$$

for all $v \in V_{\mathscr{C}_{\varnothing}(G)}$*. In particular, if* G *is connected, then* $(G_n, X^{G_n,\xi^n}) \to (G, X^{G,\xi})$ *a.s. in* $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{D}]$.

We first show that Proposition 5.10 implies Theorem 5.4, then we prove the proposition.

Proof of Theorem 5.4. First suppose that G is connected. By Lemma 5.9, Assumptions 1 and 3 imply Assumption 3′. The a.s. finite dissociability of the graphs G_n, $n \in \mathbb{N}$, G implies that $\mathscr{C}_{\varnothing_n}(G_n)$, $n \in \mathbb{N}$, $\mathscr{C}_{\varnothing}(G) = G$ consistently spatially localize the SDE (3.1.4) by Proposition 4.22. Then the result follows from Proposition 5.10.

Now suppose that G is not connected. Then by Proposition 4.22, for each $v \in V_{\mathscr{C}_{\varnothing}(G)}$, $\lim_{\substack{n\to\infty \\ v\in V_n}} X_v^{G_n,\xi^n} = X_v^{G,\xi}$ a.s.. Now let $v \in V$ be any vertex that is not in the same connected component of G as $\varnothing$. For each $n \in \mathbb{N}$ let $\varnothing_n^v = v$ if $v \in V_n$ and $\varnothing_n$ otherwise. Then because $(G_n, \xi^n) \to (G, \xi)$ a.s., conditions 1,2,4,5 and 6 of Definition 2.19 are satisfied with regard to the convergence of $(G_n, \varnothing_n^v, \xi^n)$ to (G, v, ξ). Condition 3 follows from condition 1 of Definition 2.19 which states that $v \in V_n$ for all sufficiently large n. Thus, $\varnothing_n^v \to v$ a.s.. Thus, $(G_n, \varnothing_n^v, \xi^n) \to (G, v, \xi)$ a.s.. Because finite dissociability does not depend on the choice of root, $(G_n, \varnothing_n^v, \xi^n)$, $n \in \mathbb{N}$, (G, v, ξ) are finitely dissociable. This implies that $\mathscr{C}_{\varnothing_n^v}(G_n)$, $n \in \mathbb{N}$, $\mathscr{C}_v(G)$ consistently spatially localize the SDE (3.1.4) by Proposition 4.22. Thus, by Lemma 5.10,

$$\lim_{\substack{n\to\infty\\ v\in V_n}} X_v^{G_n,\xi^n} = X_v^{G,\xi}.$$

The limit above holds for all $v \in V$, so by Lemma 2.21, $(G_n, X^{G_n,\xi^n}) \to (G, X^{G,\xi})$ a.s. as desired. $\square$

Proof of Proposition 5.10. If G is connected, then $V_{\mathscr{C}_\varnothing(G)} = V$, so (5.3.2) and the fact that $(G_n, \xi^n) \to (G,\xi)$ a.s. imply by Lemma 2.21 that $(G_n, X^{G_n,\xi^n}) \to (G, X^{G,\xi})$ a.s., which proves the final assertion of the proposition. Therefore, it suffices to prove (5.3.2).

Let $\mathbb{N}_\infty := \mathbb{N} \cup \{\infty\}$ and let $(G_\infty, \xi^\infty) := (G,\xi)$, and $(G_\infty, X^{G_\infty,\xi^\infty}) := (G, X^{G,\xi})$ respectively. Note that if $\mathbb{P}(X_v^{G_n,\xi^n} \to X_v^{G,\xi} | \mathcal{F}_0) = 1$ a.s. for each $v \in V$, then we are done. Therefore, we may assume without loss of generality that the initial data $\{(G_n,\xi^n)\}_{n\in\mathbb{N}\cup\{\infty\}}$ is deterministic. For each $n \in \mathbb{N}'$, let $(G'_n, \xi'^{,n}) = \mathscr{C}_{\varnothing_n}(G_n, \xi^n)$ and let $\mathbf{N}^{G'_n} = \mathbf{N}^{G_n}_{V_{G'_n}}$. Recall from Remark 2.24 that $(G'_n, \xi'^{,n})$ may fail to converge to (G', ξ') in $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$.

Due to Assumption $1'$ (which holds by Assumption $3'$) and the assumption of consistent spatial localization, there exists a consistent sequence of localizing maps $\{\mathcal{S}_T(\cdot; G'_n, \mathbf{N}^{G'_n})\}_{n\in\mathbb{N}_\infty}$ for the SDE (3.1.4) on $\{(G'_n, \mathbf{N}^{G'_n})\}_{n\in\mathbb{N}_\infty}$, and by Proposition 4.10, for each $n \in \mathbb{N}_\infty$, the SDE (3.1.4) is strongly well-posed for $(G'_n, \xi'^{,n})$. Thus, the $\mathbf{N}^{G'_n}$-strong solution $(G'_n, X^{G'_n,\xi'^{,n}})$ to (3.1.4) for the initial data $(G'_n, \xi'^{,n})$ is well-defined. Furthermore, this implies that for each $n \in \mathbb{N} \cup \infty$ and any $(\mathcal{F}, \mathbf{N}^{G_n})$-weak solution X^{G_n,ξ^n} to (3.1.4) for the initial data (G_n, ξ^n) and any $v \in V_{G'_n}$, $X_v^{G_n,\xi^n} = X_v^{G'_n,\xi'^{,n}}$ a.s. (see Remark 4.12). Furthermore, if $v \in V_{G'}$, then there must exist an $m \in \mathbb{N}$ such that $v \in V_{B_m(G')} = V_{B_m(G)}$. Then by Lemma 2.22, $v \in V_{B_m(G_n)} = V_{B_m(G'_n)} \subseteq V_{G'_n}$ for all sufficiently large n. Thus, it suffices to prove that for each $v \in V_{G'}$,

$$\lim_{\substack{n\to\infty\\ v\in V_{G'_n}}} X_v^{G'_n,\xi'^{,n}} = X_v^{G',\xi'} \text{ a.s.}.$$

For each $n \in \mathbb{N}$, let M_n be the maximal constant such that $B_{M_n}([G'_{n,*}], \xi'^{,n}) = B_{M_n}([G_{n,*}], \xi^n) = B_{M_n}([G_*], \xi) = B_{M_n}([G'_*], \xi')$. Note that because $([G_{n,*}], \xi^n)$ converges a.s. to $([G_*], \xi^n)$ in $\widehat{\mathscr{G}}_*[\{1\}, \mathcal{X}]$, $M_n \to \infty$ as $n \to \infty$ by Lemma 2.22 and condition 5 of Definition 2.19.

Fix $n', m, M \in \mathbb{N}$ such that $m \le M \le M_{n'}$, which implies that $B_M([G'_{n',*}]) = B_M([G'_*])$. Because the IPS are communally driven, for every $n \in \mathbb{N}$, $m \le M_n$ and $v \in V_{B_m(G')}$, $\mathbf{N}_v^{G'_n} = \mathbf{N}_v^{G'}$.

Thus, by an application of condition 2 of the consistent spatial localization property in Definition 4.7, with $\ell = M$, $n = \infty$, $n' = n'$ and $U = B_m(G')$, and noting that then $\mathcal{I} \supseteq \{M \leq M_{n'}\} = \Omega$, we see that

$$\mathcal{S}_T(U; G', \mathbf{N}^{G'}) = \mathcal{S}_T(U; G'_{n'}, \mathbf{N}^{G'_{n'}}) \text{ a.s. on } \{\mathcal{S}_T(U; G', \mathbf{N}^{G'}) \subseteq B_M(G')\}.$$

Because $\mathcal{S}_T(U; G', \mathbf{N}^{G'})$ is a.s. finite, there must exist an $\mathcal{F}_T$-measurable a.s. finite random variable $\overline{M}_m$ such that $\mathcal{S}_T(U; G', \mathbf{N}^{G'}) \subseteq B_{\overline{M}_m}(G')$ a.s.. Furthermore, by (4.2.3) this implies that

$$X_U^{G',\xi'}[T] = X_U^{B_M(G'),\xi'_{B_M(G')}}[T] \quad \text{and} \quad X_U^{G'_{n'},\xi',n'}[T] = X_U^{B_M(G'_{n'}),\xi',n'_{B_M(G'_{n'})}}[T] \text{ a.s. on } \{M \geq \overline{M}_m\}.$$

By (5.3.1) of Assumption $3'$, there exists an a.s. finite, $\mathcal{F}_T$-measurable random variable N_M such that

$$X_U^{B_M(G'),\xi'_{B_M(G')}}[T] = X_U^{B_M(G'_{n'}),\xi',n'_{B_M(G'_{n'})}}[T] \text{ a.s. on the event } \{n' \geq N_M\}.$$

The last two displays together show that a.s. on the event $\{M \geq \overline{M}_m\} \cap \{n' \geq N_M\}$,

$$X_U^{G',\xi'}[T] = X_U^{B_M(G'),\xi'_{B_M(G')}}[T] = X_U^{B_M(G'_{n'}),\xi',n'_{B_M(G'_{n'})}}[T] = X_U^{G'_{n'},\xi',n'}[T].$$

Applying the last display for each M satisfying $m \leq M \leq M_{n'}$ and noting by condition 1 of Definition 4.4 that $\overline{M}_m \geq m$, it follows that

$$X_U^{G',\xi'}[T] = X_U^{G'_{n'},\xi',n'}[T] \text{ a.s. on the event } \{\overline{M}_m \leq M_{n'}\} \cap \{n' \geq N_{\overline{M}_m}\}.$$

Sending $n' \to \infty$ and noting that $M_{n'} \to \infty$, $\overline{M}_m$ is a.s. finite and that $U = V_{B_m(G')} = V_{B_m(G'_{n'})}$ for n' such that $m \leq M_{n'}$, it follows that for each $v \in B_m(G')$,

$$\lim_{\substack{n' \to \infty \\ v \in B_m(G'_{n'})}} X_v^{G'_{n'},\xi',n'}[T] = X_v^{G',\xi'}[T] \text{ a.s..}$$

As this holds for arbitrary $T > 0$, $m \in \mathbb{N}$ and therefore arbitrary $v \in V_{G'}$,

$$\lim_{\substack{n' \to \infty \\ v \in V_{G'_{n'}}}} X_v^{G'_{n'},\xi',n'} = X_v^{G',\xi'} \text{ a.s. for all } v \in V_{\mathscr{C}_\varnothing(G)},$$

which concludes the proof. $\square$

We finish the section with a proof of Lemma 5.9.

Proof of Lemma 5.9: Write $G = (V, E, \varnothing, \overline{\kappa}, \kappa)$, $G_n = (V_n, E_n, \varnothing_n, \overline{\kappa}^n, \kappa^n)$, $n \in \mathbb{N}$, and set $M := \max_{v \in V} d_G(v, \varnothing)$. We first consider the case when $\{G_n\}_{n\in\mathbb{N}}$ and G are a.s. finite, connected graphs and additionally assume that $\max_{v\in V_n} d_{G_n}(v, \varnothing_n) \le M$ for all $n \in \mathbb{N}$, and show that for any $T \in \mathbb{R}_+$, there exists an a.s. finite, $\mathcal{F}_T$-measurable random variable $\overline{N} := \overline{N}_{M,T}$ such that,

$$\Big([G_{n,*}], X^{G_n,\xi^n}[T]\Big) = \Big([G_*], X^{G,\xi}[T]\Big) \text{ a.s. on the event } \{n \ge \overline{N}\} \cap \widehat{\mathcal{I}}_n, \tag{5.3.3}$$

where $\widehat{\mathcal{I}}_n := \{B_M([G_{n,*}], \xi^n) = B_M([G_*], \xi)$ and $X^{M,n} := X^{G_n,\xi^n}$, $n \in \mathbb{N}$, and $X^{M,\infty} := X^{G,\xi}$ are communally driven. To show (5.3.3), fix $T \in \mathbb{R}_+$. Because the IPS are communally driven, $\mathbf{N}_v^{G_n} = \mathbf{N}_v^{G}$ for each $n \in \mathbb{N}$ and $v \in V$ on the event $\widehat{\mathcal{I}}_n$. In terms of the function C from Assumption 1, define

$$\overline{\mathcal{E}} = \overline{\mathcal{E}}_T := \Big\{(v,t,r,j) \in V \times [0,T] \times (0, C(|\mathrm{cl}_v(G)|, T)] \times \mathcal{J} : \mathbf{N}_v^G\left(\{(t,r,j)\}\right) = 1\Big\}.$$

Since $|\overline{\mathcal{E}}| < \infty$ a.s., we can a.s. order the elements $\{(v_k, \tau_k, r_k, j_k)\}_{k=1}^{|\overline{\mathcal{E}}|}$ of $\overline{\mathcal{E}}$ such that $\{\tau_k\}_{k\in\mathbb{N}}$ is strictly increasing. Note that $\{\tau_k\}_{k\in\mathbb{N}}$ is the sequence of points in a time-homogeneous Poisson process, which implies that each τ_k is an absolutely continuous (Gamma distributed) non-negative random variable. Thus, for any Borel set $O \subset [0,T]$ such that $\mathrm{Leb}(O) = 0$, $\mathbb{P}(\tau_k \in O \text{ for any } k) = \sum_{k=1}^{\infty} \mathbb{P}(\tau_k \in O, k \le |\overline{\mathcal{E}}|) = 0$. Let $R := \inf\Big\{|r_k - r_{j_k}^{G,v_k}(\tau_k, X^{G,\xi})| : k = 1, \ldots, |\overline{\mathcal{E}}|\Big\}$. By the predictability of the jump rate function $r_{j_k}^{G,v_k}$, the random variables r_k and $r_{j_k}^{G,v_k}(\tau_k, X^{G,\xi})$ are independent and therefore $r_k \ne r_{j_k}^{G,v_k}(\tau_k, X^{G,\xi})$ a.s. which implies that $R > 0$ a.s.. On the event $\widehat{\mathcal{I}}_n$ (noting that $V_n = V$), let

$$R_n := \sup\left\{\left|r_{j_k}^{G,v_k}(\tau_k, X^{G,\xi}) - r_{j_k}^{G_n,v_k}(\tau_k, X_{V_n}^{G,\xi})\right| : k = 1, \ldots, |\overline{\mathcal{E}}|\right\}.$$

By Lemma 2.22 (and the fact that G is a.s. finite), $[G_{n,*}] = [G_*]$ for all sufficiently large n a.s.. By condition 5 of Definition 2.19, $\xi_v^n = \xi_v$ a.s. for all $v \in V$ and sufficiently large n. Thus, $\{\widehat{\mathcal{I}}_n\}_{n\in\mathbb{N}}$ is a sequence of events such that $\mathbb{P}(\cap_{n'\ge n}\widehat{\mathcal{I}}_n) \to 1$ so that $\lim_{n\to\infty} \mathbb{I}_{\{\widehat{\mathcal{I}}_n\}} = 1$ a.s.. By conditions 4 and 5 of Definition 2.19, it follows that for each $v \in V$ and $e \in E$,

$$\lim_{n\to\infty} \mathbb{I}_{\widehat{\mathcal{I}}_n}(\overline{\kappa}_e^n, \kappa_v^n) = (\overline{\kappa}_e, \kappa_v) \quad \text{a.s..} \tag{5.3.4}$$

Combining (5.3.4) with the absolute continuity of τ_k and the fact that (G, ξ) satisfies Assumption

3, we have for every $k \in \mathbb{N}$,

$$\lim_{n\to\infty} r_{j_k}^{G_n,v_k}\left(\tau_k, X_{V_n}^{G,\xi}\right)\mathbb{I}_{\widehat{\mathcal{I}}_n} = \lim_{n\to\infty} r_{j_k}^{(V,E,\varnothing,\overline{\kappa}_E^n,\kappa_V^n),v_k}\left(\tau_k, X^{G,\xi}\right)\mathbb{I}_{\widehat{\mathcal{I}}_n} = r_{j_k}^{G,v_k}\left(\tau_k, X^{G,\xi}\right)$$

a.s. on the event $\{k \le |\overline{\mathcal{E}}|\}$. Because $|\overline{\mathcal{E}}|$ is a.s. finite, it follows that $\lim_{n\to\infty} R_n = 0$ a.s.. Moreover, the fact that $R > 0$ a.s. implies the existence of an a.s. finite random variable $\overline{N}$ such that,

$$R_n < \frac{R}{2} \text{ on the event } \{n \ge \overline{N}\} \cap \widehat{\mathcal{I}}_n. \tag{5.3.5}$$

We now argue that $X^{G,\xi}[T] = X^{G_n,\xi^n}[T]$ on the event $\{n \ge \overline{N}\} \cap \widehat{\mathcal{I}}_n$ by making use of the fact that for each $n \in \mathbb{N}$ and $v \in V$, $X_v^{G,\xi}$ and $X_v^{G_n,\xi^n}$ are communally driven on the event $\widehat{\mathcal{I}}_n$. Fix $n \in \mathbb{N}$ and note that by the SDE (3.1.4), $X^{G,\xi}$ and X^{G_n,ξ^n} are both a.s. continuous on the random set $\{t \in [0,\infty) \setminus \{\tau_k\}_{k\in\mathbb{N}}\} \cap \widehat{\mathcal{I}}_n$. Furthermore, at time τ_k, the processes $X^{G,\xi}$ and X^{G_n,ξ^n} may either remain constant or experience a jump of size j_k at the vertex v_k. It follows from the SDE (3.1.4) that the processes will either simultaneously jump or both fail to jump if and only if

$$\operatorname{sgn}\left(r_k - r_{j_k}^{G,v_k}(\tau_k, X^{G,\xi})\right) = \operatorname{sgn}\left(r_k - r_{j_k}^{G_n,v_k}(\tau_k, X^{G_n,\xi^n})\right),$$

where $\operatorname{sgn} : \mathbb{R} \to \mathbb{R}$ is the càdlàg function given by $\operatorname{sgn}(a) = \mathbb{I}_{\{a\ge 0\}} - \mathbb{I}_{\{a<0\}}$. Suppose that $X^{G,\xi}[\tau_k) = X^{G_n,\xi^n}[\tau_k)$. Then, using the predictability of the jump rates (by the standing assumption) in the first equality below, we have

$$\begin{aligned}
&\operatorname{sgn}\left(r_k - r_{j_k}^{G_n,v_k}(\tau_k, X^{G_n,\xi^n})\right)\\
&\qquad= \operatorname{sgn}\left(r_k - r_{j_k}^{G_n,v_k}(\tau_k, X^{G,\xi})\right)\\
&\qquad= \operatorname{sgn}\left(r_k - r_{j_k}^{G,v_k}(\tau_k, X^{G,\xi}) + r_{j_k}^{G,v_k}(\tau_k, X^{G,\xi}) - r_{j_k}^{G_n,v_k}(\tau_k, X^{G,\xi})\right).
\end{aligned}$$

However, the last line of the above display is a.s. equal to $\operatorname{sgn}\left(r_k - r_{j_k}^{G,v_k}(\tau_k, X^{G,\xi})\right)$ on the event $\{n \ge \overline{N}\}$ by (5.3.5). Thus, $X^{G,\xi}[\tau_k] = X^{G_n,\xi^n}[\tau_k]$ a.s. on the event $\{n \ge \overline{N}\} \cap \widehat{\mathcal{I}}_n$. It follows that $X^{G,\xi}[\tau_{k+1}) = X^{G_n,\xi^n}[t_{k+1})$ a.s. if $|\overline{\mathcal{E}}| > k$ and $X^{G,\xi}[T] = X^{G_n,\xi^n}[T]$ a.s. if $|\overline{\mathcal{E}}| = k$. Applying induction, we see that $X^{G,\xi}[T] = X^{G_n,\xi^n}[T]$ a.s. on the event $\{n \ge \overline{N}\} \cap \widehat{\mathcal{I}}_n$, and so (5.3.3) follows from the fact that $[G_{n,*}] = [G_*]$ on the event $\widehat{\mathcal{I}}_n$.

To see why this implies Lemma 5.9 in the general case of possibly infinite, disconnected graphs, note that (5.3.3) implies that for any $m \in \mathbb{N}$ and any sequence $\{G_n\}_{n\in\mathbb{N}}, G$ we may replace $(G_n, \xi^n), n \in \mathbb{N}$ by $B_m(G_n, \xi^n), n \in \mathbb{N}$, (G, ξ) by $B_m(G, \xi)$, $\overline{N}$ by $\overline{N}_m$ and M by m in (5.3.3) to get (5.3.1). This concludes the proof of the lemma. $\square$

5.3.2 Proof of Local Weak Convergence

We now show how Corollary 5.5 follows from Theorem 5.4. The proof uses a simple argument involving the Skorokhod representation theorem and Theorem 5.4.

Proof of Corollary 5.5. Set $\langle G_\infty, \xi^\infty\rangle := \langle G, \xi\rangle$ and $\mathbb{N}_\infty := \mathbb{N} \cup \{\infty\}$. Since $\langle G_n, \xi^n\rangle \Rightarrow \langle G, \xi\rangle$ in $\mathscr{G}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$, by the Skorokhod representation theorem there exists a (complete) probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ that supports random elements $\langle \widetilde{G}_n, \widetilde{\xi}^n\rangle \overset{(\mathrm{d})}{=} \langle G_n, \xi^n\rangle$, $n \in \mathbb{N}_\infty$, such that $\langle \widetilde{G}_n, \widetilde{\xi}^n\rangle \to \langle \widetilde{G}, \widetilde{\xi}\rangle := \langle \widetilde{G}_\infty, \widetilde{\xi}^\infty\rangle$ $\widetilde{\mathbb{P}}$-a.s.. Because Assumption 2 holds, Theorem 4.2 and Remark 4.3 imply that (3.1.4) is strongly well-posed for all initial data in $\{\langle \widetilde{G}_n, \widetilde{\xi}^n\rangle\}_{n\in\mathbb{N}_\infty}$. By Theorem 2.29, there exists a sequence of connected, $\widetilde{\mathcal{F}}$-measurable representatives $(\widetilde{G}_n, \widetilde{\xi}^n)$, $n \in \mathbb{N}_\infty$, of the respective random isomorphism classes $\langle \widetilde{G}_n, \widetilde{\xi}^n\rangle$, $n \in \mathbb{N}_\infty$, such that $(\widetilde{G}_n, \widetilde{\xi}^n) \to (\widetilde{G}, \widetilde{\xi}) := (\widetilde{G}_\infty, \widetilde{\xi}^\infty)$ a.s. in $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K}\times\mathcal{X}]$. By Remark 4.17, these graphs are a.s. finitely dissociable. Lastly, suppose without loss of generality that $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ supports a filtration $\widetilde{\mathbb{F}}$ satisfying the usual conditions such that $\{(G_n, \xi^n)\}_{n\in\mathbb{N}_\infty}$ is $\widetilde{\mathcal{F}_0}$ and an i.i.d. sequence of $\widetilde{\mathbb{F}}$-Poisson processes $\mathbf{N} := \{\mathbf{N}_v\}_{v\in\mathbb{N}}$ on $\mathbb{R}_+^2 \times \mathcal{J}$ with intensity measure $\mathrm{Leb}^2 \otimes \#_{\mathcal{J}}$.

Let $X^{\widetilde{G}_n,\widetilde{\xi}^n}$, $n \in \mathbb{N}_\infty$, be strong solutions to (3.1.4) for the respective initial data $(\widetilde{G}_n, \widetilde{\xi}^n)$, $n \in \mathbb{N}_\infty$, communally driven by $\mathbf{N}$, noting that these quantities are all well defined by the strong well-posedness of (3.1.4) for $(\widetilde{G}_n, \widetilde{\xi}^n)$, $n \in \mathbb{N}_\infty$, by Theorem 4.2. Because Assumptions 1 and 3 also hold, the conditions of Theorem 5.4 are satisfied, so $(\widetilde{G}_n, X^{\widetilde{G}_n,\widetilde{\xi}^n}) \to (\widetilde{G}, X^{\widetilde{G},\widetilde{\xi}}) := (\widetilde{G}_\infty, X^{\widetilde{G}_\infty,\widetilde{\xi}^\infty})$ a.s. in $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{D}]$. By Proposition 2.26, Definition 3.14 and Lemma 3.11, this implies that the strong solutions $\langle \widetilde{G}_n, X^{\widetilde{G}_n,\widetilde{\xi}^n}\rangle \overset{(\mathrm{d})}{=} \langle G_n, X^{G_n,\xi^n}\rangle$, $n \in \mathbb{N}$, to (3.1.4) for the respective initial data $\langle \widetilde{G}_n, \widetilde{\xi}^n\rangle \overset{(\mathrm{d})}{=} \langle G_n, \xi^n\rangle$, $n \in \mathbb{N}$, converge locally a.s. to the strong solution $\langle \widetilde{G}, X^{\widetilde{G},\widetilde{\xi}}\rangle \overset{(\mathrm{d})}{=} \langle G, X^{G,\xi}\rangle$, $n \in \mathbb{N}$, to (3.1.4) for the initial data $\langle \widetilde{G}, \widetilde{\xi}\rangle \overset{(\mathrm{d})}{=} \langle G, \xi\rangle$. Thus, $\langle G_n, X^{G_n,\xi^n}\rangle \Rightarrow \langle G, X^{G,\xi}\rangle$ as desired. $\square$

5.3.3 Proof of Asymptotic Correlation Decay

This section is devoted to the proof of Theorem 5.6. Recall that for an unrooted $[\overline{\mathcal{K}}, \mathcal{K}]$-graph G and a vertex $v \in V_G$, $\mathscr{C}_v(G)$ is the connected component of G equipped with v as its root. For the remainder of the section, we fix a sequence of finite, (possibly disconnected) unrooted

$[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-random graphs (G_n, ξ^n), $n \in \mathbb{N}$, and a $\mathscr{G}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-random element $\langle \mathscr{C}_\varnothing(G, \xi)\rangle$ (henceforth, denoted just $\langle G, \xi\rangle$), all defined on a common complete probability space $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$. We additionally assume that, by extending the probability space if necessary, $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$ also supports an i.i.d. pair of vertices (o_n^1, o_n^2), each uniformly distributed on G_n, for all $n \in \mathbb{N}$. Invoking the Skorokhod representation theorem, the first and main step of the proof assumes joint local convergence of the initial data $\{\langle \mathscr{C}_{o_n^1}(G_n, \xi^n)\rangle, \langle \mathscr{C}_{o_n^2}(G_n, \xi^n)\rangle\}_{n \in \mathbb{N}}$ to the i.i.d. pair $(\langle G^{(1)}, \xi^{(1)}\rangle, \langle G^{(2)}, \xi^{(2)}\rangle)$ and proves convergence *in probability* of the corresponding pairs of strong solutions $\{\langle \mathscr{C}_{o_n^1}(G_n, X^{G_n,\xi^n})\rangle, \langle \mathscr{C}_{o_n^2}(G_n, X^{G_n,\xi^n})\rangle\}_{n\in\mathbb{N}}$ to $(\langle G^{(1)}, X^{G^{(1)},\xi^{(1)}}\rangle, \langle G^{(2)}, X^{G^{(2)},\xi^{(2)}}\rangle)$ for an appropriate choice of driving processes. The coupling proof is a fairly simple consequence of Corollary 2.30. This result then implies the correlation decay result.

Lemma 5.11. *Fix $\{(G_n, \xi^n)\}_{n\in\mathbb{N}}$ and suppose that the vertices of G_n, $n \in \mathbb{N}$ lie in $\mathbb{N}$. Suppose also that Assumption 2 holds and that the collection of isomorphism classes $\{\langle \mathscr{C}_{o_n^i}(G_n)\rangle, \langle G^{(i)}\rangle\}_{n\in\mathbb{N}, i=1,2}$ a.s. consistently spatially localizes the SDE* (3.1.4). *In addition, assume that for each $i = 1, 2,$,*

$$\langle \mathscr{C}_{o_n^i}(G_n, \xi^n)\rangle \to \langle G^{(i)}, \xi^{(i)}\rangle \tag{5.3.6}$$

and the tuple $(\{\langle \mathscr{C}_{o_n^i}(G_n, \xi^n)\rangle\}_{n\in\mathbb{N}}, \langle G^{(i)}, \xi^{(i)}\rangle)$ satisfies Assumption 3′. Then it is possible to define a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ supporting solutions $(G_n, X^{G_n,\xi^n}), n \in \mathbb{N}$, and $\langle G^{(i)}, X^{G^{(i)},\xi^{(i)}}\rangle$, $i = 1, 2$, to the SDE (3.1.4) *for the respective initial data $(G_n, \xi^n), n \in \mathbb{N}$, and $\langle G^{(i)}, \xi^{(i)}\rangle$, $i = 1, 2$, such that as $n \to \infty$,*

$$\langle \mathscr{C}_{o_n^i}(G_n, X^{G_n,\xi^n})\rangle \to \langle G^{(i)}, X^{G^{(i)},\xi^{(i)}}\rangle, \;\; i = 1, 2, \;\; \textit{in probability.} \tag{5.3.7}$$

Proof. Let $L_n = d_{G_n}(o_n^1, o_n^2)$ and note that $L_n \to \infty$ in probability by [88, Corollary 2.21]. By assumption, $\langle \mathscr{C}_{o_n^i}(G_n, \xi^n)\rangle \to \langle G^{(i)}, \xi^{(i)}\rangle$ a.s. for each $i = 1, 2$. Thus, by Corollary 2.30, for each $n \in \mathbb{N}$, there exists a doubly rooted, $\widehat{\mathcal{F}}$-measurable $[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-random graph $(\widetilde{G}_n, \widetilde{o}_n^1, \widetilde{o}_n^2, \widetilde{\xi}^n) \cong (G_n, o_n^1, o_n^2, \xi^n)$ with vertices in $\mathbb{N}$ and a doubly rooted, $\widehat{\mathcal{F}}$-measurable $[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-random graph (G, o^1, o^2, ξ) with vertices in $\mathbb{N}$ such that for each $i = 1, 2$, $(\widetilde{G}_n, \widetilde{o}_n^i, \widetilde{\xi}^n)$ converges in probability in $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$ to (G, o^i, ξ) and for each $i = 1, 2$, $\mathscr{C}_{o^i}(G, \xi)$ is a $\widehat{\mathcal{F}}$-measurable representative of $\langle G^{(i)}, \xi^{(i)}\rangle$.

Extend the probability space $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$ to a complete filtered probability $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ such that $\mathbb{F}$ satisfies the usual conditions, any $\widehat{\mathcal{F}}$-measurable random element is $\mathcal{F}_0$-measurable in $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$

and $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ supports a collection of i.i.d. $\mathbb{F}$-Poisson processes $\mathbf{N} := \{\mathbf{N}_k\}_{k\in\mathbb{N}}$ on $\mathbb{R}^2_+ \times \mathcal{J}$ with intensity $\mathrm{Leb}^2 \otimes \#_{\mathcal{J}}$. Note that by Corollary 2.23, G is a disjoint union of the two connected graphs $\mathscr{C}_{o^i}(G)$, $i = 1, 2$, each of which spatially localizes the SDE (3.1.4). By Assumption $1'$ and the spatial localization assumption, Proposition 4.10 implies that the SDE (3.1.4) is strongly well-posed for all initial data in the collection of marked graphs $\{(\widetilde{G}_n, \widetilde{\xi}^n), (G, \xi)\}_{n\in\mathbb{N}}$. Thus, for each $n \in \mathbb{N}$, let $(\widetilde{G}_n, X^{\widetilde{G}_n,\widetilde{\xi}^n})$ and $(G, X^{G,\xi})$ be the unique strong solutions to (3.1.4) communally driven by $\mathbf{N}$ in the sense of Definition 5.3.

For any subsequence $\{n_k\} \subseteq \mathbb{N}$, there must exist a further subsequence $\{n_{k_\ell}\} \subseteq \{n_k\}$ such that for each $i = 1, 2$,

$$(\widetilde{G}_{n_{k_\ell}}, \widetilde{o}^i_{n_{k_\ell}}, \widetilde{\xi}^{n_{k_\ell}}) \to (G, o^i, X^{G,\xi}) \text{ a.s..}$$

Then applying Proposition 5.10 once for each $i = 1, 2$, it follows that for each $v \in \cup_{i=1,2} V_{\mathscr{C}_{o^i}(G)} = V_G$,

$$\lim_{\substack{\ell\to\infty \\ v\in V_{\widetilde{G}_{n_{k_\ell}}}}} X_v^{\widetilde{G}_{n_{k_\ell}},\widetilde{\xi}^{n_{k_\ell}}} = X_v^{G,\xi} \text{ a.s..}$$

By Lemma 2.21, this implies that for each $i = 1, 2$,

$$\lim_{\substack{\ell\to\infty \\ v\in V_{\widetilde{G}_{n_{k_\ell}}}}} (\widetilde{G}_{n_{k_\ell}}, o^i_{n_{k_\ell}}, X^{\widetilde{G}_{n_{k_\ell}},\widetilde{\xi}_{n_{k_\ell}}}) = (G, o^i, X^{G,\xi}) \text{ a.s.,}$$

so $(\widetilde{G}_n, o^i_n, X^{\widetilde{G}_n,\widetilde{\xi}^n}) \to (G, o^i, X^{G,\xi})$ in probability. Then by Proposition 2.26,

$$(\langle \mathscr{C}_{o^i_n}(G_n, X^{G_n,\xi^n}) \rangle)_{i=1,2} = (\langle (\widetilde{G}_n, \widetilde{o}^i_n, X^{\widetilde{G}_n,\widetilde{\xi}^n}) \rangle)_{i=1,2} \to (\langle G, o^i, X^{G,\xi} \rangle)_{i=1,2} = (\langle G^{(i)}, X^{G^{(i)},\xi^{(i)}} \rangle)_{i=1,2},$$

in probability in $\mathscr{G}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{D}]$ noting that the first equality follows from Remark 3.7 and the fact that $(\langle \mathscr{C}_{o^i_n}(G_n, X^{G_n,\xi^n}) \rangle)_{i=1,2}$ is well defined for each $n \in \mathbb{N}$ is a consequence of Proposition 4.1 and Remark 4.3. $\square$

Remark 5.12. Note that the proof of Lemma 5.11 does not make direct use of the form of (3.1.4) but only relies on spatial localization of the SDE, Assumption $3'$ and Proposition 5.10, all of which could potentially be verified for solutions $X^{G,\xi}$ of more general Poisson-driven SDEs.

Applying Lemmas 2.5 and 5.11, we now prove Theorem 5.6.

Proof of Theorem 5.6: Because (G_n, ξ^n) converges in probability in the local weak sense to $\langle G, \xi\rangle$ and $|G_n| \to \infty$ in probability, Lemma 2.5 implies that

$$(\langle \mathscr{C}_{o_n^1}(G_n, \xi^n)\rangle, \langle \mathscr{C}_{o_n^2}(G_n, \xi^n)\rangle) \Rightarrow (\langle G^{(1)}, \xi^{(1)}\rangle, \langle G^{(2)}, \xi^{(2)}\rangle),$$

where $\langle G^{(i)}, \xi^{(i)}\rangle, i = 1, 2$, are i.i.d. copies of $\langle G, \xi\rangle$. By the Skorokhod representation theorem and Lemma A.3, there exists a probability space supporting the random elements $\{\langle \widetilde{G}^{(i)}, \widetilde{\xi}^{(i)}\rangle\}_{i=1,2} \stackrel{(\mathrm{d})}{=} \{\langle G^{(i)}, \xi^{(i)}\rangle\}_{i=1,2}$, and finite, $\widehat{\mathscr{G}}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{X}]$-random elements $\{(\widetilde{G}_n, \widetilde{o}_n^i, \widetilde{\xi}^n)\}_{n\in\mathbb{N}, i=1,2} \stackrel{(\mathrm{d})}{=} \{(G_n, o_n^i, \xi^n)\}_{n\in\mathbb{N}, i=1,2}$ such that

$$\langle \mathscr{C}_{\widetilde{o}_n^i}(\widetilde{G}_n, \widetilde{\xi}^n)\rangle \to \langle \widetilde{G}^{(i)}, \widetilde{\xi}^{(i)}\rangle,\ i = 1, 2,\ \text{a.s..}$$

By assumption, $\widetilde{G}_n \stackrel{(\mathrm{d})}{=} G_n$ and $\widetilde{G}^{(i)} \stackrel{(\mathrm{d})}{=} G$ are a.s. finitely dissociable for $i = 1, 2$, and $n \in \mathbb{N}$. So by Assumption 1 and Proposition 4.22, $\{\mathscr{C}_{\widetilde{o}_n^i}(\widetilde{G}_n, \widetilde{\xi}^n), \mathscr{C}_{\widetilde{o}^{(i)}}(\widetilde{G}^{(i)}, \widetilde{\xi}^{(i)})\}_{n\in\mathbb{N}, i=1,2}$ consistently spatially localizes the SDE (3.1.4). Assumptions 1 and 3 imply by Lemma 5.9 that Assumption 3′ holds for each i. Thus, by Lemma 5.11, it is possible to construct a complete, filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and a collection of $\mathbb{F}$-driving noises $\{\mathbf{N}^{\widetilde{G}_n}, \mathbf{N}^{\widetilde{G}^{(i)}}\}_{n\in\mathbb{N}, i=1,2}$ such that for each $n \in \mathbb{N}$ and $i = 1, 2$, the respective $\mathbf{N}^{\widetilde{G}_n}$ and $\mathbf{N}^{\widetilde{G}^{(i)}}$-solutions $X^{\widetilde{G}_n, \widetilde{\xi}_n}$ and $X^{\widetilde{G}^{(i)}, \widetilde{\xi}^{(i)}}$ satisfy

$$\langle \mathscr{C}_{\widetilde{o}_i^n}(\widetilde{G}_n, X^{\widetilde{G}_n, \widetilde{\xi}^n})\rangle \to \langle \widetilde{G}^{(i)}, X^{\widetilde{G}^{(i)}, \widetilde{\xi}^{(i)}}\rangle \text{ in probability as } n \to \infty.$$

By the bounded convergence theorem, for any bounded, continuous function $f : (\mathscr{G}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{D}])^2 \to \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{E}\left[f\left(\left(\langle \mathscr{C}_{\widetilde{o}_i^n}(\widetilde{G}_n, X^{\widetilde{G}_n, \widetilde{\xi}^n})\rangle\right)_{i=1,2}\right)\right] = \mathbb{E}\left[f\left(\left(\langle \widetilde{G}^{(i)}, X^{\widetilde{G}^{(i)}, \widetilde{\xi}^{(i)}}\rangle\right)_{i=1,2}\right)\right]. \tag{5.3.8}$$

By well-posedness of the SDE (3.1.4) on each of the graphs $(\widetilde{G}_n, \widetilde{\xi}^n), n \in \mathbb{N}$, and $(\widetilde{G}, \widetilde{\xi})$, which holds by Theorem 4.2, it follows that for every $n \in \mathbb{N}$, $\left(\langle \mathscr{C}_{\widetilde{o}_i^n}(\widetilde{G}_n, X^{\widetilde{G}_n, \widetilde{\xi}^n})\rangle\right)_{i=1,2} \stackrel{(\mathrm{d})}{=} \left(\langle \mathscr{C}_{o_n^i}(G_n, X^{G_n, \xi^n})\rangle\right)_{i=1,2}$ and $(\widetilde{G}^{(1)}, X^{\widetilde{G}^{(1)}, \widetilde{\xi}^{(1)}}) \stackrel{(\mathrm{d})}{=} (\widetilde{G}^{(2)}, X^{\widetilde{G}^{(2)}, \widetilde{\xi}^{(2)}}) \stackrel{(\mathrm{d})}{=} (G, X^{G, \xi})$. Together with

(5.3.8) and the independence of $\langle \widetilde{G}^{(1)}, X^{\widetilde{G}^{(1)},\widetilde{\xi}^{(1)}} \rangle$ and $\langle \widetilde{G}^{(2)}, X^{\widetilde{G}^{(2)},\widetilde{\xi}^{(2)}} \rangle$, for any bounded, continuous functions $f_1, f_2 : \mathscr{G}_*[\overline{\mathcal{K}}, \mathcal{K} \times \mathcal{D}] \to \mathbb{R}$, it follows that

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}\left[f_1(\langle \mathscr{C}_{o_n^1}(G_n, X^{G_n,\xi^n})\rangle) f_2(\langle \mathscr{C}_{o_n^2}(G_n, X^{G_n,\xi^n})\rangle)\right] \\
&\quad = \lim_{n\to\infty} \mathbb{E}\left[f_1(\langle \mathscr{C}_{\widetilde{o}_1^n}(\widetilde{G}_n, X^{\widetilde{G}_n,\widetilde{\xi}^n})\rangle) f_2(\langle \mathscr{C}_{\widetilde{o}_2^n}(\widetilde{G}_n, X^{\widetilde{G}_n,\widetilde{\xi}^n})\rangle)\right] \\
&\quad = \mathbb{E}\left[f_1(\langle \widetilde{G}^{(1)}, X^{\widetilde{G}^{(1)},\widetilde{\xi}^{(1)}}\rangle) f_2(\langle \widetilde{G}^{(2)}, X^{\widetilde{G}^{(2)},\widetilde{\xi}^{(2)}}\rangle)\right] \\
&\quad = \mathbb{E}\left[f_1(\langle \widetilde{G}^{(1)}, X^{\widetilde{G}^{(1)},\widetilde{\xi}^{(1)}}\rangle)\right] \mathbb{E}\left[f_2(\langle \widetilde{G}^{(2)}, X^{\widetilde{G}^{(2)},\widetilde{\xi}^{(2)}}\rangle)\right] \\
&\quad = \mathbb{E}\left[f_1(\langle G, X^{G,\xi}\rangle)\right] \mathbb{E}\left[f_2(\langle G, X^{G,\xi}\rangle)\right].
\end{aligned}$$

This allows us to conclude the desired result:

$$\lim_{n\to\infty} \mathrm{Cov}\left(f_1(\langle \mathscr{C}_{o_n^1}(G_n, X^{G_n,\xi^n})\rangle), f_2(\langle \mathscr{C}_{o_n^2}(G_n, X^{G_n,\xi^n})\rangle)\right) = 0.$$

$\square$

Remark 5.13. [36, Remark 7.4] From the proof of Theorem 5.6, it is not hard to see that in fact the conclusion of Theorem 5.6 holds under Assumption 3′ and the spatial localization property in place of Assumptions 1 and 3 and the finite dissociability property respectively.

Part III

The Local Equation

CHAPTER 6

The Local Equation

In this chapter we state the local equation and several results pertaining to the local equation. The proofs of all results in this chapter save for Lemma 6.3 are deferred to Chapters 9 to 8. More specifically, results in Section 6.2.1 are proved in Appendix C.2, Section 6.2.2 results are proved in Section 8, Section 6.3 results are proved in Chapter 9, results in Sections 6.4 and 6.5 are proved in Section 10.1 and finally the results in Section 6.6 are proven in Section 10.2.

In Part III we study IPS on the infinite d-regular tree $\mathbb{T}_d := (V^d, E^d, \varnothing^d)$ for $d \geq 2$. For any $\varnothing \in U \subseteq V^d$, $\mathbb{T}_d[U]$ is assumed to be the *rooted* subtree of $\mathbb{T}_d$ induced by U. For each $n \in \mathbb{N}$, let $V^d_n := \{v \in V^d : d_{\mathbb{T}_d}(\varnothing, v) \leq n\}$, $n \in \mathbb{N}$.

6.1 New Assumptions and Notation

Throughout Part III, we work exclusively with IPS on deterministic (possibly unrooted) graphs G with trivial edge marks. For this reason, we make the following minor simplifications to the notation:

1. we separate G from its vertex marks regarding G as an unmarked graph describing the topology of the IPS while we regard the initial data to be a random vector κ^G in $\mathcal{K}^{V_G}$;

2. we assume there exists a measurable mapping $\varsigma : \mathcal{K} \to \mathcal{X}$ such that the initial state of the

IPS ξ^G is a.s. given by $\xi^G_v = \varsigma(\kappa^G_v)$ for all $v \in V_G$;

3. we remove the initial state from the superscript of the IPS, that is, we denote the IPS by X^G, a random vector in the space $\mathcal{D}^{V_G}$;

4. we no longer consider jump rate functions to be indexed by the graph G, so we instead denote the rate functions by a family of functions $r^v_j : \mathbb{R}_+ \times (\mathcal{K} \times \mathcal{D})^{V_G}$, $v \in V_G$, $j \in \mathcal{J}$.

For ease, we also compile all of the major assumptions we hold throughout part III here. Some of these assumptions are new, but most are assumptions from earlier portions of the paper restated for our convenience.

Assumption 4. The unmarked (possibly rooted) graph $G = (V, E, \varnothing)$ or $G = (V, E)$ and family of Borel-measurable rate functions $\{r^v_j\}_{v \in V, j \in \mathcal{J}}$ from $\mathbb{R}_+ \times (\mathcal{K} \times \mathcal{D})^V$ to $\mathbb{R}_+$, satisfy the following five conditions:

(1) (locality) for each $v \in V$ and $j \in \mathcal{J}$, there exists a measurable function $\bar{r}^v_j : \mathbb{R}_+ \times (\mathcal{K} \times \mathcal{D})^{\mathrm{cl}_v} \to \mathbb{R}_+$ such that for any $t \in \mathbb{R}_+$ and $(\vartheta, x) \in (\mathcal{K} \times \mathcal{D})^V$,

$$r^v_j(t, \vartheta, x) = \bar{r}^v_j(t, \vartheta_{\mathrm{cl}_v}, x_{\mathrm{cl}_v});$$

(2) (predictability) for every $\vartheta \in \mathcal{K}^V$, $(t, x) \mapsto r^v_j(t, \vartheta, x)$ is predictable in the sense that for every $t > 0$, $v \in V$ and $x, y \in \mathcal{D}^V$,

$$x(s) = y(s) : \quad s \in [0, t) \quad \Rightarrow \quad r^v_j(t, \vartheta, x) = r^v_j(t, \vartheta, y);$$

(3) (boundedness) there exists a non-decreasing function $C : \mathbb{R}_+ \to \mathbb{R}_+$ such that for any $v \in V$, $j \in \mathcal{J}$ and $t \in \mathbb{R}_+$, $r^v_j(t, \cdot, \cdot) \leq C(t)$;

(4) (regularity) for every $j \in \mathcal{J}$, $v \in V$ and $(\vartheta, x) \in (\mathcal{K} \times \mathcal{D})^V$, the function $t \mapsto r^v_j(t, \vartheta, x)$ is càglàd (left continuous with right limits).

(5) (symmetry) for each $j \in \mathcal{J}$, $t \in \mathbb{R}_+$, $(\vartheta, x) \in (\mathcal{K} \times \mathcal{D})^V$, $u, v \in V$ and any unrooted automorphism $\varphi : V \to V$ such that $\varphi(v) = u$,

$$r_j^v(t,\vartheta,x) = r_j^u(t,\varphi_*(\vartheta,x));$$

Remark 6.1. Note that condition (1) of Assumption 4 is simply condition (2) of the standing assumption. Condition (2) of Assumption 4 is condition (1) of the standing assumption. Because G is a deterministic graph, the Borel measurability of each r_j^v implies condition (3) of the standing assumption. Condition (3) of Assumption 4 is simply Assumption 1 under the condition that the deterministic graph G is of bounded maximal degree. Condition (5) of Assumption 4 is implied by Assumption 2, though Assumption 2 is more stringent as it pertains to isomorphisms between different graphs. Finally, condition (4) of Assumption 4 is a new assumption. It ensures that for any càdlàg, $\mathbb{F}$-adapted $\mathcal{D}^{V_G}$-random element Z, any $\mathcal{F}_0$-measurable $\mathcal{X}^{V_G}$-random element κ^Z and any $(j,v) \in \mathbb{R}_+ \times \mathcal{J} \times V_G$, the non-negative process given by $t \mapsto r_j^v(t,Z,\kappa^Z)$ is $\mathbb{F}$-predictable. This condition is quite mild and it is satisfied by every example in Section 3.2 under the extra condition that h_I and h_R in Example 3.18 are càdlàg functions.

Remark 6.2. Recall that $V_1^d = \mathrm{cl}_\varnothing(\mathbb{T}_d)$. Under Assumption 4, if $G = \mathbb{T}_d$, then for each $j \in \mathcal{J}$, there exists a measurable function $r_j : \mathbb{R}_+ \times (\mathcal{K} \times \mathcal{D})^{V_1^d} \to \mathbb{R}_+$ so that for any unrooted automorphisms φ and $\varphi' : V^d \to V^d$ such that $\varphi^{-1}(\varnothing) = (\varphi')^{-1}(\varnothing)$ and any $(t,\vartheta,x) \in \mathbb{R}_+ \times (\mathcal{K} \times \mathcal{D})^{V^d}$,

$$r_j^{\varphi^{-1}(\varnothing)}(t,\vartheta,x) = r_j(t,\varphi_*(\vartheta_{V_1^d},x_{V_1^d})) = r_j(t,\varphi'_*(\vartheta_{V_1^d},x_{V_1^d})).$$

With our modified notation, (3.1.4) becomes,

$$X_v^G(t) = \varsigma(\kappa_v^G) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\{r \le r_j^v(s,\kappa^G,X^G)\}} \mathbf{N}_v(ds,dr,dj), \quad v \in V, t \in [0,\infty). \tag{6.1.1}$$

For the remainder of Part III, we fix a filtered probability space $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ and driving Poisson processes $\mathbf{N} := (\mathbf{N}_v)_{v\in V^d}$. We denote by X^G the corresponding weak solution to (6.1.1) for the initial data κ^G driven by $\mathbf{N}$ in $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$. By Lemma 6.3, this is also a strong solution so we simply call X^G *the solution* to (6.1.1) (for the initial data κ^G). We further define $\mu^G := \mathrm{Law}(\kappa^G, X^G)$ (note that by Lemma 6.3, μ^G is uniquely characterized by the graph G and the jump rate functions $\{r_j^v\}$). For any $t \in \mathbb{R}_+$, we write $\mu_t^G := \mathrm{Law}(\kappa^G, X^G[t])$ and for $t > 0$, $\mu_{t-}^G := \mathrm{Law}(\kappa^G, X^G[t))$. When $G = \mathbb{T}_d$, the superscript is omitted.

Recall the definition of $\mathrm{Disc}_t\,(x)$ from Section 1.4. We make use of the following fundamental result.

Lemma 6.3. *Any weak solution X^G to* (6.1.1) *is a.s. proper in the sense that for each $u, v \in V$ and $t \in \mathbb{R}_+$,* $\mathrm{Disc}_t\left(X_u^G\right) \cap \mathrm{Disc}_t\left(X_v^G\right) = \emptyset$ *a.s.. Suppose conditions (1),(2) and (3) of Assumption 4 hold. If G is a finitely dissociable graph, then the SDE* (6.1.1) *is strongly well-posed. Finally, if condition (5) of Assumption 4 also holds and κ^G is automorphism invariant, then so is (κ^G, X^G).*

Proof. To show that a weak solution X^G is a.s. proper, note that for any two vertices $u, v \in V$ and $T < \infty$, $\mathrm{Disc}_T\left(X_u^G\right) \cap \mathrm{Disc}_T\left(X_v^G\right) \neq \emptyset$ only if there exists an $s \in [0, T]$ such that $\mathbf{N}_u(\{s\} \times \mathbb{R}_+ \times \mathcal{J})$ and $\mathbf{N}_v(\{s\} \times \mathbb{R}_+ \times \mathcal{J})$ are both greater than 0, but this only happens with probability 0 by the independence of $\mathbf{N}_u$ and $\mathbf{N}_v$ and the fact that both Poisson processes have non-atomic intensity measures on $\mathbb{R}_+^2 \times \mathcal{J}$. Thus, X^G is a.s. proper.

Because G is a subgraph of $\mathbb{T}_d$, it is a bounded degree graph and therefore finitely dissociable by Proposition 4.20. By Remark 6.1, conditions (1)-(3) of Assumption 4 imply that the standing assumption and Assumption 1 hold. So by Theorem 4.2, (6.1.1) is strongly well-posed.

Lastly, suppose now that condition (5) of Assumption 4 also holds and let $\varphi : V \to V$ be any automorphism of G. Suppose also that the initial data is automorphism invariant in the sense that $\kappa' := \varphi_* \kappa \overset{(\mathrm{d})}{=} \kappa$. Then $(\kappa', \mathbf{N}') := \varphi_*(\kappa, \mathbf{N}) \overset{(\mathrm{d})}{=} (\kappa, \mathbf{N})$. Thus, condition (5) of Assumption 4 implies that

$$\begin{aligned}
\varphi_* X_v^G(t) &= X_{\varphi^{-1}(v)}^G(t) \\
&= \varsigma(\kappa_{\varphi^{-1}(v)}^G) + \int_{(0,t] \times \mathbb{R}_+ \times \mathcal{J}} j \mathbb{I}_{\left\{r \leq r_j^{\varphi^{-1}(v)}(t, \kappa^G, X^G)\right\}} \mathbf{N}_{\varphi^{-1}(v)}(ds, dr, dj) \\
&= \varphi_* \varsigma(\kappa_v^G) + \int_{(0,t] \times \mathbb{R}_+ \times \mathcal{J}} j \mathbb{I}_{\left\{r \leq r_j^v(t, \varphi_*(\kappa^G, X^G))\right\}} \varphi_* \mathbf{N}_v(ds, dr, dj),
\end{aligned}$$

so by well-posedness of (6.1.1), $\varphi_*(\kappa^G, X^G) \overset{(\mathrm{d})}{=} (\kappa^G, X^G)$ so the solution to (6.1.1) is automorphism invariant. $\square$

6.2 Semi-Global Markov Random Fields and Tiling

6.2.1 Markov Random Field Definitions

Markov random fields (MRF) generalize the concept of a Markov chains to graphical models, or random vectors indexed by the nodes of a graph. We begin with a standard definition of an MRF for finite graphs.

Definition 6.4. If $G = (V, E)$ is a finite graph and $\mathcal{Z}$ is a Polish space, $\eta \in \mathcal{P}(\mathcal{Z}^V)$ is said to be an MRF on $\mathcal{Z}^V$ with respect to G if for any partition A, B, S of V such that $S = \mathcal{N}_A(G)$ and any random element Z with $\mathrm{Law}(Z) = \eta$, Z_A is conditionally independent of Z_B given Z_S, denoted

$$Z_A \perp\!\!\!\perp Z_B | Z_S. \tag{6.2.1}$$

On infinite graphs, there are several different versions of the MRF property.

Definition 6.5. [37] Fix a locally finite graph $G = (V, E)$, a Polish space $\mathcal{Z}$ and $\eta \in \mathcal{P}(\mathcal{Z}^V)$. Then η is said to be a *global* MRF if for any Z with $\mathrm{Law}(Z) = \eta$ and any partition A, B, S of V such that $S = \mathcal{N}_A(G)$, (6.2.1) holds. Moreover, η is said to be a *local* MRF if (6.2.1) holds under the condition that A *is finite*. Finally, we define a semi-global MRF (abbreviated SGMRF) to be a measure η for which (6.2.1) holds under the weaker condition that S *is finite*.

Remark 6.6. [37] Note that many sources that focus on graphical models on finite graphs assign different meanings to local and global MRFs [63]. Our use of the term is more in keeping with the terminology of texts focused on infinite graphs such as [43] and [41].

The derivation of the local equation involves making use of the fact that under certain conditions, the trajectories of an IPS are *second-order* SGMRFs (2-SGMRFs), where higher order SGMRFs are defined below.

Definition 6.7. [37] Fix $\alpha \in \mathbb{N}$ and $\eta \in \mathcal{P}(\mathcal{Z}^V)$. Then η is said to be a *global* α-MRF if for any Z with $\mathrm{Law}(Z) = \eta$ and any partition A, B, S of V such that $S = \mathcal{N}_A^\alpha(G)$, (6.2.1) holds. Moreover, η is said to be a *local* α-MRF if (6.2.1) holds under the condition that A *is finite*. Finally, we define an α-SGMRF to be a measure η for which (6.2.1) holds under the weaker condition that S *is finite*.

In [60, Remark 3.16] it is noted that the local equations hold generally require automorphism-invariance and a *global* 2-MRF property, however we show in this paper that the *semi-global* property defined below is sufficient.

There is a second useful definition of α-SGMRF which we state below. We prove that these definitions are, in fact, equivalent in Appendix C.2.

Lemma 6.8. *[37] A measure $\eta \in \mathcal{P}(\mathcal{Z}^V)$ is an α-SGMRF if and only if for every random element Z such that $Law(Z) = \eta$,* (6.2.1) *holds for all finite, disjoint $A, B, S \subseteq V$ such that S α-separates A and B.*

Remark 6.9. [37] When G is a tree, it is shown in Appendix C.2 that η is an α-SGMRF if and only if for any connected, finite $U \subseteq V$, $\eta[U]$ is an α-MRF with respect to the graph $G[U]$. Thus, when G is a tree, our definition of an α-SGMRF extends the definition of a "Markov chain on a tree," as stated in [92, Section 2] and [41, Chapter 12], to order $\alpha \geq 1$ Markov chains in the obvious way.

Clearly, all global MRFs are SGMRFs, which in turn (since G is locally finite) are also local MRFs, and all three concepts coincide on finite graphs. Furthermore, the SGMRF property is strictly stronger than the local MRF property (see [41, Corollary 11.33]) and strictly weaker than the global MRF property (see [43, Section 2]); also see [41, Example 8.24] as well as part B of the bibliographical notes for Section 8.2 of [41] for further discussion of local and global MRFs.

6.2.2 Automorphism Invariant 2-SGMRFs

As briefly mentioned in the previous section, the reason we are interested in 2-SGMRFs is because the trajectories of an IPS are typically automorphism invariant 2-SGMRFs as we show below and prove in Section 8.1.

Proposition 6.10. *Suppose Assumption 4 holds. If $\mu_\kappa \in \mathcal{P}(\mathcal{K}^{V^d})$ is an automorphism invariant, 2-SGMRF and μ is the unique law of a weak solution to* (6.1.1) *for initial data with law μ_κ, then μ, μ_{t-} and μ_t, $t > 0$ are automorphism invariant 2-SGMRFs.*

Furthermore, it turns out that not only IPS, but *all* automorphism invariant 2-SGMRF graphical models defined on the graph $\mathbb{T}_d$ can be characterized by their V_1^d-marginals. Specifically, every automorphism invariant 2-SGMRF has a V_1^d-marginal satisfying the following strong symmetry property.

Definition 6.11. A measure $\eta \in \mathcal{P}(\mathcal{Z}^{V_1^d})$ is said to be *strongly symmetric* if and only if η is automorphism invariant with respect to $\mathbb{T}_d[V_1^d]$ and for any $Z \sim \eta$, $Z_{\{\varnothing,1\}} \overset{(\mathrm{d})}{=} Z_{\{1,\varnothing\}}$.

Proposition 6.12. *If $\overline{\eta} \in \mathcal{P}(\mathcal{Z}^{V^d})$ is automorphism invariant and 2-SGMRF, then $\overline{\eta}[V_1^d]$ is strongly symmetric. Conversely, if $\eta \in \mathcal{P}(\mathcal{Z}^{V_1^d})$ is strongly symmetric, then there exists a unique automorphism invariant, 2-SGMRF $\overline{\eta} \in \mathcal{P}(\mathcal{Z}^{V^d})$ such that $\overline{\eta}[V_1^d] = \eta$.*

Definition 6.13. Given η in Proposition 6.12, we call $\overline{\eta}$ the *tiled extension* of η.

The proof of Proposition 6.12 is deferred to Section 8.3. The proof is constructive, and the explicit construction of the tiled extension of any strongly symmetric measure is given in Section 8.3.2.

6.3 The Local Equation

6.3.1 Motivation

We begin with our main result, which is the derivation of a local equation in the style of the local equation of [60, Definition 3.5] for regular trees. Given an IPS X that solves (6.1.1) for the graph $G = \mathbb{T}_d$, the local equation is a law-dependent, history dependent SDE that characterizes the marginal law of $X_{V_1^d}$, where recall that $V_1^d = \mathrm{cl}_\varnothing(\mathbb{T}_d)$. For convenience, we introduce the labeling $V_1^d := \{\varnothing, 1, \dots, d\}$. The basic idea is to first use a filtering argument, then to exploit the fact that (κ, X) forms an automorphism invariant 2-SGMRF to derive a collection of measurable functions $\gamma_j : \mathbb{R}_+ \times (\mathcal{K} \times \mathcal{D})^{\{\varnothing,1\}} \to \mathbb{R}_+$ such that the solution Y to the equation,

$$(6.3.1) \qquad Y_v(t) = \begin{cases} \varsigma(\kappa_v) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j \mathbb{I}_{\{r \le \gamma_j(s, \kappa_{\{v,\varnothing\}}, Y_{\{v,\varnothing\}})\}} \mathbf{N}_v(ds, dr, dj) & \text{if } v \in \mathcal{N}_\varnothing, \\ \varsigma(\kappa_\varnothing) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j \mathbb{I}_{\left\{r \le r_j(s, \kappa_{V_1^d}, Y)\right\}} \mathbf{N}_\varnothing(ds, dr, dj) & \text{if v} = \emptyset, \end{cases}$$

satisfies $(\kappa_{V_1^d}, Y) \overset{(\mathrm{d})}{=} (\kappa_{V_1^d}, X_{V_1^d})$, where recall the definition of $\{r_j\}_{j\in\mathcal{J}}$ from Remark 6.2. Note that (6.3.1) is just a restatement of (6.1.1) on the graph $G = \mathbb{T}_d[V_1^d]$ where for $v \in \mathcal{N}_\varnothing$, r_j^v is replaced by γ_j. As one might expect, for Y and $X_{V_1^d}$ to have the same distribution, $\gamma_j(t, \vartheta, x)$

is, formally speaking, obtained by averaging $r_j^1(t,\kappa,X)$ conditioned on the (possibly null event) $\{(\kappa_{V_1^d}, X_{V_1^d}) = (\vartheta_{V_1^d}, x_{V_1^d})\}$. Indeed, this can be established rigorously by applying standard filtering results from point process theory. By utilizing the 2-SGMRF structure of (κ, X) established in Proposition 6.10 and the automorphism-invariance of (κ, X) it is possible to derive an autonomous representation of the functions γ_j, $j \in \mathcal{J}$, using arguments which are broadly outlined in the introduction. This also explains the dependence of γ_j on $(\kappa_{\{v,\varnothing\}}, Y_{\{v,\varnothing\}})$ instead of $(\kappa_{V_1^d}, Y_{V_1^d})$ in (6.3.1). It should be noted that these arguments are broadly similar to the arguments applied in the proof of [60, Theorem 3.12]. However, our approach includes several new contributions which we discuss in Section 6.3.4.

6.3.2 Definition of the Local Equation

Recall that the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and $\mathbb{F}$-driving Poisson processes $\mathbf{N} := (\mathbf{N}_v)_{v\in V^d}$ are assumed to be fixed. Further define the $\mathcal{F}_0$-measurable, $\mathcal{K}^{V_1^d}$-random vector $\widehat{\kappa}$. Then we define a (strong) solution to the local equation for the initial data $\widehat{\kappa}$ as follows.

Definition 6.14. A strong solution to the local equation for the initial data $\widehat{\kappa}$ is a tuple (Y, γ) such that,

(1) Y is a $\mathcal{D}^{V_1^d}$-valued, $\mathcal{H}^{\widehat{\kappa}} \vee \mathbb{H}^{\mathbf{N}} \subseteq \mathbb{F}$-adapted process;

(2) $\gamma := \{\gamma_j\}_{j\in\mathcal{J}}$ is a family of Borel measurable functions $\gamma_j : \mathbb{R}_+ \times (\mathcal{K}\times\mathcal{D})^2 \to \mathbb{R}_+$, $j \in \mathcal{J}$, satisfying conditions (2), (3) and (4) of Assumption 4 such that,

$$\gamma_j(t, \widehat{\kappa}_{\{\varnothing,1\}}, Y_{\{\varnothing,1\}}) = \mathbb{E}\left[r_j(t,\widehat{\kappa}, Y) \middle| \widehat{\kappa}_{\{\varnothing,1\}}, Y_{\{\varnothing,1\}}[t] \right] \text{ a.s.,} \tag{6.3.2}$$

for every $t \in \mathbb{R}_+$;

(3) $(\widehat{\kappa}_v, Y_v, \mathbf{N}_v)_{v\in V_1^d}$ and γ satisfy (6.3.1).

We call Y a strong solution to the local equation whenever there exists a collection of functions γ for which (Y, γ) is a strong solution to the local equation. We denote the law of the strong solution Y to the local equation by ν. The local equation is said to be *unique in law* if for any two weak solutions, Y and Y', to the local equation, $Y \stackrel{(d)}{=} Y'$, and it is said to be *pathwise unique* if for any two strong solutions (Y,γ) and (Y',γ'), to the local equation, $Y = Y'$ a.s..

Remark 6.15. It may be strange that we stipulated that $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ support $\mathbf{N} := \{\mathbf{N}_v\}_{v\in V^d}$ before Definition 6.14. To understand why we may assume this without loss of generality, note that by Definition 6.14, any strong solution to the local equation given the filtered probability space $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ supporting the $\mathbb{F}$-driving Poisson processes $\widehat{\mathbf{N}} := (\mathbf{N}_v)_{v\in V_1^d}$ is also a strong solution to the local equation with respect to any extended probability space $(\Omega',\mathcal{F}',\mathbb{F}',\mathbb{P}')$ for which $\widehat{\mathbf{N}}$ are $\mathbb{F}'$-Poisson processes. Setting $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P}) = (\Omega',\mathcal{F}',\mathbb{F}',\mathbb{P}')$ we may assume without loss of generality that $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ supports the $\mathbb{F}$-Poisson processes $\mathbf{N}$.

Remark 6.16. We choose to work with strong solutions to the local equation because, for standard definitions of weak solution, all weak solutions to the local equation are strong when $\{r_j^v\}_{v\in V^d, j\in\mathcal{J}} := \{r_j\}_{v\in V^d, j\in\mathcal{J}}$ satisfies Assumption 4 with respect to the graph $\mathbb{T}_d$. A weak solution to the local equation can be defined in the standard way by including the filtered probability space $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$, the driving Poisson processes $\mathbf{N}$ and the initial data $\widehat{\kappa}$ as part of the solution and by only requiring Y to be $\mathbb{F}$-adapted instead of $\mathcal{H}^{\widehat{\kappa}} \vee \mathbb{H}^{\mathbf{N}}$-adapted. Indeed, [60] works specifically with weak solutions to the local equation for diffusions in this sense. However, it can be noted that for pure jump processes, it is not necessary to define solutions in this way. To see why, note that for a fixed set of functions γ, (6.3.1) is equivalent to (6.1.1) where $G = \mathbb{T}_d[V_1^d]$ and with respect to the jump rate functions,

$$
r_j^{\prime,v}(s,\widehat{\vartheta},y) := \begin{cases} r_j(s,\widehat{\vartheta},y) & \text{if } v = \varnothing, \\ \gamma_j(s,\widehat{\vartheta}_{\{v,\varnothing\}}, y_{\{v,\varnothing\}}) & \text{otherwise.} \end{cases} \tag{6.3.3}
$$

If γ satisfies condition (2) of Definition 6.14, then it is easily verified that the jump rate functions $\{r_j^{\prime,v}\}$ defined above satisfy Assumption 4 with respect to G.

Furthermore, γ are deterministic functions and so do not depend upon the choice of filtered probability space $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ and $\mathbf{N}$. Then by Lemmas 6.3 and 3.11, given any collection of driving Poisson processes $\mathbf{N}$ and any filtered probability space $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$, there exists a pathwise unique weak solution Y to (6.3.1) in the sense of Definition 3.5, and Y is $\mathcal{H}^{\widehat{\kappa}} \vee \mathbb{H}^{\mathbf{N}}$-adapted. Therefore, for any weak solution $((\Omega,\mathcal{F},\mathbb{F},\mathbb{P}),\gamma,(\widehat{\kappa}_v, Y_v, \mathbf{N}_v)_{v\in V_1^d})$ to the local equation, (Y,γ) is a strong solution to the local equation given the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, the driving Poisson processes $\mathbf{N}$ and for the initial data $\widehat{\kappa}$.

As argued in Remark 6.16, (6.3.1) every weak solution to the local equation is also strong. Therefore we will simply refer to Y as a solution to the local equation if there exist functions γ such that (Y, γ) is a strong solution to the local equation. Note that because the functions $\{\gamma_s^j\}_{j \in \mathcal{J}, s \in \mathbb{R}_+}$ depend on $\nu := \mathrm{Law}(\widehat{\kappa}, Y)$, the local equation is essentially a fixed point equation in the following sense. Given a law ν, one can use (6.3.2) to define the functions γ. Then the solution Y' to (6.3.1) given the filtered rate functions γ will have law ν'. Then Y is a solution to the local equation if and only if $\nu = \nu'$.

6.3.3 Well-Posedness of the Local Equation

We now introduce our main result, which shows that for any strongly symmetric initial data, the local equation has a pathwise unique strong solution, the solution is a marginal of a corresponding IPS and the IPS itself is the tiled extension of the solution to the local equation. We prove this theorem in Chapter 9.

Theorem 6.17. *Suppose Assumption 4 holds. Let $\nu_\kappa \in \mathcal{P}(\mathcal{K}^{V_1^d})$ be a strongly symmetric probability measure. Let μ be the law of the solution to* (6.1.1) *for the initial data $\kappa \sim \mu_\kappa := \overline{\nu}_\kappa$, where $\overline{\nu}_\kappa \in \mathcal{P}(\mathcal{K}^{V^d})$ is the tiled extension of ν_κ. Then there exists a pathwise unique strong solution to the local equation with law $\nu := \mu[V_1^d]$, and μ is the tiled extension of ν. Moreover, for every $t > 0$, μ_{t-} is the tiled extension of ν_{t-}.*

6.3.4 Discussion

The arguments outlined earlier in this section describing the proof of Theorem 6.17 above are, at a high level, similar to the approach developed in [60]. However, there are four major distinctions between the Theorem 6.17 and the main result of [60] (Theorem 3.12) in the regular tree case described in [60, Section 3.2.1]. First, our assumptions are minimal. We require bounded jump rate functions that are càdlàg in time and sufficiently predictable, but we do not require the jump rate functions to satisfy any Lipschitz bounds or uniform continuity assumptions.

Second, we work with the more general initial data both in the sense that the initial data encompasses more information than simply the value of the process at time 0, and in the sense that our results hold for all strongly symmetric initial data. This extends the results of [60] which only proved the well-posedness of the local equation (for diffusions) for i.i.d. initial conditions representing only the state of the process at time 0. In fact, our work goes beyond the extension discussed in [60, Remark 3.16], as we merely require the initial data of the local equation to be strongly symmetric, whereas [60, Remark 3.16] details extending the local equations to marginals of automorphism invariant global 2-MRFs. It is very easy to verify whether or not a measure $\nu_\kappa \in \mathcal{P}(\mathcal{K}^{V_1^d})$ is strongly symmetric, but not so simple to verify whether or not it is the marginal of an automorphism invariant global 2-MRF. Furthermore, as a consequence of Proposition 6.12, the class of strongly symmetric initial data is precisely the class of marginals of automorphism invariant 2-SGMRFs which is strictly larger than the class of initial data described in [60, Remark 3.16]. Our exploration of the local equation for more general initial data allows us to begin a rigorous study of flow properties and stationarity of the local equation. We explore this in greater depth in Sections 6.4 and 6.5.

The third major difference is that we establish (under suitable conditions) the well-posedness of the local equation is strong in the sense that, by taking the filtered probability space, driving noise and initial data as given, there exists a unique solution to the local equation up to indistinguishability. This will prove useful for numerical applications of the local equation.

Lastly, the local equation of [60] applies to interacting diffusions, while ours is applied to interacting particle systems. While the key ideas of the proof are analogous to the proof of [60, Theorem 3.12], the specific arguments required to establish each intermediate result is more subtle. For example, because we are working with stochastic processes with càdlàg trajectories rather than continuous trajectories, our approach requires that solutions to the local equation impose stronger the regularity conditions on the set of functions $\{\gamma_j\}$ (see condition (2) of Definition 6.14) than those imposed on the analogous function γ in [60, Definition 3.5]. The proof that Assumption 4 is sufficient for the solution of the local equation to satisfy this regularity condition requires a careful, càglàd extension argument showing that the conditional expectation in (6.3.2) can be extended to a càglàd process even on events of measure 0 (see Appendix B).

6.4 Local Flow

Theorem 6.17 provides a useful characterization of marginals of transient IPS. Here we discuss flow properties of the local equation and how they relate to flow properties of solutions to (6.1.1). Recall that $\overline{\mathcal{D}}$ is the space of $\mathcal{X}$-valued, càdlàg functions with domain $\mathbb{R}$. For each $t \in \mathbb{R}$, define the shift operator $\theta_t : \overline{\mathcal{D}} \to \overline{\mathcal{D}}$ by,

$$(\theta_t(z))(s) = z(t+s) \text{ for all } z \in \overline{\mathcal{D}}, s \in \mathbb{R}.$$

We will abuse notation slightly by letting $\{\theta_t\}_{t\in\mathbb{R}_+}$ denote the shift operator from $\overline{\mathcal{D}}^U \to \overline{\mathcal{D}}^U$ for any $U \subseteq V^d$ and by letting it denote the pushforward operator. That is, for any $U \subseteq V^d$ and $\eta \in \mathcal{P}(\overline{\mathcal{D}}^U)$, if $Z \sim \eta$, then $\theta_t Z \sim \theta_t \eta$ for all $t \in \mathbb{R}$. In this context, we say that for any $U \subseteq V^d$, a measure $\eta \in \mathcal{P}(\overline{\mathcal{D}}^U)$ is *stationary* if $\eta = \theta_t \eta$ for all $t \in \mathbb{R}$. Then we can define stationary solutions as long as the jump rates are time-homogeneous in the following sense, where recall that $\overline{\mathcal{D}}_0 := \mathcal{D}((-\infty, 0], \mathcal{X})$:

Assumption 5. Let $\mathcal{K} := \overline{\mathcal{D}}_0$. For each $v \in V^d, j \in \mathcal{J}$, there exists a measurable function $\breve{r}^v_j : \overline{\mathcal{D}}_{0-}^{V^d} \to \mathbb{R}_+$ such that for each $(t,x) \in \mathbb{R}_+ \times \overline{\mathcal{D}}^{V^d}$,

$$r^v_j(t, x[0,\infty), x(-\infty,0]) = \breve{r}^v_j\left((\theta_t x)(-\infty,0)\right).$$

In this context, it becomes useful to identify the spaces $\overline{\mathcal{D}}_0 \times \mathcal{D}$ with $\overline{\mathcal{D}}$. For any $\vartheta \in \overline{\mathcal{D}}_0$ and $x \in \mathcal{D}$, we will denote by $[\vartheta, x] \in \overline{\mathcal{D}}$ by the function,

$$[\vartheta, x](t) := \begin{cases} \vartheta(t) & \text{if } t < 0, \\ x(t) & \text{if } t \geq 0. \end{cases}$$

If ϑ and x are in $\overline{\mathcal{D}}_0^U$ and $\mathcal{D}^U$ respectively for some countable index set U, then we will define $[\vartheta, x] \in \overline{\mathcal{D}}^U$ analogously. For any $\eta = \text{Law}(\kappa_U, X_U) \in \mathcal{P}((\overline{\mathcal{D}}_0 \times \mathcal{D})^U)$, we will likewise define $[\eta] := \text{Law}([\kappa, X]) \in \mathcal{P}(\overline{\mathcal{D}}^U)$. We will say $\eta \in \mathcal{P}(\overline{\mathcal{D}}_0^U \times \mathcal{D}^U)$ is *stationary* if and only if $[\eta]$ is stationary. Lastly, given $\eta \in \mathcal{P}(\overline{\mathcal{D}}^U)$, let $\eta_- := \eta|_{\overline{\mathcal{D}}_0^U}$ and $\eta_+ := \eta|_{\mathcal{D}^U}$ be the negative and positive time marginals of η respectively.

Definition 6.18. Suppose Assumptions 4 and 5 hold. Then we say that a measure $[\mu] \in \mathcal{P}(\overline{\mathcal{D}}^{V^d})$ is a *two-sided solution* to the SDE (6.1.1) if $[\mu]_+$ is the law of the strong solution to the SDE (6.1.1) for initial data with law $[\mu]_-$. Likewise, a measure $[\nu] \in \mathcal{P}(\overline{\mathcal{D}}^{V_1^d})$ is said to be a *two-sided solution* to the local equation if $[\nu]_+$ is the law of the strong solution to the local equation for the initial data with law $[\nu]_-$. If $[\mu]$ or $[\nu]$ is stationary, then we say the measure is a *stationary solution.*

Lastly, we define a continuous flow.

Definition 6.19. A *time-homogeneous, continuous* flow on a Polish space $\mathcal{Z}$ is a collection of Borel measurable maps $\Phi_{s,t} : \mathcal{Z} \to \mathcal{Z}$, $0 \leq s < t < \infty$ such that for any $z \in \mathcal{Z}$ and $s, t \in \mathbb{R}_+$,

1. $\Phi_{s,t}(z) = \Phi_{0,t-s}(z)$;
2. $\lim_{t \searrow 0} \Phi_{0,t}(z) = \Phi_{0,0}(z) = z$;
3. for any $0 \leq s_1 \leq s_2 \leq s_3 < \infty$, $\Phi_{s_1,s_3}(z) = \Phi_{s_2,s_3} \circ \Phi_{s_1,s_2}(z)$.

We can immediately define a correspondence between flows defined by the local equation and by the SDE (6.1.1).

Definition 6.20. For $0 \leq s \leq t < \infty$ and $\mu_\kappa \in \mathcal{P}(\overline{\mathcal{D}}_0^{V^d})$, let $\overline{\Psi}_{s,t}(\mu_\kappa) := (\theta_{t-s}[\mu])_-$, where $[\mu]$ is the unique two-sided solution to (6.1.1) such that $[\mu]_- = \mu_\kappa$. Likewise, for any strongly symmetric $\nu_\kappa \in \mathcal{P}(\overline{\mathcal{D}}_0^{V_1^d})$ let $\Psi_{s,t}(\nu_\kappa) := (\theta_{t-s}[\nu])_-$ where $[\nu]$ is the unique two-sided solution the local equation such that $[\nu]_- = \nu_\kappa$. For $\nu_\kappa \in \mathcal{P}(\overline{\mathcal{D}}_0^{V_1^d})$ not strongly continuous, let $\Psi_{s,t}(\nu_\kappa) := \nu_\kappa$.

Corollary 6.21. *Suppose Assumptions 4 and 5 hold. Then both $\overline{\Psi}$ and Ψ are time-homogeneous, continuous flows, and for any strongly symmetric $\nu_\kappa \in \mathcal{P}(\overline{\mathcal{D}}^{V_1^d})$ with tiled extension $\overline{\nu}_\kappa$ and any $0 \leq s \leq t < \infty$, $\overline{\Psi}_{s,t}(\overline{\nu}_\kappa)$ is the tiled extension of $\Psi_{s,t}(\nu_\kappa)$. This also implies that $\Psi_{s,t}(\nu_\kappa) = \overline{\Psi}_{s,t}(\overline{\nu}_\kappa)[V_1^d]$.*

We defer the proof of this result to Section 10.1.

Remark 6.22. In general, for $0 \leq s, t < \infty$, we suspect that $\Psi_{s,t} : \mathcal{P}(\overline{\mathcal{D}}_0^{V_1^d}) \to \mathcal{P}(\overline{\mathcal{D}}_0^{V_1^d})$ may fail to be continuous, in which case Ψ would not be continuous in the Feller sense.

6.5 Stationary Solutions to the Local Equation and the SDE (6.1.1)

Theorem 6.17 only addresses the relationship between transient solutions to (6.1.1) and the local equation. Here we discuss some simple corollaries that highlight the surprisingly complex relationship between stationary solutions to the local equation and stationary IPS. This section details preliminary work. We start with the following important observation. The 2-SGMRF property is not preserved under weak limits. Because of this, it is very likely that there are examples of IPS for which the local equation experiences a *discontinuity at infinite time* in which limiting distributions of the form $\lim_{t\to\infty} \theta_t[\nu]$, where $[\nu]$ is some strongly symmetric two-sided solution to the local equation, may fail to be stationary even when the limit exists. Therefore, we begin with a simple characterization of this broader class of stationary measures.

Definition 6.23. A stationary measure $\widetilde{\nu} \in \mathcal{P}(\overline{\mathcal{D}}^{V_1^d})$ is called a *symmetric limiting solution to the local equation* if there exists a strongly symmetric two-sided solution $[\nu]$ to the local equation such that $\lim_{t\to\infty} \theta_t[\nu] = \widetilde{\nu}$. Likewise, a stationary measure $\widetilde{\mu} \in \mathcal{P}(\overline{\mathcal{D}}^{V^d})$ is called a *symmetric limiting solution to the SDE* (6.1.1) if there exists an automorphism invariant, 2-SGMRF two-sided solution $[\mu]$ to (6.1.1) such that $\lim_{t\to\infty} \theta_t[\mu] = \widetilde{\mu}$.

It should be immediately clear that all symmetric limiting solutions to the local equation are strongly symmetric, and all symmetric limiting solutions to (6.1.1) are automorphism invariant. However, symmetric limiting solutions to (6.1.1) may fail to be 2-SGMRFs as the 2-SGMRF conditional structure is not necessarily preserved under weak convergence of measures. Thus, under Assumptions 4 and 5, a symmetric limiting solution to (6.1.1) is essentially any stationary solution to (6.1.1) for which there exists an automorphism invariant 2-SGMRF in the basin of attraction.

Corollary 6.24. *Suppose Assumptions 4 and 5 holds. Then for any symmetric limiting solution to* (6.1.1) $\widetilde{\mu}$*,* $\widetilde{\mu}[V_1^d]$ *is a symmetric limiting solution to the local equation.*

The proof of Corollary 6.24 is deferred to Section 10.1. Corollary 6.24 states that the limiting solutions of the local equation describe the marginals of all stationary distributions of the IPS for which there exists an automorphism invariant 2-SGMRF in the basin of attraction. However, it should be noted that it is possible for multiple symmetric limiting solutions to (6.1.1) to map to a single symmetric limiting solution to the local equation.

Looking at symmetric limiting solutions gives us a way to analyze the local equations and retrieve some information about stationary solutions to (6.1.1). However, there is no "reverse mapping", so to speak, that definitively yields a stationary solution to (6.1.1) given a symmetric limiting solution to the local equation. Stationary solutions to the local equation do provide such a mapping.

Corollary 6.25. *Under Assumptions 4 and 5, the tiled extension induces a bijection between strongly symmetric stationary solutions to the local equation and automorphism invariant, 2-SGMRF stationary solutions to* (6.1.1)*.*

The proof of Corollary 6.25 is deferred to Section 10.1. Note that it is currently an open problem under what conditions stationary solutions to (6.1.1) have 2-SGMRF trajectories.

6.6 Markov Local Equation

Studying the local equations numerically is difficult, as the functions γ are defined in terms of conditional expectations where one conditions on the trajectory of the process. Such conditional expectations are exceedingly difficult to compute numerically. In this section, we present a first approximation. The idea behind this approximation is simple: when computing the functions γ, simply disregard the memory of the process. So instead of calculating $\mathbb{E}\left[r_j(t,\widehat{\kappa},Y)|\widehat{\kappa}_{\{\varnothing,1\}},Y_{\{\varnothing,1\}}[t)\right]$, simply compute $\mathbb{E}\left[r_j(t,\widehat{\kappa},Y)|Y_{\{\varnothing,1\}}(t-)\right]$ instead. We call this the Markov approximation. In Section 7, we present a brief numerical analysis of the Markov approximation for four different interacting particle systems. In this section, we introduce the Markov approximation and present one theoretical result describing its efficacy.

Throughout the section, we assume that the IPS is time-homogeneous and Markov.

Assumption 6. For each $j \in \mathcal{J}$, there exists a function $\widetilde{r}_j : \mathcal{X}^{V_1^d} \to \mathbb{R}_+$ such that for all $(t,\widehat{\vartheta},y) \in \mathbb{R}_+ \times (\mathcal{K}\times\mathcal{D})^{V_1^d}$,

$$r_j(t,\widehat{\vartheta},y) = \widetilde{r}_j(y(t-)).$$

Recall from the end of Section 3 that $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ and $\mathbf{N}$ are assumed to be fixed. Let $\widehat{\kappa} := (\varsigma(\widehat{\kappa}_v))_{v\in V_1^d}$ be the $\mathcal{F}_0$-measurable $\mathcal{X}^{V_1^d}$-valued random variable defined above Definition 6.14 (noting that under Assumption 6 we may assume without loss of generality that $\mathcal{K} = \mathcal{X}$ and

$\varsigma : \mathcal{K} \to \mathcal{K}$ is the identity). Then the *Markov local equation* can be defined analogously to the local equation:

Definition 6.26. A (strong) solution to the Markov local equation for the initial data $\widehat{\kappa}$ is a tuple $(\widetilde{Y}, \widetilde{\gamma})$ such that,

(1) $\widetilde{Y}$ is a $\mathcal{D}^{V_1^d}$-valued, $\mathcal{H}^{\widehat{\kappa}} \vee \mathbb{H}^{\mathbf{N}} \subseteq \mathbb{F}$-adapted process;

(2) $\widetilde{\gamma} := \{\widetilde{\gamma}_j\}_{j \in \mathcal{J}}$ is a family of Borel measurable functions $\widetilde{\gamma}_j : \mathbb{R}_+ \times \mathcal{X}^2 \to \mathbb{R}_+$ such that $t \mapsto \widetilde{\gamma}_j(t, \widehat{\vartheta})$ is continuous for all $j \in \mathcal{J}$ and $\widehat{\vartheta} \in \mathcal{X}^2$ and that,

$$\widetilde{\gamma}_j(t, \widetilde{Y}_{\{\varnothing,1\}}(t-)) = \mathbb{E}\left[\widetilde{r}_j(\widetilde{Y}(t-)) | \widetilde{Y}_{\{\varnothing,1\}}(t-)\right] \text{ a.s.,} \tag{6.6.1}$$

for every $t \in \mathbb{R}_+$;

(3) $(\widehat{\kappa}_v, \widetilde{Y}_v, \mathbf{N}_v)_{v \in V_1^d}$ and $\widetilde{\gamma}$ satisfy

$$\widetilde{Y}_v(t) = \begin{cases} \varsigma(\widehat{\kappa}_v) + \int_{(0,t] \times \mathbb{R}_+ \times \mathcal{J}} j \mathbb{I}_{\left\{r \leq \widetilde{\gamma}_j(\widetilde{Y}_{\{v,\varnothing\}}(s-))\right\}} \mathbf{N}_v(ds, dr, dj) & \text{if } v \in \mathcal{N}_\varnothing, \\ \varsigma(\widehat{\kappa}_\varnothing) + \int_{(0,t] \times \mathbb{R}_+ \times \mathcal{J}} j \mathbb{I}_{\left\{r \leq \widetilde{r}_j(\widetilde{Y}(s-))\right\}} \mathbf{N}_\varnothing(ds, dr, dj) & \text{if v} = \varnothing \end{cases}. \tag{6.6.2}$$

We denote the law of a solution to the Markov local equation for $\widehat{\kappa}$ by $\widetilde{\nu}$. We say the solution is pathwise unique if for any strong solutions $(\widetilde{Y}^1, \widetilde{\gamma}^1)$ and $(\widetilde{Y}^2, \widetilde{\gamma}^2)$ to the Markov local equation, $\widetilde{Y}^1$ and $\widetilde{Y}^2$ are indistinguishable.

Unlike the local equation, solutions to the Markov local equation can be described in terms of solutions to its forward equation which is an ODE. For any $U \subseteq V^d$, $a \in \mathcal{X}^U$ and $j \in \mathbb{Z}$, let $S_{v,j}(a) \in \mathcal{X}^U$ be defined by $(S_{v,j}(a))_u = a_u + j\mathbb{I}_{\{u=v\}}$ for $u \in U$. Let $p : \mathbb{R}_+ \times \mathcal{X}^{V_1^d} \to [0,1]$ be a function solving the following initial value problem where the initial state is defined by the function $p_0 : \mathcal{X}^{V_1^d} \to [0,1]$ satisfying $\sum_{a \in \mathcal{X}^{V_1^d}} p_0(a) = 1$:

$$\begin{aligned} \frac{d}{dt} p(t,a) &= \sum_{j \in \mathcal{J}} (\widetilde{r}_j(S_{\varnothing,-j}(a)) p(t, S_{\varnothing,-j}(a)) - \widetilde{r}_j(a) p(t,a)) \\ &+ \sum_{v \in \mathcal{N}_\varnothing} \sum_{j \in \mathcal{J}} \left(\frac{\sum_{\substack{b \in \mathcal{X}^{V_1^d} \\ b_{\{\varnothing,1\}} = (a_v - j, a_\varnothing)}} \widetilde{r}_j(b) p(t,b)}{\sum_{\substack{b \in \mathcal{X}^{V_1^d} \\ b_{\{\varnothing,1\}} = (a_v - j, a_\varnothing)}} p(t,b)} p(t, S_{v,-j}(a)) - \frac{\sum_{\substack{b \in \mathcal{X}^{V_1^d} \\ b_{\{\varnothing,1\}} = a_{\{v,\varnothing\}}}} \widetilde{r}_j(b) p(t,b)}{\sum_{\substack{b \in \mathcal{X}^{V_1^d} \\ b_{\{\varnothing,1\}} = a_{\{v,\varnothing\}}}} p(t,b)} p(t,a) \right). \end{aligned} \tag{6.6.3}$$

(6.6.4)

$p(0,a) = p_0(a).$

In order for the above ODE to be well defined, we apply the convention that $\frac{0}{0} = 0$ above, and for all $a \in \mathcal{X}^{V_1^d}$, $v \in V_1^d$ and $j \in \mathcal{J}$ such that $S_{v,-j}(a) \notin \mathcal{X}^{V_1^d}$, we set $p(t, S_{v,-j}(a)) := 0$.

Proposition 6.27. *Suppose Assumptions 4 and 6 hold, $|\mathcal{X}| < \infty$ and $p_0(a) > 0$ for all $a \in \mathcal{X}^{V_1^d}$. Then there exists a unique, continuously differentiable (in its first argument) solution $p : \mathbb{R}_+ \times \mathcal{X}^{V_1^d} \to [0,1]$ to* (6.6.3)*, and there exists a pathwise unique solution $(\widetilde{Y}, \widetilde{\gamma})$ to the Markov local equation. Furthermore, for each $t \in \mathbb{R}_+$ and $a \in \mathcal{X}^{V_1^d}$, $\mathbb{P}(\widetilde{Y}(t) = a) = p(t,a)$.*

The proof of Proposition 6.27 is deferred to Section 10.2. We do not yet have many theoretical guarantees of how well the Markov local equation performs. However, we do identify one situation in which the process described by the Markov local equation exactly matches the marginal distribution of a stationary solution to the SDE (6.1.1).

First, note that under Assumptions 4 and 6, if $[\mu] \in \mathcal{P}(\overline{\mathcal{D}}^{V^d})$ is a stationary solution to (6.1.1) and $X \sim [\mu]$, then $[\mu]$ is completely characterized by Law$(X(0))$, so in the Markov case there is no need to define stationary solutions as probability measures on path space. We will say that $\pi \in \mathcal{P}(\mathcal{X}^{V^d})$ a stationary solution to (6.1.1) if there exists a stationary solution $[\mu] \in \mathcal{P}(\overline{\mathcal{D}}^{V^d})$ to (6.1.1) such that for any $X \sim [\mu]$, $X(0) \sim \pi$. Then we present the following result which we prove in Section 10.2.

Proposition 6.28. *Suppose Assumptions 4 and 6 hold. If there exists an automorphism invariant, 2-SGMRF stationary solution $\pi \in \mathcal{P}(\mathcal{X}^{V^d})$ to* (6.1.1)*, then the unique solution $p : \mathbb{R} \times \mathcal{X}^{V_1^d} \to [0,1]$ to* (6.6.3) *with respect to the initial condition $p_0(a) = \pi(\{a\} \times \mathcal{X}^{V^d \setminus V_1^d})$, $a \in \mathcal{X}^{V_1^d}$ is also stationary (in the sense that $p(t,a)$ does not depend on t).*

Remark 6.29. Note that we do *not* in general expect the transient dynamics of solutions to the Markov local equation to equal the dynamics of solutions to the local equation or the marginal trajectory of the IPS even in the case that the stationary solutions of all three processes are the same.

CHAPTER 7

Some Numerical Results

In this section we illustrate Proposition 6.28 numerically for the Ising model with no external magnetic field on $\mathbb{T}_4$. We then apply a number of standard approximation algorithms for to three Markovian IPS: the SIR model on $\mathbb{T}_2$, the voter model on $\mathbb{T}_3$ and the contact process on $\mathbb{T}_2$, all of which are defined in Section 3.2. In each case we investigate different functionals. For the SIR model, we simply plot the probability that the root particle is susceptible and infected at each time. In the voter model we investigate neighbor-neighbor correlations by plotting the number of neighbors with which the root is in agreement with as time progresses. Lastly, in the contact process we test how the various approximations perform at stationarity by using each approximation to estimate the upper-invariant measure of the contact process. This also allows us to detect differences in how well each approximation estimates the phase transition of the contact process.

In this context, we refer to solutions to the Markov local equation as the Markov approximation (MA). Additionally, we show that while the trajectory of the MA is not the same as the trajectory of the solution to the local equation, the MA still provides a reasonable approximation of the trajectory of the IPS in the neighborhood of the root in many situations. To highlight this, we also present comparisons to the mean field approximations (MFA) and pairwise approximations (PA) of these processes both of which are explicitly defined below.

7.1 The Approximations

In this section we describe the different approximations which we compare to the MA. We give some intuition regarding how each can be derived and, in the case of the pairwise approximation, state the specific ODEs we will be using for the specific models.

Consider the example of the SIR model. Then the standard MFA of the SIR model (though this model predated the networked SIR model and so was not considered an MFA at the time [54]) is given by the solution to the following ODE:

$$\begin{aligned}\frac{dp_2}{dt} &= p_1\\ \frac{dp_1}{dt} &= \lambda p_0 p_1 - p_1\\ \frac{dp_0}{dt} &= -\lambda p_0 p_1.\end{aligned} \tag{7.1.1}$$

Define $\gamma := \delta_1$ and let $A : \mathbb{R} \times \mathbb{R} \times \mathcal{P}(\{0,1,2\}) \to \mathbb{R}_+$ satisfy,

$$A(a, j, \eta) := \mathbb{I}_{\{j=1\}} \left(\mathbb{I}_{\{a=1\}} + \mathbb{I}_{\{a=0\}} \lambda \int \mathbb{I}_{\{b=1\}} \eta(db) \right).$$

Then it is easily seen that A, γ satisfy [73, Assumptions (B1)-(B4)] (or rather the sufficient conditions that Oelschläger claims imply these assumptions). Then in the notation of that paper, B^N is the infinitesimal generator describing the dynamics of the SIR model on the complete graph of N vertices with infection rate $\frac{\lambda}{N}$, and by [73, Theorem 2], the empirical measure of this process converges weakly to the law of a process $\overline{X}$ such that $\mathbb{P}(\overline{X}(t) = i) = p_i(t)$ for each $i = 0, 1, 2$. Therefore, the mean field approximation for the SIR model actually describes the limiting behavior of an SIR model in which the infection rate λ depends on the size of the graph G_n. This introduces a problem. Given that we want to apply the MFA to the SIR model on the graph $\mathbb{T}_d$, what value of λ should we use? Consider the following heuristic. Let G_n be the complete graph on n vertices and suppose that at a given time t, In vertices are infected. Then in the scaled MF model, any susceptible particle will be infected at rate $In\frac{\lambda}{n} = I\lambda$. Now consider the SIR model on the regular tree $\mathbb{T}_d$ with infection rate λ and, formally speaking, suppose at time t, each particle is independently infected with probability I. Then on average, we would expect a susceptible particle to be infected at a rate of $dI\lambda$, which is d times the average rate at which a particle is infected in the mean-field approximation. Therefore it makes sense to approximate the SIR model on $\mathbb{T}_d$ with infection rate λ by its MFA with infection rate λd. Indeed, this heuristic is a standard one in the application of the MFA to contact processes on heterogeneous graphs [76].

More generally, we can derive this modified MFA using a process very similar to the derivation of the MA. The basic idea is simple. When we derive the local equation, we first show that IPS has an automorphism invariant law and satisfies the 2-SGMRF property on path space. The MA can be seen as an approximation stemming from the assumption that at any given time t, $\mathrm{Law}(X(t))$ is also an automorphism invariant 2-SGMRF. For the MFA, we replace this assumption with the assumption that $X(t) = (X_v(t))_{v\in V^d}$ is always i.i.d.. Let $\{X_v^{MFA}\}_{v\in V^d}$ be an i.i.d. collection of càdlàg processes on $\mathcal{X}$ such that for a given Poisson process $\mathbf{N}$ on $\mathbb{R}_+^2 \times \mathcal{J}$ with intensity $\mathrm{Leb}^2 \otimes \#_{\mathcal{J}}$, independent of X_u^{MFA}, $u \neq \varnothing$, and $X_\varnothing^{MFA}(0)$,

$$X_\varnothing^{MFA}(t) = X_\varnothing^{MFA}(0) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r\le\widetilde{r}_j(X_{\mathrm{cl}_\varnothing}^{MFA})\right\}}\mathbf{N}(ds,dr,dj).$$

Note that X^{MFA} satisfies (6.1.1) for $v = \varnothing$. Then the function $p^{MFA} : \mathbb{R}_+ \times \mathcal{X} \to [0,1]$ defined by $p^{MFA}(t,a) = \mathbb{P}(X_\varnothing^{MFA}(t) = a)$ for $(t,a) \in \mathbb{R}_+ \times \mathcal{X}$ satisfies the following initial value problem for $p_0^{MFA} : \mathcal{X} \to [0,1]$ given by $p_0^{MFA}(a) = \mathbb{P}(X_\varnothing^{MFA}(0) = a)$:

$$\frac{d}{dt}p^{MFA}(t,a) = \sum_{j\in\mathcal{J}} \sum_{\substack{b\in\mathcal{X}^{V_1^d}\\ b_\varnothing=a-j}} \widetilde{r}_j(b) \prod_{v\in V_1^d} p(t,b_v) - \sum_{j\in\mathcal{J}} \sum_{\substack{b\in\mathcal{X}^{V_1^d}\\ b_\varnothing=a}} \widetilde{r}_j(b) \prod_{v\in V_1^d} p(t,b_v) \tag{7.1.2}$$

$$p^{MFA}(0,a) = p_0^{MFA}(a). \tag{7.1.3}$$

It is easily shown that (7.1.2) is equivalent to the ODE derived in [76] for the contact process. For the SIR model, the ODE above is equivalent to (7.1.1) with λ replaced by λd. Lastly, for the linear voter model, the above ODE simplifies to $\frac{d}{dt}p^{MFA} \equiv 0$. As these are standard results, we omit rigorous justification of (7.1.2) noting simply that such justification would follow along the same lines as the proof of Proposition 6.27.

The PA is a refinement of the MFA which takes pairwise correlations between neighboring particles into account. Let X^{PA} be the process representing the pair approximation. The idea behind the pairwise approximation is to derive a forward equation ODE that describes at any given time t the probability that the pair $(X_u^{PA}(t), X_v^{PA}(t))$ takes a given value for a typical edge $\{u,v\} \in E^d$. The approximating assumption made is that for any pair of edges $\{u,u'\}$ and $\{v,v'\}$ composed of distinct vertices the random variables $(X_u^{PA}(t), X_{u'}^{PA}(t))$ and $(X_v^{PA}(t), X_{v'}^{PA}(t))$ are independent. Furthermore, given a path $\{u, v, w\}$, $X_u^{PA}(t)$ and $X_w^{PA}(t)$ are conditionally independent given $X_v^{PA}(t)$. For the contact process, the full derivation is given in [68] with some small modifications. First, as we are working with the regular tree $\mathbb{T}_d$, we do not need to consider heterogeneities in the degree of the graph. Second, the pairwise approximation of [68] is computed with respect to the scaled contact process used for the MFA, so we replace λ by λd in the computation. We note that the SIR model has similar dynamics, to the contact process and therefore simply adapt the pairwise ODE from [68] to the SIR process as described below. Note that a pairwise approximation for the (non-Markovian) SIR model can be found in [84], however this model does not take network topology into account and so we choose to work with the modified ODE from [68] for consistency. Lastly, the pairwise approximation from the voter model that we use can be found in [91]. We write out the ODEs below for each of the three processes setting $p_{ab}^{PA}(t) = p_{ba}^{PA}(t) = \mathbb{P}(X_{\{u,v\}}^{PA}(t) = (a, b))$ for $(a, b) \in \mathcal{X}^2$ and a typical edge $\{u, v\}$. We start with the contact process pairwise approximation from [68]:

$$\begin{aligned}
\frac{d}{dt}p_{01}^{PA} &= -p_{01}^{PA}(1+\lambda) + p_{11}^{PA} + \lambda(d-1)\left(\frac{p_{01}^{PA}(p_{00}^{PA} - p_{01}^{PA})}{p_{01}^{PA} + p_{00}^{PA}}\right) \\
\frac{d}{dt}p_{00}^{PA} &= -2\lambda(d-1)\left(\frac{p_{01}^{PA}p_{00}^{PA}}{p_{01}^{PA} + p_{00}^{PA}}\right) + 2p_{01}^{PA}
\end{aligned} \tag{7.1.4}$$

Then the equation above determines $p_{11}^{PA} = 1 - p_{00}^{PA} - 2p_{01}^{PA}$. As mentioned earlier SIR model is a straightforward modification of the above equation. We simply need to change the denominator to reflect the existence of a 2 state and change all transitions from the 1 state to the 0 state into transitions from the 1 state to the 2 state.

$$\begin{aligned}
\frac{d}{dt}p_{00}^{PA} &= -2\lambda(d-1)\left(\frac{p_{01}^{PA}p_{00}^{PA}}{p_{02}^{PA} + p_{01}^{PA} + p_{00}^{PA}}\right) \\
\frac{d}{dt}p_{01}^{PA} &= -p_{01}^{PA}(1+\lambda) + \lambda(d-1)\left(\frac{p_{01}^{PA}(p_{00}^{PA} - p_{01}^{PA})}{p_{02}^{PA} + p_{01}^{PA} + p_{00}^{PA}}\right) \\
\frac{d}{dt}p_{02}^{PA} &= -\lambda(d-1)\left(\frac{p_{01}^{PA}p_{02}^{PA}}{p_{02}^{PA} + p_{01}^{PA} + p_{00}^{PA}}\right) + p_{01}^{PA} \\
\frac{d}{dt}p_{11}^{PA} &= -2p_{11}^{PA} + 2\lambda(d-1)\left(\frac{p_{01}^{PA}p_{01}^{PA}}{p_{02}^{PA} + p_{01}^{PA} + p_{00}^{PA}}\right) + 2\lambda p_{01}^{PA}
\end{aligned}$$

$$\frac{d}{dt}p_{12}^{PA} = -p_{12}^{PA} + p_{11}^{PA} + \lambda(d-1)\left(\frac{p_{01}^{PA}p_{02}^{PA}}{p_{02}^{PA}+p_{01}^{PA}+p_{00}^{PA}}\right), \tag{7.1.5}$$

where $p_{22}^{PA} = 1 - (p_{00}^{PA} + p_{11}^{PA}) - 2(p_{01}^{PA} + p_{02}^{PA} + p_{12}^{PA})$.

Lastly we apply equation (5) of [91], where we note that $\rho = 2p_{01}^{PA}$ and we set $p_0 = \mathbb{P}(X_{\varnothing}^{PA}(0) = 0)$:

$$\frac{d}{dt}p_{01}^{PA} = \frac{2p_{01}^{PA}}{d}\left((d-1)\left(1 - \frac{p_{01}^{PA}}{p_0(1-p_0)}\right) - 1\right). \tag{7.1.6}$$

Again there is an alternative approach that can be taken to derive the pairwise approximation. Suppose that the law of the process X^{PA} at any given time is an automorphism invariant 1-SGMRF. Then by repeating arguments analogous to those used to derive the Markov local equation and prove Proposition 6.27, one can derive the following coupled ODE for $p_a^{PA} : \mathbb{R}_+ \to [0,1]$, $a \in \mathcal{X}^2$:

$$\begin{aligned}\frac{d}{dt}p_a^{PA}(t) = &\sum_{v\in\{\varnothing,1\}}\sum_{j\in\mathcal{J}}\sum_{\substack{b\in\mathcal{X}^{\mathrm{cl}_v}\\ b_{\{\varnothing,1\}}=S_{v,-j}(a)}} \widetilde{r}_j(b)\frac{\prod_{u\in\mathcal{N}_v} p_{b_{\{u,v\}}}^{PA}(t)}{\left(\sum_{c\in\mathcal{X}} p_{(b_v,c)}^{PA}(t)\right)^{d-1}}\\ &- \sum_{v\in\{\varnothing,1\}}\sum_{j\in\mathcal{J}}\sum_{\substack{b\in\mathcal{X}^{\mathrm{cl}_v}\\ b_{\{\varnothing,1\}}=a}} \widetilde{r}_j(b)\frac{\prod_{u\in\mathcal{N}_v} p_{b_{\{u,v\}}}^{PA}(t)}{\left(\sum_{c\in\mathcal{X}} p_{(b_v,c)}^{PA}(t)\right)^{d-1}}.\end{aligned} \tag{7.1.7}$$

It can be shown through direct computation that (7.1.7) is equivalent to (7.1.4), (7.1.5) and (7.1.6), so we omit the full justification of (7.1.7).

7.2 Numerics

Because equations (6.6.3) and (7.1.7) are unstable, we use the implicit Runge-Kutta method of the Radau IIA family of order 5 (which is the default Radau method used by the solve_ivp method in the Python's SciPy library) to solve it. In all cases, our program took less than a second to numerically solve all relevant ODEs.

A. Ferromagnetic Ising Model: All finite state extremal local MRFs on trees are SGMRFs [92, Corollary 2], which means that given no magnetic field, there exists a unique Ising model that is also a SGMRF at high temperatures. At low temperatures, there exist two automorphism invariant extremal Ising models (given by positive and negative boundary conditions respectively) which must also be SGMRFs. Thus, these both satisfy the conditions of Proposition 6.28. We ran the

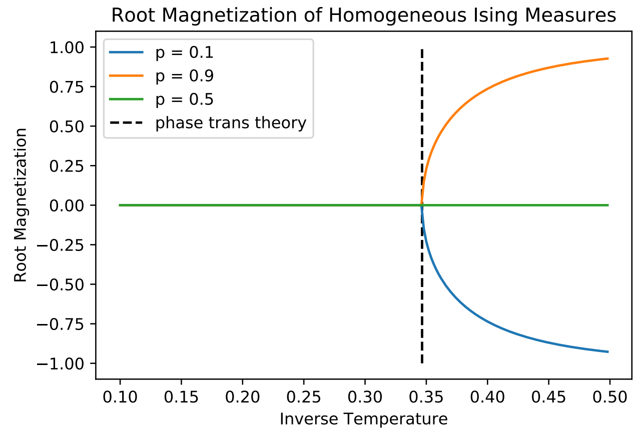
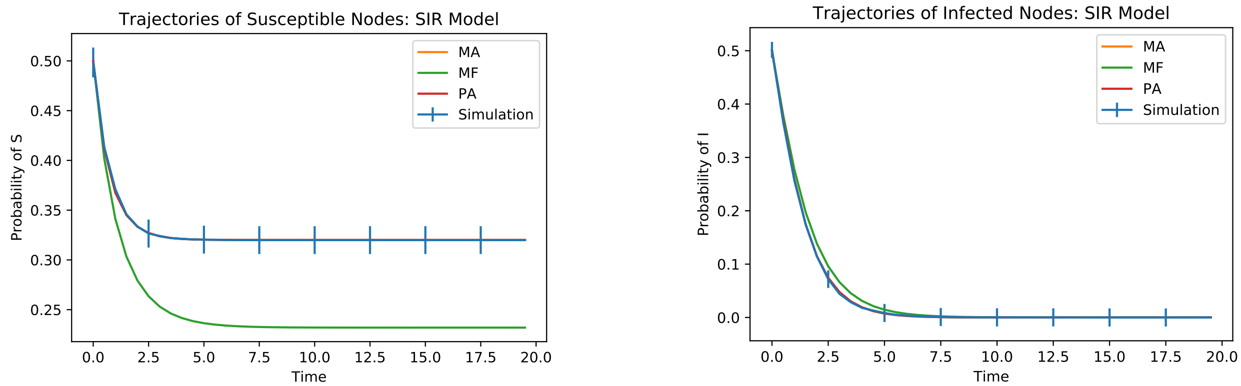


Figure 7.1: Root Magnetization of the Ising model for different values of β with different initializations. The phase transition matches the theoretical phase transition.

ODE solver for 2000 units of time starting with $p_0(\cdot)$ describing a product of probability measures on $\{\pm 1\}$ placing respective weights of 0.1,0.5 and 0.9 on the set $\{1\}$. In Figure 7.1, we plot the root magnetization (expected value of the Ising model at the root at time 2000) for different values of the inverse temperature, β varying from 0.1 to 0.5 with a step size of 0.02. This was all evaluated on the tree $\mathbb{T}_4$ and this process captured the unique Ising model for $\beta < \beta_c$, where β_c is the critical inverse temperature. For $\beta > \beta_c$ the Markov local equation captured the two SGMRFs generated by positive and negative boundary conditions. It also captured the non-extremal Ising measure generated with free boundary conditions. The vertical dashed line marks the known theoretical value of $\beta_c := \frac{1}{2} \log \frac{3+1}{3-1} = \frac{1}{2} \log(2)$ [41, Equation (12.28)]. Note that the results in Figure 7.1 are exact by Proposition 6.28.

B. SIR Model: We first simulate the SIR model with $\lambda = \frac{1}{2}$ on a cycle graph ($\mathbb{T}_2$) with $N = 1001$ nodes, where each node is initially infected or susceptible with probability $1/2$ each, and compare it against the MA, MFA and the PA. We consider 100 time steps averaged over 10000 samples. Error bars are 95% confidence intervals adjusted using the conservative Bonferroni correction to account for multiple error bars. Figures 7.2a and 7.2b plot the evolution of the probability of a node being susceptible or infected, respectively, as a function of time. As we can see, the outputs of both the MA and PA are nearly indistinguishable from the simulation of the SIR model.



(a) Trajectory of S probabilities.

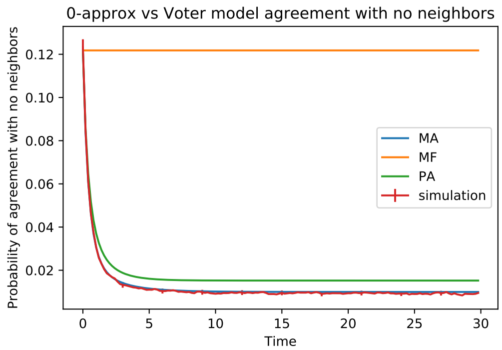
(b) Trajectory of I probabilities.

Figure 7.2: The evolution of an infection via SIR.

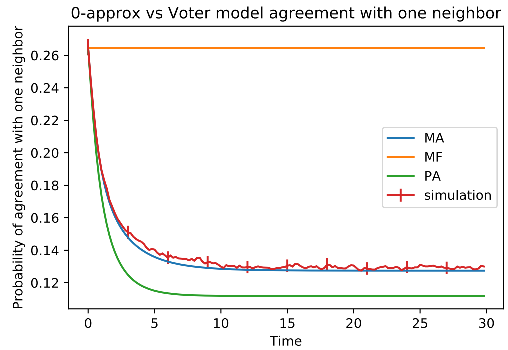
C. Voter model: Now consider the linear voter model (LVM) on $\mathbb{T}_3$. On an infinite tree it is known not to converge to consensus [65, Chapter V, Theorem 1.8]. In our simulation, we approximate $\mathbb{T}_3$ by a tree with 9 generations below the root. We initialize the process with each voter supporting position 1 independently with probability 0.3. We averaged the trajectory of 10^5 independent simulations; errors are given by 95% confidence intervals adjusted using the Bonferroni correction. As the figure shows, even the simple linear voter model exhibits some complex behavior as time progresses. While the MFA is unable to capture any of this complexity, the PA does capture the qualitative behavior of the process while the MA follows the simulation closely.

D. Contact Process: Finally, we test the stationary behavior of the Markov local equation for the contact process which has a non-trivial stationary distribution that is not a 2-SGMRF. The contact process is a simple model with interesting qualitative behavior. At a low infection rate λ, the disease converges to extinction, but for large λ, it admits a non-trivial stationary distribution. In particular, there exists a unique non-trivial automorphism invariant extremal stationary probability measure of the supercritical contact process called the *upper-invariant measure* [65, Chapter VI, Theorem 1.5].

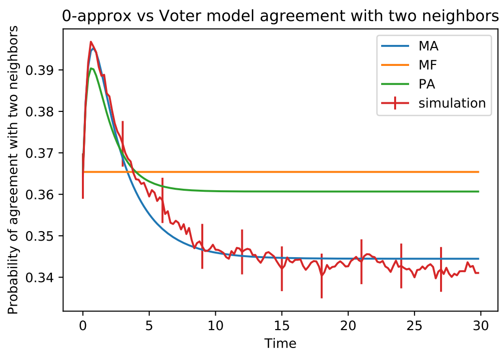
We display the total variation distance between the upper-invariant measure predicted by direct simulation and the upper-invariant measure predicted by each of the MA, the PA and the MFA. All total variation distances are calculated on the distribution of the root and its neighbors in $\mathbb{T}_2$ for varying λ. The contact process upper invariant measure was computed on 1001 nodes initialized



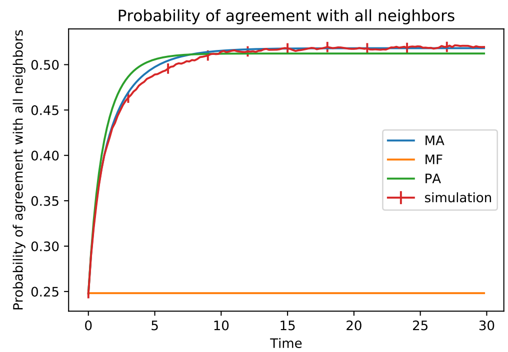
(a) Agreement with no neighbors.



(b) Agreement with one neighbor.



(c) Agreement with two neighbors.



(d) Agreement with all neighbors.

Figure 7.3: Agreement between neighbors of the voter model over time.

with all nodes infected. The boundary nodes were fixed at 1. We ran the simulation for 10^3 time steps, and then estimated the upper invariant measure by the average time spent at each state over the next 10^4 time steps. The error bars are 95% confidence intervals calculated taking into account auto-correlations which we computed separately [32, Theorem 8.3.2]. Error bars are adjusted using the Bonferroni correction.

As we can see, none of these approximations perform well. However, although the MA underestimates the critical point significantly, the PA and the MFA are significantly worse. The critical point is known to lie between 1.539 and 1.942 [66] and it is believed to be approximately 1.6494 based on numerical approximations [72, page 360]. However, for λ significantly larger and smaller than its critical value, the MA provides a moderately accurate approximation of the upper invariant measure of the contact process.

D. Concluding Remarks. Above, we use the Ising model to illustrate how the Markov local

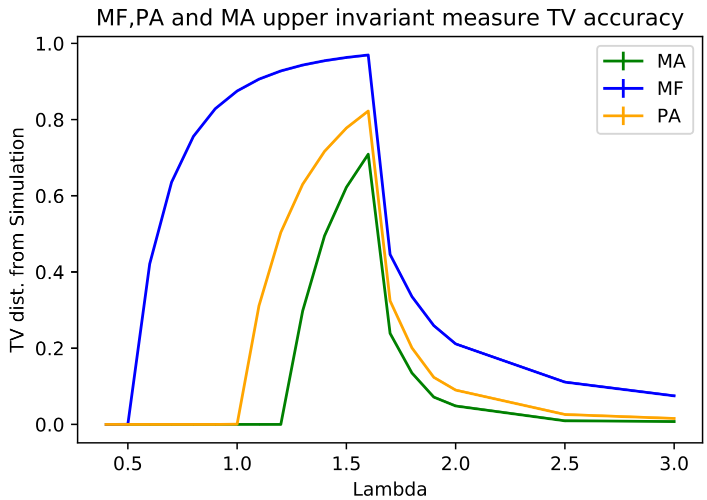


Figure 7.4: Contact Process: TV distance between MA and simulation

equation can be used to compute phase transitions. In the remaining examples, the MFA performs poorly, and while the PA performs better, it is still consistently outdone by the MA. The above results suggest that the MA is reasonably accurate for transient functionals of IPS, and is significantly faster to implement than simulations. Figure 7.4 demonstrates that the MA can also approximate stationary solutions well under certain conditions. Indeed, by Corollary 6.24, the inaccuracy of the MA near criticality stems from the difference between the MA and the solution to the local equation. Interesting directions of future work include understanding long-time behavior, developing tractable numerical algorithms to the local equations, quantifying errors of approximations to the local equations such as MA and more refined versions, and applications, involving more intensive simulations, to a variety of IPS models.

CHAPTER 8

Proofs of Intermediate Results

8.1 A Proof that IPS are Automorphism-Invariant 2-SGMRFs

In this section, we prove Proposition 6.10, which states that under certain conditions, the IPS is an automorphism invariant 2-SGMRF. The fact that the IPS is automorphism invariant follows directly from Lemma 6.3, so it suffices to prove that the IPS forms a 2-SGMRF. A rough outline of this proof is as follows. Fix G, κ, and $\{r_j^v\}_{v \in V, j \in \mathcal{J}}$ as in the proposition, let $X = \{X_v\}_{v \in V}$ be a solution to the SDE (6.1.1), and let μ denote its law, and recall that $\mu_{t-} = \mathrm{Law}(\kappa, X[t))$. Also, assume that $\mu_\kappa := \mathrm{Law}(\kappa)$ is a 2-SGMRF on $\mathcal{K}^V$ with respect to G. For any $t > 0$, our proof that μ_{t-} is also a 2-SGMRF can be broken into three main steps. First, in Section 8.1.1, we construct a sequence of $\mathcal{X}^V$-valued "reference processes" $\widehat{X}^n, n \in \mathbb{N}$, with initial data κ, such that for any partition A, B, S of V for which $S = \mathcal{N}_A^2$ is finite, we have

$$(\kappa_A, \widehat{X}_A^n[t)) \perp\!\!\!\perp (\kappa_B, \widehat{X}_B^n[t)) | (\kappa_S, \widehat{X}_S^n[t)), \quad t \in (0, \infty) \tag{8.1.1}$$

for n sufficiently large (depending on S). Then in Section 8.1.2, we compute the Radon-Nikodym derivative of the law on path space of the IPS with respect to that of the corresponding reference process $\widehat{X}^n$ by first establishing a duality relation between IPS and point processes (see Proposition 8.12) and then leveraging results from point process theory. In particular, this yields an explicit factorization of the Radon-Nikodym derivative, which is combined in Section 8.1.3 with a result from [81] to prove that μ_{t-} is a 2-SGMRF with respect to G for all $t > 0$. Finally, the result is extended to the case $t = \infty$ via a standard martingale argument akin to that used in [59, Section 3.2].

8.1.1 Reference Processes and their Conditional Independence Properties

For the remainder of the paper, we assume that $\mathcal{X} = \mathbb{Z}$. Note that this is without loss of generality because the IPS X may be regarded as a process with state space $\mathbb{Z}^V$ such that $X_v(t) \in \mathcal{X}$ almost surely for all $v \in V$ and $t \in \mathbb{R}_+$. Given initial data κ and recalling the definition of ς from Section 6.1, let $W \subseteq V$ be any finite vertex set and define the jump rate functions

$$\widehat{r}_j^{v,U}(\cdot,\cdot,\cdot) = \begin{cases} r_j^v(\cdot,\cdot,\cdot) & \text{if } v \notin U, \\ 1 & \text{otherwise.} \end{cases} \tag{8.1.2}$$

Then let $\widehat{X}^W$ be the almost surely unique solution to the SDE

$$\widehat{X}_v^W(t) = \varsigma(\kappa_v) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r \leq \widehat{r}_j^{v,W}(s,\kappa,\widehat{X}^W)\right\}} \mathbf{N}_v(ds,dr,dj) \text{ for } v \in V, \tag{8.1.3}$$

which is strongly well-posed by Lemma 6.3. Now, fix an arbitrary vertex $\varnothing \in V$, and define

$$V_n^G := \{v \in V : d_G(v,\varnothing) \leq n\}, \qquad n \in \mathbb{N}.$$

We then construct an associated sequence of representative processes $\{\widehat{X}^n\}_{n\in\mathbb{N}}$ as defined below.

Definition 8.1. For each $n \in \mathbb{N}$, let $\widehat{\mu}^n := \mathrm{Law}(\kappa, \widehat{X}^n)$ denote the joint distribution of the initial data κ and the solution $\widehat{X}^n := \widehat{X}^{V_n^G}$ to the SDE (8.1.3) with $W = V_n^G$. Also, let $\widehat{\mu}_\kappa^n := \mathrm{Law}(\kappa)$ denote the first marginal, and for $t > 0$, let $\widehat{\mu}_t^n$ and $\widehat{\mu}_{t-}^n$ denote the joint laws of $(\kappa, \widehat{X}^n[t])$ and $(\kappa, \widehat{X}^n[t))$, respectively.

We now state the main result of this section, Proposition 8.2, which shows that $(\kappa, \widehat{X}^n)$ has a conditional structure that partially resembles a 2-SGMRF.

Proposition 8.2. *[37] Suppose Assumptions 4 holds. Fix $n \geq 2$. Let $A, B, S \subseteq V$ form a partition of V such that $S = \mathcal{N}_A^2 \subseteq V_{n-1}^G$ and also suppose $\kappa_A \perp\!\!\!\perp \kappa_B | \kappa_S$. Then for any $t \in (0,\infty)$,*

$$(\kappa_A, \widehat{X}_A^n[t)) \perp\!\!\!\perp (\kappa_B, \widehat{X}_B^n[t)) | (\kappa_S, \widehat{X}_S^n[t)). \tag{8.1.4}$$

The proof of Proposition 8.2, given below, relies on the following abstract technical lemma, which provides sufficient conditions under which conditional independence properties can be transferred from one collection of random elements to another.

Lemma 8.3. *[37] For each $i, j = 1, \dots, 3$, let Z_i^j be a random element taking values in some Polish space $\mathcal{Z}_i^j$ that satisfies the following properties:*

1. *$\{Z_i^2\}_{i=1,2,3}$ is a set of mutually independent random elements that is independent of $\{Z_i^1\}_{i=1,2,3}$;*
2. *$\mathcal{H}^{Z_i^1} \subseteq \mathcal{H}^{Z_i^3} \subseteq \mathcal{H}^{Z_i^{\{1,2\}}}$ for $i = 1,2,3$, where $Z_i^{\{1,2\}} = (Z_i^1, Z_i^2)$.*

Then

$$Z_1^1 \perp\!\!\!\perp Z_2^1 | Z_3^1 \tag{8.1.5}$$

implies $Z_1^3 \perp\!\!\!\perp Z_2^3 | Z_3^3$.

Relegating the proof of the lemma to Appendix C.1, we now apply it to prove Proposition 8.2.

Proof of Proposition 8.2: Fix a partition $A, B, S \subseteq V$ satisfying the conditions of the proposition, and for notational convenience, set $D_1 := A, D_2 := B, D_3 := S$. To prove Proposition 8.2, it clearly suffices to show that the conditions of Lemma 8.3 are satisfied by $Z_i^1 = \kappa_{D_i}$, $Z_i^2 = \mathbf{N}_{D_i}$ and $Z_i^3 = (\kappa_{D_i}, \widehat{X}^n_{D_i}[t))$, $i = 1,2,3$, since then Lemma 8.3 implies that $Z_1^3 \perp\!\!\!\perp Z_2^3 | Z_3^3$, or equivalently, $(\kappa_A, \widehat{X}^n_A[t)) \perp\!\!\!\perp (\kappa_B, \widehat{X}^n_B[t)) | (\kappa_S, \widehat{X}^n_S[t))$, which is what we want to prove.

First, note that $\kappa_{D_1} \perp\!\!\!\perp \kappa_{D_2} | \kappa_{D_3}$ by assumption and so (8.1.5) holds. Further, since D_i, $i = 1,2,3$, are disjoint, $\{\mathbf{N}_{D_i}\}_{i=1,2,3}$ are mutually independent and, by assumption, also independent of $\{\kappa_{D_i}\}_{i=1,2,3}$. Thus, condition 1 of Lemma 8.3 also holds. It only remains to verify the measurability condition stated in condition 2. The first inclusion $\mathcal{H}^{Z_i^1} \subseteq \mathcal{H}^{Z_i^3}$ holds trivially for $i = 1,2,3$. To prove the second inclusion, note that $\widehat{X}^n$ is the a.s. unique solution to the SDE (8.1.3) with $W = V_n^G$ and so the locality assumption in condition (1) of Assumption 4 and (8.1.2) imply that its marginal $\widehat{X}^n_A$ on the set A solves the following SDE:

$$\widehat{X}^n_v(t) = \begin{cases} \varsigma(\kappa_v) + \int_{(0,t]\times(0,1]\times\mathcal{J}} j\mathbf{N}_v(ds,dr,dj) & \text{if } v \in V_n^G \cap A, \\ \varsigma(\kappa_v) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r \le \widetilde{r}^v_j(s, \widehat{X}^n_{\mathrm{cl}_v}, \kappa_{\mathrm{cl}_v})\right\}}\mathbf{N}_v(ds,dr,dj) & \text{if } v \in A \setminus V_n^G, \end{cases} \tag{8.1.6}$$

where $\widetilde{r}_j^v$ is the function from condition (1) of Assumption 4. We now claim that $\mathrm{cl}_v \subseteq A$ for every $v \in A \setminus V_n^G$, and thus the SDE (8.1.6) for the marginal $\widehat{X}_A$ is autonomously defined. The claim holds because A, B and S form a partition of V and for any $v \in A \setminus V_n^G$, the fact that $v \in A$ and the assumption on S imply $\mathrm{cl}_v \cap B \subseteq \mathcal{N}_A \subseteq \mathcal{N}_A^2(G) = S \subseteq V_{n-1}^G$, whereas the fact that $v \notin V_{n-1}^G$ implies $\mathrm{cl}_v \cap V_{n-1}^G = \emptyset$.

We now show that the SDE (8.1.6) is strongly well-posed. Fix any solution space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ supporting the driving Poisson processes $\widetilde{\mathbf{N}}_A$ and κ_A, and consider an extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{\mathbb{P}})$ of the solution space that also supports i.i.d. Poisson processes $\widetilde{\mathbf{N}}_{V\setminus A}$ and $\kappa_{V\setminus A}$ such that $\widetilde{\mathbf{N}} := (\widetilde{\mathbf{N}}_v)_{v\in V}$ is a collection of driving Poisson processes of (8.1.3) (in the sense of Definition 3.3) with $W = V_n^G$. Let $\widetilde{Y}^A$ and $\widetilde{Z}^A$ be two weak solutions to (8.1.6) on the solution space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with the same initial data κ_A, and (by strong well-posedness of (8.1.3)) let $\widetilde{X}$ be the a.s. unique solution to (8.1.3) with $W = V_n^G$ and initial data κ. Then, as argued above, the processes

$$\widetilde{Y}_v := \begin{cases} \widetilde{Y}_v^A & \text{if } v \in A \\ \widetilde{X}_v & \text{if } v \notin A \end{cases} \quad \text{and} \quad \widetilde{Z}_v := \begin{cases} \widetilde{Z}_v^A & \text{if } v \in A \\ \widetilde{X}_v & \text{if } v \notin A \end{cases},$$

are both solutions to (8.1.3) on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with $W = V_n^G$. Therefore $\widetilde{X} = \widetilde{Y} = \widetilde{Z}$ a.s., and so in particular $\widetilde{Y}^A = \widetilde{Z}^A$ a.s.. Thus, (8.1.6) is strongly well-posed which completes the proof of the claim. By the strong well-posedness of (8.1.6), $\widehat{X}_A^n$ is the a.s. unique solution to (8.1.6) driven by $\mathbf{N}_A$, and by Lemma 3.11, $(\kappa_A, \widehat{X}_A^n[t])$ must be $\mathcal{H}^{\kappa_A} \vee \mathcal{H}_{t-}^{\mathbf{N}_A}$-measurable. This proves the second inclusion in condition 2 for the case $i = 1$.

The case $i = 2$ can be argued similarly. We first claim that $\mathcal{N}_B^2(G) \subseteq S$. Indeed, if the claim were not true, then since A, B, S is a (disjoint) partition of V, it must be true that $\mathcal{N}_B^2(G) \cap A \neq \emptyset$ or equivalently, there must exist $u \in A$ and $v \in B$ such that $d_G(u, v) \leq 2$. However, since $S = \mathcal{N}_A^2(G)$ by assumption, this implies $v \in \mathcal{N}_A^2(G) \cap B = S \cap B$, which contradicts the fact that $S \cap B = \emptyset$. Given the claim, an identical argument as that used for $i = 1$ shows that $(\kappa_B, \widehat{X}_B^n[t])$ is $\mathcal{H}^{\kappa_B} \vee \mathcal{H}_{t-}^{\mathbf{N}_B}$-measurable, thus proving the second inclusion in condition 2 of Lemma 8.3 when $i = 2$. The proof for the case $i = 3$ is much simpler. By (8.1.3) with $W = V_n^G$. for each $v \in S \subset V_n^G$, $\widehat{X}_v^n$ is given by the one-dimensional, $\mathcal{H}^{\kappa_S} \vee \mathbb{H}^{\mathbf{N}_S}$-adapted process $t \mapsto \varsigma(\kappa_v) + \int_{(0,t]\times(0,1]\times\mathcal{J}} j\mathbf{N}_v(ds, dr, dj)$. This shows that $\widehat{X}_v^n$ is a measurable function of $(\kappa_v, \mathbf{N}_v)$. Since this is true for each $v \in S$, this proves condition 2 of Lemma 8.3 also holds for $i = 3$. This completes the verification of the conditions of

Lemma 8.3, and thus proves the proposition. $\square$

8.1.2 Change of Measure Results on Infinite Graphs

In this section, we state the form of the Radon-Nikodym derivative, $d\mu_{t-}/d\widehat{\mu}^n_{t-}$ of the law of the IPS X on the time interval $[0,t)$ and its initial data κ with respect to that of the reference process $\widehat{X}^n$ in the same time interval and its initial data κ on an event of $\widehat{\mu}^n_{t-}$-full measure. To state our result we first introduce the notions of proper trajectories and their so-called jump characteristics. Recall the definition of a *proper process* from Lemma 6.3.

Definition 8.4. [37] Fix $U \subseteq V$ finite, $t \in (0,\infty)$, and let $x \in \mathcal{D}^U_{t-}$ be proper. Then the *jump characteristics* of x are the elements $\{(t_k(x), j_k(x), v_k(x))\} \subset (0,\infty) \times \mathcal{J} \times U$, where $\{t_k(x)\} := \mathrm{Disc}_\infty(x)$ is an increasing sequence, and $\Delta x_{v_k(x)}(t_k(x)) = j_k(x)$ for each $k < |\mathrm{Disc}_\infty(x)| + 1$. When the trajectory x is clear from the context, we simply write $\{(t_k, j_k, v_k)\}$ for $\{(t_k(x), j_k(x), v_k(x))\}$.

We now establish conditions under which the jump characteristics of a process exist.

Lemma 8.5. *[37] Suppose the jump rate functions $\{r^v_j\}_{v\in V, j\in\mathcal{J}}$ satisfy conditions (2) and (4) of Assumption 4. For any finite $W \subseteq V$, let $\widehat{X}^W$ be any weak solution to the SDE* (8.1.3) *for the initial data κ and let X be any weak solution to* (6.1.1) *for the same initial data κ. Then for any finite $U \subseteq V$, the jump characteristics of $\widehat{X}^W_U$ and X_U are almost surely well defined.*

Proof. Note that $\widehat{X}^\emptyset$ is a weak solution to (6.1.1) so we may write $X = \widehat{X}^\emptyset$ without loss of generality. Thus, it suffices to prove that the jump characteristics of $\widehat{X}^W_U$ exist. This occurs precisely when $\widehat{X}^W_U$ is proper and its discontinuity times can be listed in increasing order. However, $\widehat{X}^W$ (and therefore $\widehat{X}^W_U$) is a.s. proper by Lemma 6.3. Furthermore, because U is finite, $\mathcal{X}^U$ is discrete. Therefore, $|\mathrm{Disc}_t\left(\widehat{X}^W_U\right)| \leq \infty$ for every $t \in \mathbb{R}_+$. Thus, the jump characteristics $\{(t_k(\widehat{X}^W_U), j_k(\widehat{X}^W_U), v_k(\widehat{X}^W_U))\}$ are well defined because the elements of $\mathrm{Disc}_\infty\left(\widehat{X}^W_U\right)$ can be ordered. $\square$

Our next result is a general change of measure result that yields the Radon-Nikodym derivative of μ_{t-} with respect to $\widehat{\mu}^W_{t-}$ for $t > 0$ and $W \subseteq V$ finite in terms of the jump characteristics of $\widehat{X}^W_W$. Lemma 8.5 ensures that this Radon-Nikodym derivative is a.s. well defined. We defer the proof of the result to Section 8.2.2.

Proposition 8.6. *[37] Let G be a deterministic (not necessarily locally finite) graph. Suppose the jump rate functions $\{r_j^v\}_{v\in V, j\in\mathcal{J}}$ satisfy conditions (2) and (4) of Assumption 4. Let $W \subseteq V$ be any finite set and suppose that $\widehat{\mu}^W$ is the law of any weak solution $\widehat{X}^W$ to the SDE* (8.1.3) *for the initial data κ. For any $(t, j, v) \in \mathbb{R}_+ \times \mathcal{J} \times V$, assume that*

$$\int_0^t r_j^v(s, \kappa, \widehat{X}^W)\, ds < \infty \text{ a.s..} \tag{8.1.7}$$

Define $\mathbb{G} := \mathcal{H}^\kappa \vee \mathbb{H}^{\widehat{X}^W}$ and the process

$$L_t^W := \left[\prod_{0<t_k\le t} r_{j_k}^{v_k}(t_k, \kappa, \widehat{X}^W)\right] \exp\left(-\sum_{(j,v)\in\mathcal{J}\times W} \int_{(0,t]} \left(r_j^v(s, \kappa, \widehat{X}^W) - 1\right) ds\right), \tag{8.1.8}$$

where $\{(t_k, j_k, v_k)\}$ are the jump characteristics of $\widehat{X}_W^W$. Then $L^W := (L_t^W)_{t\in\mathbb{R}_+}$ is a.s. well defined and it is a $\mathbb{G}$-local martingale. If L^W is also a $\mathbb{G}$-martingale, then there exists a measure $\widetilde{\mu} \in \mathcal{P}((\mathcal{K}\times\mathcal{D})^V)$ such that $\frac{d\widetilde{\mu}_{t-}}{d\widehat{\mu}_{t-}^W}(\kappa, \widehat{X}^W[t]) = L_{t-}^W$ for all $t > 0$ a.s., and there exists a weak solution X to the SDE (6.1.1) *for the initial data κ such that $Law(\kappa, X) = \widetilde{\mu}$.*

Remark 8.7. [37] The conditions of Proposition 8.6 are not sufficient to guarantee well-posedness of (6.1.1) and (8.1.3). In the absence of well-posedness (and assuming L^W is a $\mathbb{G}$-Martingale), for each probability measure $\widehat{\mu}^W$ describing the law of a weak solution to (8.1.3), there exists a unique measure $\widetilde{\mu}$ such that $\frac{d\widetilde{\mu}_{t-}}{d\widehat{\mu}_{t-}^W}(\kappa, \widehat{X}^W[t]) = L_{t-}^W$ for each $t > 0$, and $\widetilde{\mu}$ is the law of a weak solution to (6.1.1).

Remark 8.8. [37] By allowing the jump rate functions to violate condition (1) of Assumption 4, the graph structure G becomes irrelevant to Proposition 8.6. Thus, Proposition 8.6 is applicable to solutions of a more general class of infinite-dimensional Poisson-driven SDEs not necessarily arising as locally interacting processes with respect to some graph.

When G is finite and the IPS is Markov, this result is known (e.g. [19, Example 15.2.10]). Even in the non-Markov case, the result follows by applying duality (discussed in Section 8.2.1) to characterize the IPS in terms of a point process and then applying any change of measure theorem for point processes (e.g. [19, Theorem 15.2.7], [28, Proposition 14.4.III] and [18, VIII T10]), along with a few simple bounding arguments to prove that the candidate Radon-Nikodym derivative is indeed a martingale (rather than a local martingale). However, the case of infinite graphs is more subtle due to the possibility of explosions, and Proposition 8.9 in fact addresses the more general case when the graph may not be locally finite, necessitating careful justification. The more general setting is of interest for the study of analogous properties of IPS on random graphs. The proof of Proposition 8.6 is given in Section 8.2.2. As a corollary, we now state the change of measure result that we use to prove Proposition 6.10.

Corollary 8.9. *[37] Suppose Assumption 4 holds and that G is a graph of bounded maximal degree. For any $n \in \mathbb{N}$ and $t \in (0,\infty)$,*

(8.1.9)
$$\frac{d\mu_{t-}}{d\widehat{\mu}^n_{t-}}(\kappa, \widehat{X}^n[t]) = \left[\prod_{0<t_k^n<t} r^{v_k^n}_{j_k^n}(t_k^n, \kappa, \widehat{X}^n)\right] \exp\left(-\sum_{(j,v)\in\mathcal{J}\times V_n^G} \int_{(0,t)} \left(r_j^v(s,\kappa,\widehat{X}^n) - 1\right) ds\right) a.s.,$$

where $\{(t_k^n, j_k^n, v_k^n)\}_k$ is the set of jump characteristics of $\widehat{X}^n_{V_n^G}$ in the sense of Definition 8.4.

Proof. Fix $n \in \mathbb{N}$, let d be the maximal degree of any vertex in G and recall $\widehat{X}^n = \widehat{X}^{V_n^G}$. Applying Proposition 8.6 with $W = V_n^G$, the process $L^n := L^{V_n^G}$ in (8.1.8) is a $\mathbb{G}$-local martingale. We first prove that it is in fact a $\mathbb{G}$-martingale. Let $\{\theta_\ell\}_{\ell\in\mathbb{N}}$ be a localizing sequence for L^n and for each $\ell \in \mathbb{N}$, let $\tau_\ell := \inf\{t : |\mathrm{Disc}_t\left(\widehat{X}^n_{V_n^G}\right)| \geq \ell\}$ where the infimum of an empty set is taken to be infinite. Note that τ_ℓ is a $\mathbb{G}$-stopping time. Because $\widehat{X}^n_{V_n^G}$ is a.s. càdlàg and $\mathcal{X}^{V_n^G}$ is a discrete space, $\widehat{X}^n_{V_n^G}$ can only have finitely many discontinuities in any finite time interval. Thus, $\lim_{\ell\to\infty}\tau_\ell = \infty$ a.s., so $\{\tau_\ell \wedge \theta_\ell\}_{\ell\in\mathbb{N}}$ is also a localizing sequence for L^n. Recall the definition of $C : \mathbb{N}\times\mathbb{R}_+ \to \mathbb{R}_+$ from condition (3) of Assumption 4. Then for each $t \in \mathbb{R}_+$,

$$\sup_{\ell\in\mathbb{N}} |L^n_{t\wedge\tau_\ell\wedge\theta_\ell}| \leq |C(d+1,t)|^{\ell \wedge \left|\mathrm{Disc}_t\left(\widehat{X}^n_{V_n^G}\right)\right|} \exp\left(t|V_n^G||\mathcal{J}|\right) \leq |C(d+1,t)|^{\left|\mathrm{Disc}_t\left(\widehat{X}^n_{V_n^G}\right)\right|} \exp\left(t|V_n^G||\mathcal{J}|\right).$$

However, note that $\left|\mathrm{Disc}_t\left(\widehat{X}^n_{V_n^G}\right)\right| \sim \mathrm{Poiss}(|V_n^G||\mathcal{J}|t)$ because $\widehat{r}_j^{v,V_n^G}(\cdot,\kappa,\widehat{X}^n) = 1$ whenever $(j,v) \in \mathcal{J}\times V_n^G$. This implies

$$\mathbb{E}\left[\sup_{\ell\in\mathbb{N}} |L^n_{t\wedge\tau_\ell\wedge\theta_\ell}|\right] \leq \exp\left(t|V_n^G||\mathcal{J}|\right) \mathbb{E}\left[|C(d+1,t)|^{\left|\mathrm{Disc}_t\left(\widehat{X}^n_{V_n^G}\right)\right|}\right] < \infty,$$

which shows that for each $t \geq 0$, the sequence $\{L^n_{t\wedge\tau_\ell\wedge\theta_\ell}\}_{\ell\in\mathbb{N}}$ is dominated by a random variable with a finite first moment and is therefore uniformly integrable. By [56, Proposition 1.8], it follows that L^n is a $\mathbb{G}$-martingale and for any $t \in \mathbb{R}_+$, $\mathbb{E}[L^n_t] = 1$.

Proposition 8.6 then implies that the measure $\widetilde{\mu} \in \mathcal{P}((\mathcal{K}\times\mathcal{D})^V)$ defined by $d\widetilde{\mu}_{t-}/d\widehat{\mu}^n_{t-}(\kappa,\widehat{X}^n) = L^n_{t-}$ a.s. for all $t > 0$ is the law of a weak solution $\widetilde{X}$ to (6.1.1) for the initial data κ on the graph

G. By Lemma 6.3, the SDE (6.1.1) is well-posed which implies that $(\kappa, \widetilde{X}) \stackrel{(d)}{=} (\kappa, X)$. This shows $d\mu_{t-}/d\widehat{\mu}^n_{t-}(\kappa, \widehat{X}^n) = d\widetilde{\mu}_{t-}/d\widehat{\mu}^n_{t-}(\kappa, \widehat{X}^n) = L^n_{t-}$ a.s. for every $t > 0$, as desired. $\square$

8.1.3 Proof of Proposition 6.10

As mentioned earlier, the proof of the MRF properties on finite time intervals proceeds by transfering the analogous conditional independence properties for the reference processes established in Proposition 8.2 to the original IPS. This makes use of a result from [81] paraphrased and slightly modified in Lemma 8.10 below.

Lemma 8.10. *[37] Let $(\widetilde{\Omega}, \widetilde{\mathcal{F}})$ be a measurable space, let $\widetilde{\mathcal{F}}_i \subset \widetilde{\mathcal{F}}$, $i = 0, 1, 2$, be sub-σ-algebras, and let $\widetilde{\mathbb{P}}_0$ and $\widetilde{\mathbb{P}}_1$ be two probability measures on $(\widetilde{\Omega}, \widetilde{\mathcal{F}})$ such that $\widetilde{\mathbb{P}}_1 \ll \widetilde{\mathbb{P}}_0$. Assume that $\widetilde{\mathcal{F}}_1$ and $\widetilde{\mathcal{F}}_2$ are conditionally independent with respect to $\widetilde{\mathbb{P}}_0$ given $\widetilde{\mathcal{F}}_0$. If in addition ρ, the Radon-Nikodym derivative $d\widetilde{\mathbb{P}}_1/d\widetilde{\mathbb{P}}_0$ with respect to $\vee_{i=1}^3 \widetilde{\mathbb{F}}_i$, satisfies $\rho = \rho_1\rho_2$ almost surely for $\widetilde{\mathcal{F}}_i \vee \widetilde{\mathcal{F}}_0$-measurable random variables ρ_i, $i = 1, 2$, then $\widetilde{\mathcal{F}}_1$ and $\widetilde{\mathcal{F}}_2$ are also conditionally independent with respect to $\widetilde{\mathbb{P}}_1$ given $\widetilde{\mathcal{F}}_0$.*

Proof. If the filtrations were required to be complete, then this result would follow from [81, Theorem 3.6] under the stronger assumption that $\widetilde{\mathbb{P}}_1$ is equivalent to $\widetilde{\mathbb{P}}_0$. However, as elaborated below, the same argument used in [81] also shows that the result holds under the weaker assumptions of a not necessarily complete filtration and under absolute continuity $\widetilde{\mathbb{P}}_1 \ll \widetilde{\mathbb{P}}_0$ (rather than equivalence).

Let Z be any bounded, $\widetilde{\mathcal{F}}_1$-measurable random variable. Then, applying [14, Proposition B.41], the $\widetilde{\mathbb{P}}_0$-conditional independence of $\widetilde{\mathcal{F}}_1$ and $\widetilde{\mathcal{F}}_2$ given $\widetilde{\mathcal{F}}_0$ and the $\widetilde{\mathcal{F}}_i$-measurability of ρ_i, $i = 1, 2$, we have $\widetilde{\mathbb{P}}_1$-a.s.,

$$\mathbb{E}^{\widetilde{\mathbb{P}}_1}\left[Z|\widetilde{\mathcal{F}}_0 \vee \widetilde{\mathcal{F}}_2\right] = \frac{\mathbb{E}^{\widetilde{\mathbb{P}}_0}\left[Z\rho_1\rho_2|\widetilde{\mathcal{F}}_0 \vee \widetilde{\mathcal{F}}_2\right]}{\mathbb{E}^{\widetilde{\mathbb{P}}_0}\left[\rho_1\rho_2|\widetilde{\mathcal{F}}_0 \vee \widetilde{\mathcal{F}}_2\right]} = \frac{\rho_2\mathbb{E}^{\widetilde{\mathbb{P}}_0}\left[Z\rho_1|\widetilde{\mathcal{F}}_0 \vee \widetilde{\mathcal{F}}_2\right]}{\rho_2\mathbb{E}^{\widetilde{\mathbb{P}}_0}\left[\rho_1|\widetilde{\mathcal{F}}_0 \vee \widetilde{\mathcal{F}}_2\right]} = \frac{\mathbb{E}^{\widetilde{\mathbb{P}}_0}\left[Z\rho_1|\widetilde{\mathcal{F}}_0\right]}{\mathbb{E}^{\widetilde{\mathbb{P}}_0}\left[\rho_1|\widetilde{\mathcal{F}}_0\right]}.$$

Thus, $\mathbb{E}^{\widetilde{\mathbb{P}}_1}\left[Z|\widetilde{\mathcal{F}}_0 \vee \widetilde{\mathcal{F}}_2\right]$ is $\widetilde{\mathcal{F}}_0$-measurable and therefore $\widetilde{\mathbb{P}}_1$-a.s. equal to $\mathbb{E}^{\widetilde{\mathbb{P}}_1}\left[Z|\widetilde{\mathcal{F}}_0\right]$, so $\widetilde{\mathcal{F}}_1$ and $\widetilde{\mathcal{F}}_2$ are conditionally independent with respect to $\widetilde{\mathbb{P}}_1$ given $\widetilde{\mathcal{F}}_0$. $\square$

Proof of Proposition 6.10: First note that for any $t \geq 0$ μ, μ_t and μ_{t-} (if $t > 0$) are all automorphism invariant by Lemma 6.3.

Fix $n \in \mathbb{N}$. By Lemma 6.3, the SDEs (6.1.1) and (8.1.3) (the latter with $W = V_n^G$) which define X and $\widehat{X}^n$, respectively, are both strongly well-posed. Suppose that $A, B, S \subseteq V$ partition V with $|S| < \infty$ and let n be sufficiently large so that $S = \mathcal{N}_A^2 \subseteq V_{n-1}^G$. Then by assumption, we have $\kappa_A \perp\!\!\!\perp \kappa_B | \kappa_S$ and so if we fix $t > 0$, it follows from Proposition 8.2 that $\widetilde{\mathcal{F}}_1 := \mathcal{B}((\mathcal{K} \times \mathcal{D}_{t-})^A)$ is independent of $\widetilde{\mathcal{F}}_2 := \mathcal{B}((\mathcal{K} \times \mathcal{D}_{t-})^B)$ given $\widetilde{\mathcal{F}}_0 := \mathcal{B}((\mathcal{K} \times \mathcal{D}_{t-})^S)$ with respect to $\widehat{\mu}_{t-}^n$. We would like to apply Lemma 8.10 with $(\widetilde{\Omega}, \widetilde{\mathcal{F}}) = \Big((\mathcal{K} \times \mathcal{D}_{t-})^V, \mathcal{B}((\mathcal{K} \times \mathcal{D}_{t-})^V)\Big)$, $\widetilde{\mathcal{F}}_i$, $i = 0, 1, 2$ as just defined, $\widetilde{\mathbb{P}}_0 = \widehat{\mu}_{t-}^n = \text{Law}(\kappa, \widehat{X}^n[t])$, $\widetilde{\mathbb{P}}_1 = \mu_{t-} = \text{Law}(\kappa, X[t])$, $\rho = d\mu_{t-}/d\widehat{\mu}_{t-}^n$. To verify the remaining conditions of the lemma, note that by Corollary 8.9, $\widetilde{\mathbb{P}}_1 \ll \widetilde{\mathbb{P}}_0$ on $\mathcal{B}((\mathcal{K} \times \mathcal{D}_{t-})^V)$ with Radon-Nikodym derivative $d\mu_{t-}/d\widehat{\mu}_{t-}^n = \rho$.

We now verify the last remaining condition of Lemma 8.10, which involves showing that the Radon-Nikodym derivative $\rho = d\mu_{t-}/d\widehat{\mu}_{t-}^n$ factorizes in the right way. To this end, recall the definition of $\{\bar{r}_j^v\}_{v \in V, j \in \mathcal{J}}$ from condition (1) of Assumption 4, and for $v \in V$, define $g_v : (\mathcal{K} \times \mathcal{D}_{t-})^{\text{cl}_v(G)} \to \mathbb{R}_+$ as follows:

$$g_v(\vartheta, x) := \left[\prod_{0 < t_k(x_v) < t} \bar{r}_{j_k(x_v)}^v(t_k(x_v), \vartheta, x) \right] \exp\left(- \sum_{j \in \mathcal{J}} \int_{(0,t)} (\bar{r}_j^v(t, \vartheta, x) - 1)\, ds \right)$$

if x is proper, and set $g_v(\vartheta, x) \equiv 0$ (any arbitrary value would have sufficed) for any other $x \in \mathcal{D}_{t-}$. Noting that the jump characteristics $\widehat{X}_{V_n^G}^n$ are a.s. well defined by Lemma 8.5, and letting $\{(t_k^n, j_k^n, v_k^n)\}_{k \in \mathbb{N}}$ denote the associated jump characteristics, Corollary 8.9 and the locality property of Assumption 4 imply

$$\frac{d\mu_{t-}}{d\widehat{\mu}_{t-}^n} \left(\kappa, \widehat{X}^n[t] \right) = \prod_{v \in V_n^G} g_v \left(\kappa_{\text{cl}_v}, \widehat{X}_{\text{cl}_v}^n[t] \right) \text{ a.s..} \tag{8.1.10}$$

This implies that $\frac{d\mu_{t-}}{d\widehat{\mu}_{t-}^n} : (\mathcal{K} \times \mathcal{D}_{t-})^V \to \mathbb{R}_+$ may be defined by $\frac{d\mu_{t-}}{d\widehat{\mu}_{t-}^n}(\vartheta, x) := \prod_{v \in V_n^G} g_v(\vartheta_{\text{cl}_v}, x_{\text{cl}_v})$. Thus, for each $(\vartheta, x) \in (\mathcal{K} \times \mathcal{D}_{t-})^V$, we can rewrite the Radon-Nikodym derivative as

$$\frac{d\mu_{t-}}{d\widehat{\mu}_{t-}^n} (\vartheta, x) = \left(\prod_{v \in \text{cl}_A \cap V_n^G} g_v \left(\vartheta_{\text{cl}_v}, x_{\text{cl}_v} \right) \right) \left(\prod_{u \in V_n^G \setminus \text{cl}_A} g_u \left(\vartheta_{\text{cl}_u}, x_{\text{cl}_u} \right) \right). \tag{8.1.11}$$

Since $\mathcal{N}_A^2(G) = S$, the term in the first bracket on the right-hand side of (8.1.11) depends only on $(\kappa_{A \cup S}, \widehat{X}_{A \cup S}^n[t])$ and is thus $\mathcal{B}((\mathcal{K} \times \mathcal{D}_{t-})^{A \cup S} = \widetilde{\mathcal{F}}_1 \vee \widetilde{\mathcal{F}}_0$-measurable while the term in the second bracket depends only on $(\kappa_{B \cup S}, \widehat{X}_{B \cup S}^n[t])$ and is thus $\mathcal{B}((\mathcal{K} \times \mathcal{D}_{t-})^{B \cup S} = \widetilde{\mathcal{F}}_2 \vee \widetilde{\mathcal{F}}_0$-measurable. This completes the verification of all the conditions of Lemma 8.10, which allows us to conclude that

$$(\kappa_A, X_A[t)) \perp\!\!\!\perp (\kappa_B, X_B[t)) | (\kappa_S, X_S[t)). \tag{8.1.12}$$

Thus, we have shown that μ_{t-} is a 2-SGMRF with respect to G. Next, note that for any $t > 0$, $X(t-) = X(t)$ almost surely. Thus, for any $U \subseteq V$, $\mathcal{H}^{\kappa_U, X_U[t)} = \mathcal{H}^{\kappa_U, X_U[t]}$, and by Lemma C.2, $(\kappa_A, X_A[t)) \perp\!\!\!\perp (\kappa_B, X_B[t)) | (\kappa_S, X_S[t))$ implies that $\mathcal{H}^{\kappa_A, X_A[t)} \perp\!\!\!\perp \mathcal{H}^{\kappa_B, X_B[t)} | \mathcal{H}^{\kappa_S, X_S[t)}$ which we have just shown implies $\mathcal{H}^{\kappa_A, X_A[t]} \perp\!\!\!\perp \mathcal{H}^{\kappa_B, X_B[t]} | \mathcal{H}^{\kappa_S, X_S[t]}$. Applying Lemma C.2 once more yields,

$$(\kappa_A, X_A[t]) \perp\!\!\!\perp (\kappa_B, X_B[t]) | (\kappa_S, X_S[t]). \tag{8.1.13}$$

So μ_t is also a 2-SGMRF.

We now extend this to the infinite time interval to show that the same is true for μ. The argument we use is similar to the one applied in the proof of [59, Theorem 2.4]. Fix an element $a \notin \mathcal{X}$. For any $x \in \mathcal{D}$ and $t \in \mathbb{R}_+$, we may embed the truncated function $x[t)$ in $\mathcal{D}^a := \mathcal{D}(\mathbb{R}_+, \mathcal{X} \cup \{a\})$ by setting $x[t)(s) = a$ whenever $s \geq t$ and $x[t)(s) = x(s)$ when $s < t$. Thus, $x[t) \to x$ as $t \to \infty$ in $\mathcal{D}^a$. Let $A, B, S \subseteq V$ be a partition such that $S = \mathcal{N}_A^2$ is finite. For each $U \subset V$, define $f_U : (\mathcal{K} \times \mathcal{D}^a)^U \to \mathbb{R}$ to be bounded and continuous. Then, as justified below the display, we have

$$\begin{aligned}
&\mathbb{E}\left[f_A(\kappa_A, X_A) f_B(\kappa_B, X_B) f_S(\kappa_S, X_S)\right] \\
&\quad = \lim_{s_2 \to \infty} \lim_{s_1 \to \infty} \mathbb{E}\left[f_A(\kappa_A, X_A[s_2)) f_B(\kappa_B, X_B[s_2)) f_S(\kappa_S, X_S[s_1))\right] \\
&\quad = \lim_{s_2 \to \infty} \lim_{s_1 \to \infty} \mathbb{E}\left[\mathbb{E}\left[f_A(\kappa_A, X_A[s_2)) | \kappa_S, X_S[s_1)\right] \mathbb{E}\left[f_B(\kappa_B, X_B[s_2)) | \kappa_S, X_S[s_1)\right] f_S(\kappa_S, X_S[s_1))\right] \\
&\quad = \lim_{s_2 \to \infty} \mathbb{E}\left[\mathbb{E}\left[f_A(\kappa_A, X_A[s_2)) | \kappa_S, X_S\right] \mathbb{E}\left[f_B(\kappa_B, X_B[s_2)) | \kappa_S, X_S\right] f_S(\kappa_S, X_S)\right] \\
&\quad = \mathbb{E}\left[\mathbb{E}\left[f_A(\kappa_A, X_A) | \kappa_S, X_S\right] \mathbb{E}\left[f_B(\kappa_B, X_B) | \kappa_S, X_S\right] f_S(\kappa_S, X_S)\right],
\end{aligned}$$

where the first equality uses the bounded convergence theorem and continuity of the functions f_A, f_B, f_S, the second equality uses equation (8.1.12) with $t = s_1$, the third equality uses $\sigma(\kappa_S, X_S) = \vee_{s \in \mathbb{R}_+} \sigma(\kappa_S, X_S[s))$, the Doob martingale convergence theorem, the continuity of f_S and the bounded convergence theorem, and the final equality holds due to the bounded convergence theorem and the continuity of f_A and f_B. This proves μ is a 2-SGMRF with respect to G, as desired. ∎

8.2 Proof of the Radon-Nikodym Derivative Characterization

The goal of this section is to prove the change of measure result in Proposition 8.6. The proof, which is given in Section 8.2.2, relies on a key duality characterization of the IPS that is first established in Section 8.2.1 (see Proposition 8.12 therein).

8.2.1 Dual Processes

We start in Section 8.2.1 by introducing some standard notation related to point processes.

Point Processes

Fix a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ supporting a filtration $\mathbb{G} \subseteq \mathbb{F}$ and a Polish space $\mathcal{Z}$. Recall from Section 1.4 that $\mathscr{N}(\mathcal{Z})$ is the (Polish) space of locally finite, non-negative integer-valued measures on $\mathcal{Z}$. A point process P on $\mathcal{Z}$ is a random element taking values in $\mathscr{N}(\mathcal{Z})$. For every bounded set $B \subseteq \mathcal{Z}$, there exists a finite set of points $\{z_i\}_{i=1}^N \subseteq B$ such that $P(\{z_i\}) > 0$ for all $i = 1, \dots, N$, and $P(B \setminus \{z_i\}_{i=1}^N) = 0$. These points are called *events*. A point process P on $\mathcal{Z}$ is *simple* if $\sup_{z \in \mathcal{Z}} P(\{z\}) \in \{0, 1\}$ almost surely. For $\widehat{\mathcal{Z}} := \mathbb{R}_+ \times \mathcal{Z}$, a point process P on $\widehat{\mathcal{Z}}$ is referred to as a *marked* point process on $\mathbb{R}_+$ with marks in $\mathcal{Z}$. If P has events $\{(t_i, \rho_i)\}_{i=1}^N \subset \widehat{\mathcal{Z}}$, then $\{\rho_i\}_{i=1}^N$ are said to be the *marks* of P. If $P([0, T] \times \mathcal{Z}) < \infty$ a.s. for all $T \in \mathbb{R}_+$, then P is said to be *non-explosive* or *locally finite*. All point processes P that we consider will be *simple* in the sense that $\sup_{t \in \mathbb{R}_+} P(\{t\} \times \mathcal{Z}) \in \{0, 1\}$ almost surely. A marked point process P on $\widehat{\mathcal{Z}}$ is said to be $\mathbb{G}$-adapted if for every $t \in \mathbb{R}_+$ and $A \in \mathcal{B}([0, t] \times \mathcal{Z})$, $P(A)$ is $\mathcal{G}_t$-measurable. For any point process P on $\widehat{\mathcal{Z}}$, $\mathbb{H}^P$ is defined to be the minimal filtration satisfying the usual conditions such that P is $\mathbb{H}^P$-adapted.

Suppose that $\mathcal{Z}$ is equipped with a non-negative, locally finite *reference measure* ℓ. Then a random function $\Gamma : \Omega \times \mathbb{R}_+ \times \mathcal{Z} \to \mathbb{R}_+$ is said to be $\mathbb{G}$-*mark predictable* if it is measurable with respect to $\mathscr{P}(\mathbb{G}) \otimes \mathcal{B}(\mathcal{Z})$, where $\mathscr{P}(\mathbb{G})$ is the predictable σ-algebra generated by $\mathbb{G}$ [28, page 379]. As an immediate consequence, it follows that $t \mapsto \int_{z \in A} \Gamma(t, z)\, \ell(dz)$ is $\mathbb{G}$-predictable for all $A \in \mathcal{B}(\mathcal{Z})$. The $\mathbb{G}$-*intensity* of P with respect to the measure ℓ is a $\mathbb{G}$-mark predictable process Γ such that for each bounded $A \subseteq \mathcal{Z}$, the process $t \mapsto P([0, t] \times A) - \int_{z \in A} \int_0^t \Gamma(s, z)\, ds\, \ell(dz)$ is a $\mathbb{G}$-local martingale [28, Definitions 14.1.I,14.3.I].

Duality

Recalling the Definition 8.4 of the jump characteristics $\{(t_k(x), j_k(x), v_k(x))\}$ of any proper trajectory x, we now define the notion of a dual to x.

Definition 8.11. [37] For any (not necessarily finite) $U \subseteq V$, and any proper $x \in \mathcal{D}^U$, the *dual* of x is the simple marked point process $p \in \mathscr{N}(\mathbb{R}_+ \times \mathcal{J} \times U)$ with event times $\bigcup_{v\in U}\{t_k(x_v)\}$ and marks $\bigcup_{v\in U}\{(j_k(x_v), v_k(x_v))\}$.

In what follows, $\#_{\mathcal{S}}$ denotes the counting measure on any countable set $\mathcal{S}$.

Proposition 8.12. *[37] Suppose that $G = (V, E)$ is a deterministic (but not necessarily locally finite) graph, and the jump rate functions $\{r_j^v\}_{v\in V, j\in\mathcal{J}}$ satisfy* (8.1.7) *and conditions (2) and (4) of Assumption 4. Let X be any weak solution to the SDE* (6.1.1) *for given initial data κ on the solution space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, and define $\mathbb{G} := \mathcal{H}^\kappa \vee \mathbb{H}^X \subseteq \mathbb{F}$. Then for any deterministic (not necessarily finite) subset $U \subseteq V$, the following properties are satisfied:*

(a) The dual P_U of X_U has a $\mathbb{G}$-intensity on $\mathcal{J} \times U$ with respect to the reference measure $\#_{\mathcal{J}\times U}$ that is given explicitly by

$$\Lambda_U(t,\rho) := r_j^v(t,\kappa,X) \quad \text{for} \quad \rho = (j,v) \in \mathcal{J}\times U,\ t > 0. \tag{8.2.1}$$

(b) Let $\widetilde{X}$ be an a.s. proper $\mathcal{X}^V$-valued càdlàg process and let $\widetilde{\kappa}$ be a $\mathcal{K}^V$-valued random element, both defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and such that $\widetilde{X}_v(0) = \varsigma(\widetilde{\kappa}_v)$ for each $v \in V$. Set $\widetilde{\mathbb{G}} := \mathcal{H}^{\widetilde{\kappa}} \vee \mathbb{H}^{\widetilde{X}}$. If the dual of $\widetilde{X}$ is a simple marked point process $\widetilde{P}$ with mark space $\mathcal{J} \times U$ and $\widetilde{\mathbb{G}}$-predictable intensity $\widetilde{\Lambda}$ with respect to the reference measure $\#_{\mathcal{J}\times U}$ given by

$$\widetilde{\Lambda}(t,\rho) := r_j^v(t,\widetilde{X},\widetilde{\kappa}) \quad \text{for} \quad \rho = (j,v) \in \mathcal{J}\times U,\ t > 0, \tag{8.2.2}$$

then it is possible to extend the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to support a filtration $\widehat{\mathbb{G}} \supseteq \widetilde{\mathbb{G}}$ satisfying the usual conditions and a collection of i.i.d. $\widehat{\mathbb{G}}$-Poisson processes $\widetilde{\mathbf{N}} := \{\widetilde{\mathbf{N}}_v\}_{v\in U}$ on $\mathbb{R}^2 \times \mathcal{J}$ with intensity measure $\mathrm{Leb}^2 \times \#_{\mathcal{J}}$ *such that $\widetilde{X}$ satisfies*

$$\widetilde{X}_v(t) = \varsigma(\widetilde{\kappa}_v) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r\le r_j^v(s,\widetilde{X},\widetilde{\kappa})\right\}}\,\widetilde{\mathbf{N}}_v(ds,dr,dj) \text{ for } v\in U. \tag{8.2.3}$$

Remark 8.13. [37] When G is finite, similar duality results are well known for countable state

Markov chains (see [28, Example 10.3(a)]), and extensions to the general (non-Markovian) setting considered here can also be easily deduced by applying Poisson embedding theorems such as [19, Theorems 15.3.3 and 15.3.4] (which would require a few simple scaling arguments to account for the slightly more restrictive assumptions [19] places on the driving Poisson processes, see the proof of Proposition 8.12(a)). However, when G is infinite, a key challenge is that the dual point process (of X and $\widehat{X}$) can become explosive in the sense that it may have infinitely many events on any finite interval, and so the results from [19] cannot be directly applied. In this case, our proofs proceed by first establishing duality for X_W over all $W \subseteq U$ such that W is finite and then extending the results to the (possibly) infinite subset U.

Proof of Proposition 8.12: **Proof of (a):** Define $\overline{\mathbb{G}} := \mathbb{G} \vee \mathbb{H}^{\mathbf{N}_U}$. Let $W \subseteq U$ be finite. Define the point process $\overline{\mathbf{N}}^W$ on $(\mathbb{R}_+^2 \times \mathcal{J} \times W)$ by

$$\overline{\mathbf{N}}^W(\{(t,r,j,v)\}) = \mathbf{N}_v\left(\left\{\left(t, \frac{r}{|\mathcal{J}||W|}, j\right)\right\}\right), \quad (t,r,j,v) \in \mathbb{R}_+^2 \times \mathcal{J} \times W.$$

Since $\{\mathbf{N}_v\}_{v \in W}$ are i.i.d. $\mathbb{F}$-Poisson processes and they are $\overline{\mathbb{G}}$-adapted (by definition), it follows that they are also i.i.d. $\overline{\mathbb{G}}$-Poisson processes. Thus $\overline{\mathbf{N}}^W$ is also a $\overline{\mathbb{G}}$-Poisson process on $\mathbb{R}_+^2 \times \mathcal{J} \times W$. Let $Q_W = \#_{\mathcal{J}\times W}/(|\mathcal{J}||W|)$ be the uniform distribution on $\mathcal{J} \times W$. For any $\overline{A} \in \mathcal{B}(\mathbb{R}_+^2)$ and $B \subseteq \mathcal{J} \times W$, let $\overline{A}' := \left\{\left(t, \frac{r}{|\mathcal{J}||W|}\right) : (t,r) \in \overline{A}\right\}$ and define $B_v := \{j \in \mathcal{J} : (j,v) \in B\}$ so that $B = \cup_{v \in W} B_v \times \{v\}$. For every $\overline{A} \in \mathcal{B}(\mathbb{R}_+^2)$,

$$\mathbb{E}\left[\overline{\mathbf{N}}^W(\overline{A} \times B)\right] = \mathbb{E}\left[\sum_{(j,v)\in B} \mathbf{N}_v(\overline{A}' \times \{j\})\right] = \frac{\mathrm{Leb}^2(\overline{A})|B|}{|\mathcal{J}||W|} = \mathrm{Leb}^2(\overline{A}) Q_W(B),$$

which implies that $\overline{\mathbf{N}}^W$ has intensity measure $\mathrm{Leb}^2 \otimes Q_W$. Then note that for any $A \in \mathcal{B}(\mathbb{R}_+)$ and $B \subseteq \mathcal{J} \times W$,

$$P_W(A \times B) = \int_{\mathbb{R}_+^2 \times \mathcal{J} \times W} \mathbb{I}_{\{s \in A\}} \mathbb{I}_{\{(j,v)\in B\}} \mathbb{I}_{\{r \le |W||\mathcal{J}| r_j^v(s,\kappa,X)\}} \overline{\mathbf{N}}^W(ds, dr, dj, dv).$$

Since this form coincides with Equation (15.45) of [19], by [19, Theorem 15.3.3], P_W has $\overline{\mathbb{G}}$-intensity,

$$\Lambda_W(t,j,v) = |\mathcal{J}||W|r_j^v(t,\kappa,X)Q_W(\{(j,v)\}) = \frac{|\mathcal{J}||W|r_j^v(t,\kappa,X)}{|\mathcal{J}||W|} = r_j^v(t,\kappa,X).$$

Given that $P_U|_{\mathbb{R}_+\times\mathcal{J}\times W} = P_W$ for every finite $W \subseteq U$, it follows that P_U has $\overline{\mathbb{G}}$-intensity Λ where for every $(t,j,v) \in \mathbb{R}_+ \times \mathcal{J} \times U$,

$$\Lambda_U(t,j,v) = \Lambda_{\{v\}}(t,j,v) = r_j^v(t,\kappa,X).$$

By condition (4) of Assumption 4, for each $(j,v) \in \mathcal{J} \times U$, $\Lambda_U(\cdot,j,v)$ has a.s. càglàd trajectories and by condition (2) of Assumption 4 it is $\mathbb{G}$-adapted. Thus, Λ_U is $\mathbb{G}$-mark predictable (in the sense defined in Section 8.2.1) and hence, Λ_U is also the $\mathbb{G}$-intensity of P_U.

Proof of (b): The proof proceeds via an explicit construction of a Poisson embedding of $\widetilde{P}$, that is, an extension of the probability space supporting an i.i.d. collection of Poisson processes $\widetilde{\mathbf{N}} := \{\widetilde{\mathbf{N}}_v\}_{v\in U}$ for which (8.2.3) holds. Despite the explosive nature of $\widetilde{P}$ note that for every finite $W \subset U$ and $T \in (0,\infty)$, $\mathrm{Disc}_T\left(\widetilde{X}_W\right)$ must be a.s. finite (since $\widetilde{X}_W$ is càdlàg and $\mathcal{X}$ is discrete). Thus, $\widetilde{P}_W := \widetilde{P}|_{\mathbb{R}_+\times\mathcal{J}\times W}$, which is the dual of $\widetilde{X}_W$, must be non-explosive. We may then use more standard arguments to show that $\widetilde{\mathbf{N}}_v$, $v \in W$, (which may be viewed as a Poisson embedding of $\widetilde{P}_W$) are i.i.d. driving Poisson processes. Because this holds for all finite $W \subseteq U$, this implies that $\widetilde{\mathbf{N}}_v$, $v \in U$, are also i.i.d. driving Poisson processes. We finish by proving that $\widetilde{X}$ and $\widetilde{\mathbf{N}}$ satisfy (8.2.3).

Our construction of $\{\widetilde{\mathbf{N}}_v\}_{v\in U}$ is based on a standard construction of Poisson embeddings of non-explosive point processes that can be found in numerous texts that discuss Poisson embeddings (e.g. see the proofs of [19, Theorem 15.3.4], [28, Proposition 14.7.1(b)] and [50, pages 469-478] as well as the statement of [20, Lemma 4]). However, we were unable to find any accessible and rigorous proofs that such constructions are, in fact, Poisson embeddings of marked point processes, so we include the full proof here for completeness. Our proof that, for finite $W \subseteq U$, $\{\widetilde{\mathbf{N}}_v\}_{v\in W}$ is a collection of i.i.d. Poisson processes closely follows the proof of [19, Theorem 15.3.4] which is stated for marked point processes but proven in the specific case of unmarked point processes.

Let $\widehat{\mathbf{N}} := \{\widehat{\mathbf{N}}_v\}_{v\in U}$ be a collection of i.i.d. Poisson processes independent of $\widetilde{\mathcal{G}}_\infty$ with intensity measure $\mathrm{Leb}^2 \otimes \#_{\mathcal{J}}$. Additionally, define the collection of i.i.d. uniform $[0,1]$-random variables $\{R_k^v\}_{k\in\mathbb{N},v\in U}$ to be independent of $\widetilde{\mathcal{G}}_\infty$ and $\widehat{\mathbf{N}}$. Finally, for each $v \in U$, let $\{(t_k^v, j_k^v)\}_{k\in\mathbb{N}}$ be the collection of events in $\widetilde{P}(dt, dj, \{v\})$ ordered so that $\{t_k^v\}$ is strictly increasing. In the event that $\widetilde{P}(\mathbb{R}_+ \times \mathcal{J} \times \{v\}) := K < \infty$, we say that $t_k^v = \infty$ for $k > K$. Note that $\{t_k^v\}$ is well defined as a consequence of Lemma 8.5. Thus, $\widetilde{P}(\cdot \times \mathcal{J} \times \{v\})$ must also have finitely many events in any finite interval. Then we can define a point process $\widetilde{\mathbf{N}}$ (which we will later show is a Poisson point process) by,

$$
\begin{aligned}
\widetilde{\mathbf{N}}(A\times B) &:= \sum_{v\in U} \widetilde{\mathbf{N}}_v(A_v \times B_v) \\
&:= \sum_{v\in U} \int_{A_v\times B_v} \mathbb{I}_{\left\{r > r_j^v(s,\widetilde{X},\widetilde{\kappa})\right\}} \widehat{\mathbf{N}}_v(ds,dr,dj) + \sum_{v\in U}\sum_{k\in\mathbb{N}} \mathbb{I}_{\left\{\left(t_k^v, R_k^v r^v_{j_k^v}(t_k^v,\widetilde{X},\widetilde{\kappa}), j_k^v\right)\in A_v\times B_v\right\}},
\end{aligned} \tag{8.2.4}
$$

where $A\times B := \cup_{v\in U} A_v \times B_v \times \{v\}$, and $\{A_v\}_{v\in U}$ and $\{B_v\}_{v\in U}$ are finite collections of Borel-measurable subsets of $\mathbb{R}_+^2$ and $\mathcal{J}$, respectively. In the event that $t_k^v = \infty$ for some $v \in U$ and $k \in \mathbb{N}$, $\{t_k^v\} \times \mathbb{R}_+ \cap A_v = \emptyset$ because $A_v \in \mathcal{B}(\mathbb{R}_+^2)$, so (8.2.4) is still well defined. Next, let $P^{R,J}$ be the point process on $\mathbb{R}_+ \times [0,1] \times \mathcal{J} \times U$ with events $\{(t_k^v, R_k^v, j_k^v, v)\}_{k\in\mathbb{N},v\in U}$ and define

$$
\widehat{\mathbb{G}} := \widetilde{\mathbb{G}} \vee \mathbb{H}^{\widehat{\mathbf{N}}} \vee \mathbb{H}^{P^{R,J}}.
$$

Note that $\widehat{\mathbb{G}}$ satisfies the usual conditions. Fix any finite $W \subseteq U$ and let $\widetilde{\mathbf{N}}_W = \widetilde{\mathbf{N}}|_{\mathbb{R}_+^2\times\mathcal{J}\times W}$. Let $H : \mathbb{R}_+^2 \times \mathcal{J} \times W \to \mathbb{R}_+$ be a non-negative, left-continuous (in its first input), $\widehat{\mathbb{G}}$-mark predictable random function. Then note that the process $(t,r,j,v) \mapsto \mathbb{I}_{\left\{r>r_j^v(t,\widetilde{X},\widetilde{\kappa})\right\}} H(t,r,j,v)$ is also $\widehat{\mathbb{G}}$-mark predictable. Lastly, note that because $\widehat{\mathbf{N}}_W := \{\widehat{\mathbf{N}}_v\}_{v\in W}$ is independent of $\widetilde{\mathcal{G}}_\infty$ and $\{R_k^v\}_{v,k}$, it is also a collection of i.i.d. $\widehat{\mathbb{G}}$-Poisson processes with intensity measure $\mathrm{Leb}^2 \otimes \#_{\mathcal{J}}$. Then by [19, Theorem 15.1.22], it follows that

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{v\in W} \int_{\mathbb{R}_+^2\times\mathcal{J}} \mathbb{I}_{\left\{r>r_j^v(s,\widetilde{X},\widetilde{\kappa})\right\}} H(s,r,j,v)\, \widehat{\mathbf{N}}_v(ds,dr,dj)\right] \\
&\qquad = \mathbb{E}\left[\int_{\mathbb{R}_+^2\times\mathcal{J}\times W} \mathbb{I}_{\left\{r>r_j^v(s,\widetilde{X},\widetilde{\kappa})\right\}} H(s,r,j,v)\, ds\, dr\, \#_{\mathcal{J}\times W}(dj,dv)\right].
\end{aligned} \tag{8.2.5}
$$

Let $P_W^{R,J} := P^{R,J}|_{\mathbb{R}_+\times[0,1]\times\mathcal{J}\times W}$ and note that $P_W^{R,J}$ is a $\widehat{\mathbb{G}}$-adapted non-explosive point process.

Furthermore, because $P_W^{R,J}(\cdot,[0,1],\cdot,\cdot) = \widetilde{P}_W(\cdot,\cdot,\cdot)$ and $\{R_j^v\}_{v\in U, j\in\mathcal{J}}$ are i.i.d. and independent of $\widetilde{P}_W$ with density 1, $P_W^{R,J}$ has $\widehat{\mathbb{G}}$-intensity $\Lambda^{R,J}(t,\rho,j,v) := r_j^v(t,\widetilde{X},\widetilde{\kappa})$ with respect to the reference measure $\mathrm{Leb}\otimes\mathrm{Leb}|_{[0,1]}\otimes\#_{\mathcal{J}\times W}$. Then the definition of $P_W^{R,J}$ and [19, Theorem 15.1.22] imply

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{k\in\mathbb{N},v\in W} H(t_k^v, R_k^v r_{j_k^v}^v(t_k^v,\widetilde{X},\widetilde{\kappa}), j_k^v, v)\right]\\
&\qquad = \mathbb{E}\left[\int_{\mathbb{R}_+\times(0,1]\times\mathcal{J}\times W} H(s,\rho r_j^v(s,\widetilde{X},\widetilde{\kappa}),j,v)\, P^{R,J}(ds,d\rho,dj,dv)\right]\\
&\qquad = \mathbb{E}\left[\int_{\mathbb{R}_+\times(0,1]\times\mathcal{J}\times W} H(s,\rho r_j^v(s,\widetilde{X},\widetilde{\kappa}),j,v) r_j^v(s,\widetilde{X},\widetilde{\kappa})\, ds\, d\rho\, \#_{\mathcal{J}\times W}(dj,dv)\right]\\
&\qquad = \mathbb{E}\left[\int_{\mathbb{R}_+^2\times\mathcal{J}\times W} \mathbb{I}_{\{r\le r_j^v(s,\widetilde{X},\widetilde{\kappa})\}} H(s,r,j,v)\, ds\, dr\, \#_{\mathcal{J}\times W}(dj,dv)\right],
\end{aligned}
\tag{8.2.6}
$$

where the last equality uses the change of variables $r = \rho r_j^v(s,\widetilde{X},\widetilde{\kappa})$. When combined, (8.2.4), (8.2.5) and (8.2.6) together yield

$$
\begin{aligned}
&\mathbb{E}\left[\int_{\mathbb{R}_+^2\times\mathcal{J}\times W} H(s,r,j,v)\widetilde{\mathbf{N}}_W(ds,dr,dj,dv)\right]\\
&= \mathbb{E}\left[\sum_{v\in W}\int_{\mathbb{R}_+^2\times\mathcal{J}} \mathbb{I}_{\{r> r_j^v(s,\widetilde{X},\widetilde{\kappa})\}} H(s,r,j,v)\widehat{\mathbf{N}}(ds,dr,dj)\right]\\
&\qquad + \mathbb{E}\left[\sum_{v\in W}\sum_{k\in\mathbb{N}} H(t_k^v, R_k^v r_j^v(t_k^v,\widetilde{X},\widetilde{\kappa}), j_k^v, v)\right]\\
&= \mathbb{E}\left[\int_{\mathbb{R}_+^2\times\mathcal{J}\times W} \mathbb{I}_{\{r> r_j^v(s,\widetilde{X},\widetilde{\kappa})\}} H(s,r,j,v)\, ds\, dr\, \#_{\mathcal{J}\times W}(dj,dv)\right]\\
&\qquad + \mathbb{E}\left[\int_{\mathbb{R}_+^2\times\mathcal{J}\times W} \mathbb{I}_{\{r\le r_j^v(s,\widetilde{X},\widetilde{\kappa})\}} H(s,r,j,v)\, ds\, dr\, \#_{\mathcal{J}\times W}(dj,dv)\right]\\
&= \mathbb{E}\left[\int_{\mathbb{R}_+^2\times\mathcal{J}\times W} H(s,r,j,v)\, ds\, dr\, \#_{\mathcal{J}\times W}(dj,dv)\right].
\end{aligned}
$$

By [19, Theorem 15.1.22], since H is an arbitrary left-continuous, non-negative $\widehat{\mathbb{G}}$-mark predictable function, $\widetilde{\mathbf{N}}_W$ is a $\widehat{G}$-Poisson process on $\mathbb{R}_+^2\times\mathcal{J}\times W$ with intensity measure $\mathrm{Leb}^2\otimes\#_{\mathcal{J}\times W}$. Because we fixed $W\subseteq U$ to be finite and arbitrary, it follows that $\widetilde{\mathbf{N}}$ is a $\widehat{\mathbb{G}}$-Poisson process on $\mathbb{R}_+^2\times\mathcal{J}\times U$ with intensity measure $\mathrm{Leb}^2\otimes\#_{\mathcal{J}\times U}$. For each $v\in V$, let $\widetilde{\mathbf{N}}_v(ds,dr,dj) := \widetilde{\mathbf{N}}(ds,dr,dj,\{v\})$. To complete the proof of the proposition, note that by (8.2.4), for every $v\in U$ and $t\in\mathbb{R}_+$,

$$
\begin{aligned}
\widetilde{X}_v(0) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r\le r_j^v(s,\widetilde{X},\widetilde{\kappa})\right\}}\widetilde{\mathbf{N}}_v(ds,dr,dj) &= \widetilde{X}_v(0) + \sum_{\substack{k\in\mathbb{N}\\ t_k^v\in(0,t]}} j_k^v\mathbb{I}_{\left\{R_k^v r_{j_k^v}^v(t_k^v,\widetilde{X},\widetilde{\kappa})\le r_{j_k^v}^v(t_k^v,\widetilde{X},\widetilde{\kappa})\right\}} \\
&= \widetilde{X}_v(0) + \sum_{s\in\mathrm{Disc}_t\left(\widetilde{X}_v\right)} \Delta\widetilde{X}_v(s) \\
&= \widetilde{X}_v(t).
\end{aligned}
$$

This proves (8.2.3) and completes the proof of the proposition.

□

8.2.2 Proof of Proposition 8.6

As alluded to earlier, the main difficulty in the proof of Proposition 8.6 is that X and $\widehat{X}^W$ will typically have explosive duals, and thus we cannot apply standard point process change of measure theorems directly. However, the proof will exploit a classical change of measure result from [19], which is reproduced in Appendix D.3 and relies on the notion of local characteristics of a marked point process that we now define in the case that the mark space of the point process is finite.

Definition 8.14. [37] Fix a complete, filtered probability space $(\Omega, \mathcal{G}, \mathbb{G}, \eta)$ and let N_Z be a $\mathbb{G}$-adapted marked point process on $\mathbb{R}_+ \times \mathcal{Z}$ for some finite space $\mathcal{Z}$ equipped with the counting measure $\#_{\mathcal{Z}}$ with $\mathbb{G}$-intensity $(t,z) \mapsto \lambda(t,z)$. If there exist functions $t \mapsto \lambda_g(t) := \sum_{z\in\mathcal{Z}} \lambda(t,z)$ and $\Psi : \mathbb{R}_+ \times \mathcal{Z} \to \mathbb{R}_+$ such that for any $t \in \mathbb{R}_+$, $\sum_{z\in\mathcal{Z}} \Psi(t,z) = 1$ and

$$
\lambda(t,z) = \lambda_g(t)\Psi(t,z) \text{ for all } (t,z) \in \mathbb{R}_+ \times \mathcal{Z},
$$

then we say that N_Z admits the $(\eta,\mathbb{G})$-local characteristics (λ_g, Ψ).

Proof of Proposition 8.6: For simplicity of notation, we fix the finite set W, let $\widehat{X} = \widehat{X}^W$, $\widehat{\mu} = \widehat{\mu}^W$ and let $L = L^W$ be the quantity from (8.1.8) and for all $t \in \mathbb{R}_+$, $z = (j,v) \in \mathcal{J} \times V$, define

$$
r(t,z) := r_j^v(t,\kappa,\widehat{X}) \quad \text{and} \quad \widehat{r}(t,z) := \widehat{r}_j^{v,W}(t,\kappa,\widehat{X}) := \begin{cases} 1 & \text{if } v \in W, \\ r_j^v(t,\kappa,\widehat{X}) & \text{if } v \notin W. \end{cases}
$$

By Lemma 8.5, the jump characteristics of $\widehat{X}_W$ are a.s. well defined, so L is also well defined. Let $\widehat{\mathbb{G}} = \mathcal{H}^\kappa \vee \mathbb{H}^{\widehat{X}}$. From (8.1.8) it follows that L is $\widehat{\mathbb{G}}$-adapted and a.s. càdlàg. Fix any $n \in \mathbb{N}$ and finite $U \subseteq V$ such that $W \subseteq U$. Let $\widehat{P}$ be the dual of $\widehat{X}$ and $\widehat{P}_U := \widehat{P}|_{\mathbb{R}_+ \times \mathcal{J} \times U}$ be the dual of $\widehat{X}_U$ in the sense of Definition 8.11. Then by Proposition 8.12(a), $\widehat{P}_U$ has $\widehat{\mathbb{G}}$-intensity $\widehat{\Lambda}_U(t,z) := \widehat{r}(t,z)$ for all $(t,z) \in \mathbb{R}_+ \times \mathcal{J} \times U$. We define the space $\mathcal{Z} := \mathcal{J} \times U$ and the point process $N_Z := \widehat{P}_U$, where $\widehat{P}_U$ admits the $(\widehat{\mu}, \widehat{\mathbb{G}})$-local characteristics

$$\widehat{\lambda}_U(t) := \sum_{z' \in \mathcal{J} \times U} \widehat{r}(t,z') \quad \text{ and } \quad \widehat{\Psi}_U(t,z) := \frac{\widehat{r}(t,z)}{\widehat{\lambda}_U(t)} \qquad \text{for all } t \in \mathbb{R}_+, z \in \mathcal{J} \times U.$$

Also, define λ_U and Ψ_U analogously, but with $r(\cdot,\cdot)$ in place of $\widehat{r}(\cdot,\cdot)$. Then note that the dual $\widehat{P}_U$ is locally finite because the jump characteristics of $\widehat{X}_U$ exist and that the jump characteristics $\{(t_k, z_k) := (t_k, (j_k, v_k))\}_{k \in \mathbb{N}}$ are precisely the events of $\widehat{P}_U$. Let $N_Z := \widehat{P}_U$, $\Psi := \widehat{\Psi}_U$, $\lambda_g := \widehat{\lambda}_U$, $\theta := \theta_U$ and $h := h_U$ where θ_U and h_U are defined by

$$\theta_U(t) := \frac{\lambda_U(t)}{\widehat{\lambda}_U(t)} = \frac{\sum_{z' \in \mathcal{J} \times U} r(t,z')}{|\mathcal{J}||W| + \sum_{\zeta' \in \mathcal{J} \times (U \setminus W)} r(t,\zeta')} \quad \text{ and } \quad h_U(t,\zeta) := \frac{\frac{r(t,z)}{\lambda_U(t)}}{\frac{\widehat{r}(t,z)}{\widehat{\lambda}_U(t)}}, \quad t > 0, z \in \mathcal{J} \times U.$$

Then the quantity $L(t) := L_t^W = L_t$ from (8.1.8) satisfies (D.3.3):

$$L(t) = \left(\prod_{t_n \in (0,t]} \theta(t_n) h(t_n, z_n) \right) \exp\left(- \sum_{z \in \mathcal{Z}} \int_{(0,t]} (\theta(s)h(s,z) - 1) \lambda_g(s) \Psi(s,z)\, ds \right).$$

Because $\theta(t)h(t,z) = 1$ for $z \in \mathcal{J} \times (U \setminus W)$, L does not depend upon the choice of U so long as $W \subseteq U$. Then by the definitions above and (8.1.7),

$$\int_0^t \theta_U(s) \widehat{\lambda}_U(s)\, ds = \int_0^t \sum_{z' \in \mathcal{J} \times U} r(s,z')\, ds = \sum_{(j,v) \in \mathcal{J} \times U} \int_0^t r_j^v(s, \widehat{X}, \kappa)\, ds < \infty \text{ a.s..}$$

Likewise, for any $t \in \mathbb{R}_+$,

$$\sum_{z \in \mathcal{J} \times U} h_U(t,z)\, \widehat{\Psi}_U(t,z) = \sum_{z \in \mathcal{J} \times U} \frac{r(t,z)}{\lambda_U(t)} = 1.$$

In summary, on the complete probability space $(\Omega, \widehat{\mathbb{G}}, \mathbb{P})$, $\widehat{P}_U$ is a locally finite point process on $\mathbb{R}_+$ with marks in $\mathcal{J} \times U$ that admits the $(\mathbb{P}, \widehat{\mathbb{G}})$-local characteristics (λ, Ψ). Then θ_U is non-negative and $\widehat{\mathbb{G}}$-predictable while h_U is non-negative and $\widehat{\mathbb{G}}$-mark predictable. Finally, we have shown that

(D.3.1), (D.3.2) and (D.3.3) all hold for $N_Z = \widehat{P}_U$, $\Psi = \widehat{\Psi}_U$, $\lambda_g = \widehat{\lambda}_U$, $\theta = \theta_U$ and $h = h_U$. Thus, the conditions of Theorem D.1 are satisfied, and so L is a $\widehat{G}$-local martingale. This proves the first assertion of the proposition.

Next, suppose that L is also a $\widehat{\mathbb{G}}$-martingale. Then $\mathbb{E}\left[L(t)\right] = \mathbb{E}\left[L(0)\right] = 1$ for all $t \in \mathbb{R}_+$. Fix a probability measure $\widetilde{\mathbb{P}}$ on the space $(\Omega, \widetilde{\mathbb{G}})$, where $\widetilde{\mathbb{G}}$ is the $\widetilde{\mathbb{P}}$-completion of $\widehat{\mathbb{G}}$. For $t \geq 0$, fix $\mathbb{P}_t := \mathbb{P}|_{\sigma(\kappa, \widehat{X}[t])}$ and $\widetilde{\mathbb{P}}_t := \widetilde{\mathbb{P}}|_{\sigma(\kappa, \widehat{X}[t])}$ and suppose that for each $t \geq 0$, $d\widetilde{\mathbb{P}}_t / d\mathbb{P}_t = L(t)$ a.s.. Define $\widetilde{\mu} := \mathrm{Law}(\kappa, \widehat{X})$ with respect to $\widetilde{\mathbb{P}}$. Because $L(t)$ does not depend on the choice of $U \supseteq W$, neither does $\widetilde{\mathbb{P}}_t$. Fix $t \in \mathbb{R}_+$ and define $\widetilde{L}_t : (\mathcal{K} \times \mathcal{D}_t)^V \to \mathbb{R}_+$ by,

$$\widetilde{L}_t(\vartheta, x) := \left[\prod_{0 < t_k \leq t} r_{j_k}^{v_k}(t_k, \vartheta, x) \right] \exp\left(- \sum_{(j,v) \in \mathcal{J} \times W} \int_{(0,t]} \left(\bar{r}_j^v(s, \vartheta, x) - 1 \right) ds \right),$$

if the jump characteristics $\{(t_k, v_k, j_k)\}$ of x_W exist, and 0 otherwise. Then $\widetilde{L}_t(\kappa, \widehat{X}[t]) = L(t)$ $\mathbb{P}_t$-a.s., so for any $A \in \mathcal{B}((\mathcal{K} \times \mathcal{D}_t)^V)$,

$$\widetilde{\mu}_t(A) = \widetilde{\mathbb{P}}_t((\kappa, X) \in A) = \mathbb{E}^{\mathbb{P}_t}\left[L(t) \mathbb{I}_{\{(\kappa, X) \in A\}} \right] = \int_A \widetilde{L}_t(\vartheta, x)\, \widehat{\mu}_t(d\vartheta, dx).$$

Thus, $\frac{d\widetilde{\mu}_t}{d\widehat{\mu}_t}(\kappa, \widehat{X}[t]) = L(t)$ $\mathbb{P}$-a.s.. Lastly, because $\widehat{X}(t) = \widehat{X}(t-)$ $\mathbb{P}_t$ (and therefore $\widetilde{\mathbb{P}}_t$)-a.s., $\frac{d\widetilde{\mu}_{t-}}{d\widehat{\mu}_{t-}}(\kappa, \widehat{X}[t)) = \frac{d\widetilde{\mu}_t}{d\widehat{\mu}_t}(\kappa, \widehat{X}[t]) = L(t) = L(t-)$ a.s.. All that remains is to prove that $\widetilde{\mu} = \mathrm{Law}(\kappa, X)$ for some weak solution X to (3.1.4).

It follows from Theorem D.1 that $\widehat{P}_U$ admits the $(\widetilde{\mathbb{P}}, \widetilde{\mathbb{G}})$-local characteristics

$$\widetilde{\lambda}(t) = \sum_{z' \in \mathcal{J} \times U} r(t, z') \quad \text{and} \quad \widetilde{\Psi}(t, z) = \frac{r(t, z)}{\widetilde{\lambda}(t)}.$$

Thus, for $(t, z) \in \mathbb{R}_+ \times \mathcal{J} \times U$, $\widehat{P}_U$ has $\widetilde{\mathbb{G}}$-intensity,

$$\widetilde{\Lambda}_U(t, z) = r(t, z) = r_j^v(t, \kappa, \widehat{X}),$$

for $z = (j, v) \in \mathcal{J} \times U$ with respect to $\widetilde{\mathbb{P}}$. However, since $\widehat{P}_U = \widehat{P}|_{\mathbb{R}_+ \times \mathcal{J} \times U}$, and the above display holds for all finite $U \supseteq W$, this implies that $\widehat{P}$ has $\widetilde{\mathbb{G}}$-intensity,

$$\widetilde{\Lambda}(t, z) = r_j^v(t, \kappa, \widehat{X}),$$

for $z = (j, v) \in \mathcal{J} \times V$ with respect to $\widetilde{\mathbb{P}}$.

Note that G is a deterministic graph with a countable vertex set, and the jump rate functions

$\{r_j^v\}$ satisfy conditions (2) and (4) of Assumption 4 by assumption. Moreover, because $\widehat{X}_v(0) = \varsigma(\kappa_v)$ $\mathbb{P}$-a.s. for all $v \in V$, and because $\widetilde{\mathbb{P}}_0 \ll \mathbb{P}_0$, the same must hold $\widetilde{\mathbb{P}}$-a.s.. Then by the above display, under $\widetilde{\mathbb{P}}$, the intensity of the dual $\widehat{P}$ of $\widehat{X}$ satisfies (8.2.2). Thus, Proposition 8.12(b) states that, by extending the probability space if necessary, we may assume without loss of generality that $(\Omega, \mathbb{F}, \widetilde{\mathbb{P}})$ supports a filtration $\overline{\mathbb{G}} \supseteq \widetilde{\mathbb{G}}$ satisfying the usual conditions and a collection of i.i.d. $\overline{\mathbb{G}}$-driving Poisson processes $\widetilde{\mathbf{N}} := \{\widetilde{\mathbf{N}}_v\}_{v\in V}$ such that $\widehat{X}$ satisfies (6.1.1) driven by $\widetilde{\mathbf{N}}$ $\widetilde{\mathbb{P}}$-a.s. for the initial data κ. Thus, with respect to $\widetilde{\mathbb{P}}$, $\widehat{X}$ is a weak solution to (6.1.1) for the initial data κ and $\widetilde{\mu} = \mathrm{Law}(\kappa, \widehat{X})$. This concludes the proof of the proposition. $\square$

8.3 Proof of the Tiling Proposition

We prove Proposition 6.12 by decomposing the statement of the Proposition into four smaller statements. First, in Section 8.3.1, we directly show that any two automorphism invariant 2-SGMRF probability measures whose V_1^d marginals match must be equal and that the marginals of any automorphism invariant 2-SGMRF must be strongly symmetric. Then in Section 8.3.2, we provide a direct construction of the tiled extension of a measure and prove that the tiled extension is well defined when the measure is strongly symmetric. Then in Section 8.3.3 we show that the tiled extension is automorphism invariant, and in Section 8.3.4 we show that it is a 2-SGMRF. Then Section 8.3.5 combines these elements into a complete proof of Proposition 6.12.

In this section, we work with automorphisms of $\mathbb{T}_d$ and its subgraphs. Therefore, it is useful to have a labeling scheme that helps clearly identify the vertices of $\mathbb{T}_d$. To this end, we work with the following canonical labeling scheme for the vertices of G, known as Ulam-Harris-Neveu labeling (see [60, Section 2.1.2] and the references therein). We define the following labeling on V^d:

$$V^d := \{\varnothing\} \cup \left(\{1, \ldots, d\} \times \bigcup_{k=0}^{\infty} \{1, \ldots, d-1\}^k \right).$$

For each vertex $v \in V^d$, we denote the ith child of v by the word vi. If $v = \varnothing$, then $vi = \varnothing i = i$. In addition, recall the partial order on V^d in which for $u, v \in V^d$, $u < v$ if u is descended from v. That is, there exists a $w \in \bigcup_{k=1}^{\infty} \mathbb{N}^k$ such that $u = vw$.

8.3.1 Uniqueness of the Marginal

We prove that the V_1^d-marginals of automorphism invariant 2-SGMRF measures are distinct and strongly symmetric.

Lemma 8.15. *Let $\eta, \eta' \in \mathcal{P}(\mathcal{Z}^{V^d})$ be distinct automorphism invariant 2-SGMRFs. Then $\eta[V_1^d]$ is strongly symmetric and $\eta[V_1^d] \neq \eta'[V_1^d]$.*

Proof. Strong symmetry of $\eta[V_1^d]$ holds by automorphism invariance. In particular, $\eta[\{\varnothing, 1\}] = \varphi_* \eta[\{\varnothing, 1\}] = \eta[\{1, \varnothing\}]$ for any automorphism $\varphi : V^d \to V^d$ that transposes $\varnothing$ and 1.

To prove uniqueness, we argue by contradiction and assume that $\eta[V_1^d] = \eta'[V_1^d]$. Then this may be regarded as the base case of the inductive claim that $\eta[V_{n-1}^d] = \eta'[V_{n-1}^d]$ implies $\eta[V_n^d] = \eta'[V_n^d]$ for all $n > 1$. To prove the inductive claim, fix any $n > 1$ and suppose that $\eta[V_{n-1}^d] = \eta'[V_{n-1}^d]$. For each $v \in \mathcal{L}(V_{n-1}^d)$ let $f_v : \mathcal{Z}^{\mathrm{cl}_v(\mathbb{T}_d)} \to \mathbb{R}$ be a bounded measurable function and let $\{v_i\}_{i=1}^{d(d-1)^{n-2}}$ be an enumeration of the vertices of $\mathcal{L}(V_{n-1}^d)$. Let $Z \sim \eta$ and $Z' \sim \eta'$. Then by the 2-SGMRF property, it follows that for any disjoint $U_1, U_2 \subseteq \mathcal{L}(V_{n-1}^d)$,

$$Z_{\mathrm{cl}_{U_1}} \perp\!\!\!\perp Z_{\mathrm{cl}_{U_2}} | Z_{V_{n-1}^d}.$$

Repeatedly applying the above display for $U_1 = \{v_k\}$ and $U_2 = \bigcup_{i=k+1}^{d(d-1)^{n-2}} \{v_i\}$, $k = 1, \ldots, d(d-1)^{n-2}$ yields,

$$\begin{aligned}
\mathbb{E}\left[\prod_{i=1}^{d(d-1)^{n-2}} f_{v_i}(Z_{\mathrm{cl}_{v_i}}) \middle| Z_{V_{n-1}^d}\right] &= \mathbb{E}\left[f_{v_1}(Z_{\mathrm{cl}_{v_1}}) \middle| Z_{V_{n-1}^d}\right] \mathbb{E}\left[\prod_{i=2}^{d(d-1)^{n-2}} f_{v_i}(Z_{\mathrm{cl}_{v_i}}) \middle| Z_{V_{n-1}^d}\right] \\
&= \prod_{i=1}^{d(d-1)^{n-2}} \mathbb{E}\left[f_{v_i}(Z_{\mathrm{cl}_{v_i}}) \middle| Z_{\{v_i, \pi_{v_i}\}}\right].
\end{aligned} \tag{8.3.1}$$

Moreover, for any i let φ_i be an automorphism of $\mathbb{T}_d$ that maps v_i to $\varnothing$ and π_{v_i} to 1. Then it must map c_{v_i} to $\{2, \ldots, d\}$, so by automorphism invariance,

$$Z_{\mathrm{cl}_{v_i}} \overset{(\mathrm{d})}{=} Z_{V_1^d} \Rightarrow \left(\mathbb{E}\left[f_{v_i}(Z_{\mathrm{cl}_{v_i}}) \middle| Z_{\{v_i, \pi_{v_i}\}}\right], Z_{\{v_i, \pi_{v_i}\}}\right) \overset{(\mathrm{d})}{=} \left(\mathbb{E}\left[f_{v_i}(Z_{V_1^d}) \middle| Z_{\{\varnothing, 1\}}\right], Z_{\{\varnothing, 1\}}\right).$$

Note that both equalities (in distribution) hold for Z' as well. Then because $Z_{V_1^d} \overset{(\mathrm{d})}{=} Z'_{V_1^d}$, there must exist for each i a bounded, measurable function $\psi_i : \mathcal{Z}^{\{v_i,\pi_{v_i}\}} \to \mathbb{R}$ such that,

$$\mathbb{E}\left[f_{v_i}(Z_{\mathrm{cl}_{v_i}})\middle|Z_{\{v_i,\pi_{v_i}\}}\right] = \psi_i(Z_{\{v_i,\pi_{v_i}\}}) \text{ a.s. and } \mathbb{E}\left[f_{v_i}(Z'_{\mathrm{cl}_{v_i}})\middle|Z'_{\{v_i,\pi_{v_i}\}}\right] = \psi_i(Z'_{\{v_i,\pi_{v_i}\}}) \text{ a.s..}$$

Then the inductive assumption $Z_{V^d_{n-1}} \stackrel{(\mathrm{d})}{=} Z'_{V^d_{n-1}}$ implies that $\prod_{i=1}^{d(d-1)^{n-2}} \psi_i(Z_{\{v_i,\pi_{v_i}\}}) \stackrel{(\mathrm{d})}{=} \prod_{i=1}^{d(d-1)^{n-2}} \psi_i(Z'_{\{v_i,\pi_{v_i}\}})$. Together with (8.3.1) and the above display, this implies

$$\mathbb{E}\left[\prod_{i=1}^{d(d-1)^{n-2}} f_{v_i}(Z_{\mathrm{cl}_{v_i}})\middle|Z_{V^d_{n-1}}\right] \stackrel{(\mathrm{d})}{=} \mathbb{E}\left[\prod_{i=1}^{d(d-1)^{n-2}} f_{v_i}(Z'_{\mathrm{cl}_{v_i}})\middle|Z'_{V^d_{n-1}}\right].$$

Given that the set of functions $f : \mathcal{Z}^{\mathcal{L}(V^d_n)} \to \mathbb{R}$ such that $\mathbb{E}\left[f(Z_{\mathcal{L}(V^d_n)})|Z_{V^d_{n-1}}\right] \stackrel{(\mathrm{d})}{=} \mathbb{E}\left[f(Z'_{\mathcal{L}(V^d_n)})|Z'_{V^d_{n-1}}\right]$ is a monotone class, it immediately follows that $\mathbb{E}\left[f(Z_{\mathcal{L}(V^d_n)})\middle|Z_{V^d_{n-1}}\right] \stackrel{(\mathrm{d})}{=} \mathbb{E}\left[f(Z'_{\mathcal{L}(V^d_n)})\middle|Z'_{V^d_{n-1}}\right]$ for all bounded, measurable function $f : \mathcal{Z}^{\mathcal{L}(V^d_n)} \to \mathbb{R}$ by the monotone class theorem [32, Theorem 5.2.2]. Thus, $\eta[V^d_n|V^d_{n-1}] = \eta'[V^d_n|V^d_{n-1}]$, and by the inductive assumption, $\eta[V^d_{n-1}] = \eta'[V^d_{n-1}]$ so $\eta[V^d_n] = \eta'[V^d_n]$ as desired. Then by induction, it follows that $\eta = \eta'$ which contradicts the assumption that η and η' are distinct. $\square$

8.3.2 The Tiling Construction

In this section we describe a tiling construction that, given a strongly symmetric $\eta \in \mathcal{P}(\mathcal{Z}^{V^d_1})$, outputs a measure $\overline{\eta} \in \mathcal{P}(\mathcal{Z}^{V^d})$ such that $\overline{\eta}[V^d_1] = \eta$. We will also prove that the tiling construction is well defined. In the next two sections we will show that the tiling construction always outputs automorphism invariant 2-SGMRF measures.

Let $\eta \in \mathcal{P}(\mathcal{Z}^{V^d_1})$ be strongly symmetric. Let $V'_1 := \{2, \dots, d\}$ and $V'_0 := \{\varnothing, 1\}$ so that $\eta[V'_1|V'_0] = \eta(dz_{\{2,\dots,d\}}|z_{\{\varnothing,1\}})$. Then define $\overline{\eta}^1 := \eta$ and for each $n > 1$ define,

$$\overline{\eta}^n(dz_{V^d_n}) := \left(\otimes_{v\in\mathcal{L}(V^d_{n-1})}\eta[V'_1|V'_0]\left(dz_{c_v(\mathbb{T}_d)}\middle|z_{\{v,\pi_v\}}\right)\right) \otimes \overline{\eta}^{n-1}(dz_{V^d_{n-1}}). \tag{8.3.2}$$

We will show that this construction implies that $\overline{\eta}^n[V^d_{n-1}] = \overline{\eta}^{n-1}$, and for $\overline{Z}^n \sim \overline{\eta}^n$, the collection of random elements $\left(\overline{Z}^n_{\mathrm{cl}_v}\right)_{v\in\mathcal{L}(V^d_{n-1})}$ are mutually conditionally independent given $\overline{Z}^n_{V^d_{n-1}}$ and equal in distribution to Z for $Z \sim \eta$.

Definition 8.16. For any strongly symmetric $\eta \in \mathcal{P}(\mathcal{Z}^{V_1^d})$, let $\overline{\eta} \in \mathcal{P}(\mathcal{Z}^{V^d})$ be the unique probability measure such that $\overline{\eta}[V_n^d] = \overline{\eta}^n$ for any $n \in \mathbb{N}$. Then $\overline{\eta}$ is said to be the *tiled extension* of η.

Lemma 8.17. *For any strongly symmetric $\eta \in \mathcal{P}(\mathcal{Z}^{V_1^d})$, there exists a unique tiled extension $\overline{\eta}$ of η in the sense of Definition 8.16.*

Proof. Note that if $\overline{\eta}^n$ is well defined and satisfies $\overline{\eta}^n[V_{n-1}^d] = \overline{\eta}^{n-1}$ for all $n > 1$, then $\overline{\eta}$ is well defined by the Kolmogorov extension theorem [86, Theorem 2.4.3] (noting that all Borel probability measures are inner regular). Furthermore, $\overline{\eta}^1 = \eta$ is clearly well defined, so it suffices to prove that given any $n > 1$, if $\overline{\eta}^{n-1}$ is well defined, then there exists a unique measure $\overline{\eta}^n \in \mathcal{P}(\mathcal{Z}^{V_n^d})$ satisfying (8.3.2). Furthermore, because the right-hand side of (8.3.2) depends entirely on the measures $\overline{\eta}^{n-1}$ and η, it suffices to prove that $\overline{\eta}^n$ exists.

The probability measure $\overline{\eta}^n$ exists if and only if for each $v \in \mathcal{L}(V_{n-1}^d)$, $\eta[V_1'|V_0'](\cdot|z_{\{v,\pi_v\}})$ is defined for $\overline{\eta}^{n-1}[\{v,\pi_v\}]$-a.s. $z_{\{v,\pi_v\}} \in \mathcal{Z}^{\{\varnothing,1\}}$. Therefore, we simply need to show that $\overline{\eta}^{n-1}[\{v,\pi_v\}]$ is absolutely continuous with respect to $\eta[\{\varnothing,1\}]$.

In fact, we claim that $\overline{\eta}^{n-1}[\{v,\pi_v\}] = \eta[\{\varnothing,1\}]$. If $n = 2$, then note that $\overline{\eta}^1[\{v,\pi_v\}] = \eta[\{k,\varnothing\}]$ for some $k \in \mathbb{N}$. Then by the strong symmetry of η, $\overline{\eta}^1[\{v,\pi_v\}] = \eta[\{1,\varnothing\}] = \eta[\{\varnothing,1\}]$. We prove the claim by using the $n = 2$ case as a base case and making the inductive assumption that $n \geq 3$ and $\overline{\eta}^{n-2}[\{v,\pi_v\}] = \eta[\{\varnothing,1\}]$ for all $v \in \mathcal{L}(V_{n-2}^d)$. For $v \in \mathcal{L}(V_{n-2}^d)$ and $i \in \{1,\ldots,d\}$, by (8.3.2) and the inductive assumption,

$$\begin{aligned}\overline{\eta}^{n-1}[\mathrm{cl}_v](dz_{\mathrm{cl}_v}) &= \eta[V_1'|V_0'](dz_{c_v}|z_{\{v,\pi_v\}}) \otimes \overline{\eta}^{n-2}[\{v,\pi_v\}](dz_{\{v,\pi_v\}}) \\ &= \eta[V_1'|V_0'](dz_{c_v}|z_{\{v,\pi_v\}}) \otimes \eta[\{\varnothing,1\}](dz_{\{v,\pi_v\}}) \\ &= \eta(dz_{\mathrm{cl}_v}).\end{aligned}$$

Thus, $\overline{\eta}^{n-1}[\{vi,v\}] = \eta[\{\varnothing,1\}]$ for any $v \in \mathcal{L}(V_{n-2}^d)$ and $i \in \{1,\ldots,d\}$. By induction, this implies the claim.

□

8.3.3 Automorphism-Invariance of the Tiled Measure

In this section we establish the automorphism-invariance of the tiled measure. The proof of this result (stated below) is deferred to the end of the section.

Lemma 8.18. *The tiled extension $\overline{\eta}$ of η is automorphism invariant for any strongly symmetric $\eta \in \mathcal{P}(\mathcal{Z}^{V_1^d})$.*

To prove this, we first show that there exists a small set of automorphisms that generate a group that is dense in the group of unrooted automorphisms of $\mathbb{T}_d$. It therefore suffices to prove that the tiled measure is invariant with respect to these automorphisms, which we do by employing an inductive argument on $\overline{\eta}^n$. To prove automorphism-invariance, we define the following specific automorphisms of $\mathbb{T}_d$.

Definition 8.19. For any $\sigma \in S_d$ where S_d is the symmetric group defined over $\{1,\ldots,d\}$, define the automorphism $\varphi^\sigma : V^d \to V^d$ by,

$$\varphi^\sigma(v) = \begin{cases} \varnothing & \text{if } v = \varnothing, \\ \sigma(k)u & \text{if } v = ku \leq k \in \{1,\ldots,d\}. \end{cases}$$

Additionally define the inversion $\varphi^1 : V^d \to V^d$ by

$$\varphi^1(v) = \begin{cases} 1 & \text{if } v = \varnothing, \\ \varnothing & \text{if } v = 1, \\ (k+1)u & \text{if } v = 1ku < 1 \text{ for } k \in \{1,\ldots,d-1\}, \\ 1(k-1)u & \text{if } v = ku \text{ for } k \in \{2,\ldots,d\}. \end{cases}$$

Lastly we write $\mathcal{E} := \{\varphi^1, \varphi^\sigma : \sigma \in S_d\}$.

We will show in the next lemma that the group generated by $\mathcal{E}$ (henceforth denoted by $\langle\mathcal{E}\rangle$) is a dense subgroup of the automorphism group of $\mathbb{T}_d$ (denoted $\mathrm{Aut}(\mathbb{T}_d)$) with respect to the topology of pointwise convergence, and therefore any probability measure that is invariant with respect to automorphisms in $\mathcal{E}$ is automorphism invariant.

Lemma 8.20. *Any $\eta \in \mathcal{P}(\mathcal{Z}^{V^d})$ such that $\varphi_*\eta = \eta$ for every $\varphi \in \mathcal{E}$ is automorphism invariant.*

Proof. As mentioned in the preceding paragraph, we claim that $\langle \mathcal{E} \rangle$ is dense in $\mathrm{Aut}(\mathbb{T}_d)$. We start with a proof of the result given the claim. Note that any automorphism $\varphi \in \langle \mathcal{E} \rangle$ is expressible as a finite composition of automorphisms in $\mathcal{E}$, so $\varphi_*\eta = \eta$. Now consider any automorphism $\overline{\varphi} \in \mathrm{Aut}(\mathbb{T}_d)$. By the claim, there must exist a sequence of automorphisms $\varphi_n \in \langle \mathcal{E} \rangle$, $n \in \mathbb{N}$ converging pointwise to $\overline{\varphi}$. Then for any $Z \sim \eta$, $\varphi_{n,*}Z \to \overline{\varphi}_*Z$ as $n \to \infty$ and $\varphi_{n,*}\eta = \eta$ for all $n \in \mathbb{N}$ so $\overline{\varphi}_*Z \sim \eta$. Thus, $\overline{\varphi}_*\eta = \eta$, which proves that η is automorphism invariant.

We now prove the claim. Let $\mathrm{Aut}_*(\mathbb{T}_d)$ be the group of rooted automorphisms of $\mathbb{T}_d$. Then we make two further claims:

(a) for any $v \in V^d$, there exists an automorphism $\varphi^v \in \langle \mathcal{E} \rangle$ such that $\varphi^v(v) = \varnothing$;

(b) $\langle \mathcal{E} \rangle \cap \mathrm{Aut}_*(\mathbb{T}_d)$ is dense in $\mathrm{Aut}_*(\mathbb{T}_d)$.

Suppose claims (a) and (b) hold. For any $\varphi \in \mathrm{Aut}(\mathbb{T}_d)$, let $v \in V^d$ be such that $\varphi(v) = \varnothing$. Then by claim (a), there exists $\varphi^v \in \langle \mathcal{E} \rangle$ such that $\varphi^v(v) = \varnothing$. Thus, $\varphi \circ (\varphi^v)^{-1}(\varnothing) = \varphi(v) = \varnothing$, so $\varphi \circ (\varphi^v)^{-1} \in \mathrm{Aut}_*(\mathbb{T}_d)$. By claim (b), there exists a sequence $\varphi_n \in \langle \mathcal{E} \rangle, n \in \mathbb{N}$ converging pointwise to $\varphi \circ (\varphi^v)^{-1}$. Then $\varphi_n \circ \varphi^v \in \langle \mathcal{E} \rangle$ for every n and converges pointwise to φ, so φ is in the closure of $\langle \mathcal{E} \rangle$. Conclude that $\langle \mathcal{E} \rangle$ is dense in $\mathrm{Aut}(\mathbb{T}_d)$. Thus, claims (a) and (b) imply the claim we made at the beginning of the proof.

Proof of (a): For $k, k' \in \{1, \ldots, d\}$, let $(k, k') \in S_d$ be the transposition that swaps k and k'. If $v = \varnothing$, then let φ^v be the identity. For for any $n \in \mathbb{N}$ and $v = k_1 k_2 \ldots k_n \in \mathcal{L}(V_n^d)$,

$$\psi^v(v) := (\varphi^1 \circ \varphi^{(k_n+1,1)}) \circ \cdots \circ (\varphi^1 \circ \varphi^{(k_1,1)})(v) = \varnothing.$$

Proof of (b): For any permutation $\sigma' \in S_{d-1}$, define

$$\sigma(k) := \begin{cases} 1 & \text{if } k = 1, \\ \sigma'(k-1) + 1 & \text{if } k \in \{2, \ldots, k+1\}. \end{cases}$$

Then $\sigma \in S_d$. Fix any $v \in V^d$ and let $\sigma' \in S_{d-1}$ if $v \neq \varnothing$ letting σ be defined as above, and let $\sigma' \in S_d$ if $v = \varnothing$. Define,

$$\psi^{v,\sigma'} := \begin{cases} (\psi^v)^{-1} \circ \varphi^\sigma \circ \psi^v & \text{if } v \neq \varnothing, \\ \varphi^{\sigma'} & \text{otherwise.} \end{cases}$$

The following facts can be directly verified in the case that $v \neq \varnothing$:

(a) for any $w \nleq v$, $\psi^v(w) \leq 1$;

(b) if $w = v$ then $\psi^v(w) = \varnothing$;

(c) if $w \not< v$ then (a) and (b) imply that $\varphi^\sigma \circ \psi^v(w) = \psi^v(w)$;

(d) if $w \not< v$ then (c) implies that $\psi^{v,\sigma'}(w) = w$;

(e) if $w = viu < v$ for some $i \in \mathbb{N}$, then $\psi^v(w) = (i+1)u$;

(f) $\varphi^\sigma((i+1)u) = (\sigma'(i)+1)u$;

(g) if $w = viu < v$ for some $i \in \mathbb{N}$, then (d) and (e) imply that $\psi^{v,\sigma'}(w) = v\sigma'(i)u$.

Combining the items above, it follows that

$$\psi^{v,\sigma'}(w) = \begin{cases} w & \text{if } w \not< v, \\ v\sigma'(i)u & \text{if } w = viu < v, i \in \{1, \ldots, d\}, \end{cases} \tag{8.3.3}$$

where (8.3.3) holds trivially when $v = \varnothing$. Then (8.3.3) implies that the set $\{\psi^{v,\sigma'}\}_{v \in V^d \setminus \{\varnothing\}, \sigma' \in S_{d-1}} \cup \{\psi^{\varnothing,\sigma'}\}_{\sigma' \in S_d}$ is precisely the set of elementary automorphisms of $\mathrm{Aut}_*(\mathbb{T}_d)$ defined in [1, page 3600] and by definition, these automorphisms are all elements of $\langle \mathcal{E} \rangle$. Thus, $\langle \mathcal{E} \rangle \cap \mathrm{Aut}_*(\mathbb{T}_d)$ is dense in $\mathrm{Aut}_*(\mathbb{T}_d)$ [1, page 3600].

As both (a) and (b) above hold, the claim is proven. □

We can now prove Lemma 8.18.

Proof of Lemma 8.18: By Lemma 8.20, it suffices to prove that $\varphi_* \overline{\eta} = \overline{\eta}$ for all $\varphi \in \mathcal{E}$. Fix any $\sigma \in S_d$ and let $\varphi = \varphi^\sigma$. Then by strong symmetry,

$$\varphi_*\overline{\eta}^1 = \varphi_*\eta = \eta = \overline{\eta}^1.$$

Suppose $\varphi_*\overline{\eta}^{n-1} = \overline{\eta}^{n-1}$ for some $n > 1$. Note that $\varphi_* z = (z_{\varphi^{-1}(v)})_{v\in V^d}$ and $\varphi_*\overline{\eta}^n(dz_{V_n^d}) = \overline{\eta}^n(d(z_{\varphi^{-1}(v)})_{v\in V_n^d}) = \overline{\eta}^n(d(\varphi_* z))$. In the first equality below, we apply (8.3.2). Then we apply the definition of φ_* to substitute $\varphi_* z$ for z. In the third equality below, we use the fact that φ^{-1} induces a bijection from $\mathcal{L}(V_n^d)$ to $\mathcal{L}(V_n^d)$ to reorder the product of the regular conditional probabilities to remove φ^{-1} from the indices of z (which is permissible by Fubini-Tonelli's theorem). In the fourth equality, we apply the inductive assumption. In the final equality, we again apply (8.3.2).

$$\begin{aligned}
\varphi_*\overline{\eta}^n(dz_{V_n^d}) &= \varphi_*\left(\otimes_{v\in\mathcal{L}(V_{n-1}^d)}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right)\otimes\overline{\eta}^{n-1}(dz_{V_{n-1}^d})\\
&= \left(\otimes_{v\in\mathcal{L}(V_{n-1}^d)}\eta[V_1'|V_0'](dz_{c_{\varphi^{-1}(v)}(\mathbb{T}_d)}|z_{\{\varphi^{-1}(v),\varphi^{-1}(\pi_v)\}})\right)\otimes\overline{\eta}^{n-1}(d(z_{\varphi^{-1}(v)})_{v\in V_{n-1}^d})\\
&= \left(\otimes_{v\in\mathcal{L}(V_{n-1}^d)}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right)\otimes\overline{\eta}^{n-1}(d\varphi_* z_{V_{n-1}^d})\\
&= \left(\otimes_{v\in\mathcal{L}(V_{n-1}^d)}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right)\otimes\overline{\eta}^{n-1}(dz_{V_{n-1}^d})\\
&= \overline{\eta}^n(dz_{V_n^d}).
\end{aligned}$$

Then by Definition 8.16, for every $n\in\mathbb{N}$,

$$\varphi_*\overline{\eta}[V_n^d] = \varphi_*\overline{\eta}^n = \overline{\eta}^n = \overline{\eta}[V_n^d],$$

so $\varphi_*\overline{\eta} = \overline{\eta}$.

Now let $\varphi = \varphi^1$. This case is more involved because $\varphi(V_n^d) \neq V_n^d$ so that for $z \in \mathcal{Z}^{V_n^d}$, $\varphi_* z \notin \mathcal{Z}^{V_n^d}$. This implies that $\mathcal{P}(\mathcal{Z}^{\varphi(V_n^d)}) \ni \varphi_*\overline{\eta}^n \neq \overline{\eta}^n \in \mathcal{P}(\mathcal{Z}^{V_n^d})$ for all $n\in\mathbb{N}$. For each $n\in\mathbb{N}$, let $\mathcal{T}_n := V_n^d \cup \{v \leq 1 : d_{\mathbb{T}_d}(v,\varnothing) = n+1\}$. Then it is easily verified that $\varphi^1(\mathcal{T}_n) = \mathcal{T}_n$ and $\mathrm{cl}_{\mathcal{T}_{n-1}}(\mathbb{T}_d) = \mathcal{T}_n$. Thus, it suffices to prove that for each $n\in\mathbb{N}$, $\varphi_*\overline{\eta}^n[\mathcal{T}_{n-1}] = \overline{\eta}^n[\mathcal{T}_{n-1}]$. Note also that $\varphi^{-1} = \varphi$. First, for $n = 1$, the result follows from the strong symmetry of η.

$$\varphi_*\overline{\eta}^1[\mathcal{T}_0](dz_{\mathcal{T}_0}) = \varphi_*\eta[\{\varnothing,1\}](dz_{\{\varnothing,1\}}) = \eta[\{\varnothing,1\}](dz_{\{1,\varnothing\}}) = \eta[\{\varnothing,1\}](dz_{\{\varnothing,1\}}) = \overline{\eta}^1[\mathcal{T}_0](dz_{\mathcal{T}_0}).$$

For each $m \in \mathbb{N}$, define $L^{m,1} := \{v \leq 1 : d_{\mathbb{T}_d}(v, \varnothing) = m+1\}$ and $L^{m,2} := \mathcal{L}(V^d_m) \cap \{v \not\leq 1\}$. Then note that $\mathcal{L}(\mathcal{T}_m) = L^{m,1} \cup L^{m,2}$ and $\mathcal{L}(V^d_m) = L^{m-1,1} \cup L^{m,2}$. Then by (8.3.2) (noting that $\mathcal{T}_n = V^d_{n+1} \setminus L^{n+1,2}$ and that $L^{n-1,1} = \{v : c_v(\mathbb{T}_d) \subseteq \mathcal{L}(V^d_{n+1}) \setminus L^{n+1,2}\} = \{v : c_v(\mathbb{T}_d) \cap (\mathcal{L}(V^d_{n+1}) \setminus L^{n+1,2}) \neq \emptyset\}$),

$$\begin{aligned}\overline{\eta}^{n+1}[\mathcal{T}_n](dz_{\mathcal{T}_n}) &= \int_{\mathcal{Z}^{L^{n+1,2}}} \left(\otimes_{v\in\mathcal{L}(V^d_n)}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right) \otimes \overline{\eta}^n(dz_{V^d_n}) \\ &= \left(\otimes_{v\in L^{n-1,1}}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right) \otimes \overline{\eta}^n(dz_{V^d_n}),\end{aligned} \tag{8.3.4}$$

for any $n \in \mathbb{N}$. Then applying Definition 8.16, the fact that $\mathcal{L}(V^d_{n-1}) = L^{n-1,2} \cup L^{n-2,1}$ is a disjoint union and (8.3.4) (with n replaced by $n-1$),

$$\begin{aligned}\overline{\eta}^n(dz_{V^d_n}) &= \left(\otimes_{v\in\mathcal{L}(V^d_{n-1})}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right) \otimes \overline{\eta}^{n-1}(dz_{V^d_{n-1}}) \\ &= \left(\otimes_{v\in L^{n-1,2}}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right) \otimes \left(\otimes_{v\in L^{n-2,1}}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right) \otimes \overline{\eta}^{n-1}(dz_{V^d_{n-1}}) \\ &= \left(\otimes_{v\in L^{n-1,2}}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right) \otimes \overline{\eta}^n[\mathcal{T}_{n-1}](dz_{\mathcal{T}_{n-1}}).\end{aligned}$$

Applying (8.3.4), the above display and the fact that $\mathcal{L}(\mathcal{T}_{n-1}) = L^{n-1,1} \cup L^{n-1,2}$ is a disjoint union yields,

$$\begin{aligned}&\overline{\eta}^{n+1}[\mathcal{T}_n](dz_{\mathcal{T}_n}) \\ &= \left(\otimes_{v\in L^{n-1,1}}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right) \otimes \overline{\eta}^n(dz_{V^d_n}) \\ &= \left(\otimes_{v\in L^{n-1,1}}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right) \otimes \left(\otimes_{v\in L^{n-1,2}}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right) \otimes \overline{\eta}^n[\mathcal{T}_{n-1}](dz_{\mathcal{T}_{n-1}}) \\ &= \left(\otimes_{v\in\mathcal{L}(\mathcal{T}_{n-1})}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right) \otimes \overline{\eta}^n[\mathcal{T}_{n-1}](dz_{\mathcal{T}_{n-1}}).\end{aligned}$$

Now suppose that $\varphi_*\overline{\eta}^n[\mathcal{T}_{n-1}] = \overline{\eta}^n[\mathcal{T}_{n-1}]$ for some $n \in \mathbb{N}$. The remainder of the argument is nearly identical to the argument that $\varphi^\sigma_*\overline{\eta} = \overline{\eta}$ though the justifications differ. In the first equality below, we apply the above display. Then we apply the definition of φ_* to substitute $\varphi_* z$ for z using the fact that $\varphi = \varphi^{-1}$ to slightly simplify the expression. In the third equality below, we use the fact that φ induces a bijection from $\mathcal{L}(\mathcal{T}_n)$ to itself to reorder the product of the regular conditional probabilities to remove φ from the indices of z (which is permissible by Fubini-Tonelli's theorem).

In the fourth equality, we apply the inductive assumption. In the final equality, we again apply the above display:

$$\begin{aligned}
\varphi_*\overline{\eta}^{n+1}[\mathcal{T}_n](dz_{\mathcal{T}_n}) &= \varphi_*\left(\otimes_{v\in\mathcal{L}(\mathcal{T}_{n-1})}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right)\otimes\overline{\eta}^n[\mathcal{T}_{n-1}](dz_{\mathcal{T}_{n-1}}) \\
&= \left(\otimes_{v\in\mathcal{L}(\mathcal{T}_{n-1})}\eta[V_1'|V_0'](dz_{c_{\varphi(v)}(\mathbb{T}_d)}|z_{\{\varphi(v),\varphi(\pi_v)\}})\right)\otimes\overline{\eta}^n[\mathcal{T}_{n-1}](d(z_{\varphi(v)})_{v\in\mathcal{T}_{n-1}}) \\
&= \left(\otimes_{v\in\mathcal{L}(\mathcal{T}_{n-1})}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right)\otimes\overline{\eta}^n[\mathcal{T}_{n-1}](d(\varphi_* z_{\mathcal{T}_{n-1}}) \\
&= \left(\otimes_{v\in\mathcal{L}(\mathcal{T}_{n-1})}\eta[V_1'|V_0'](dz_{c_v(\mathbb{T}_d)}|z_{\{v,\pi_v\}})\right)\otimes\overline{\eta}^n[\mathcal{T}_{n-1}](dz_{\mathcal{T}_{n-1}}) \\
&= \overline{\eta}^{n+1}[\mathcal{T}_n](dz_{\mathcal{T}_n}).
\end{aligned}$$

It immediately follows by Definition 8.16 that for each $n \in \mathbb{N}$,

$$\varphi_*\overline{\eta}[\mathcal{T}_n] = \varphi_*\overline{\eta}^{n+1}[\mathcal{T}_n] = \overline{\eta}^{n+1}[\mathcal{T}_n] = \overline{\eta}[\mathcal{T}_n],$$

so $\varphi_*\overline{\eta} = \overline{\eta}$. This completes the proof. □

8.3.4 The Tiled Measure is a 2-SGMRF

By Remark 6.9, to prove that the tiled measure is a 2-SGMRF, it is sufficient to show that its connected, finite marginals are 2-MRFs. This is accomplished by making use of the properties of conditional independence derived in [81] and summarized in Lemma C.3. Note that all of the results of [81] are stated for complete filtrations. However, we may disregard this assumption by Lemma C.2.

Lemma 8.21. *If η is strongly symmetric and $\overline{\eta}$ is the tiled extension of η, then $\overline{\eta}$ is a 2-SGMRF.*

Proof of Lemma 8.21: We prove that $\overline{\eta}$ is a 2-SGMRF using an inductive method. Recall the alternative definition of a 2-SGMRF from Lemma 6.8. Let $\overline{Z} \sim \overline{\eta}$ and let $U \subset V^d$ be finite and connected. Suppose also that $V_1^d \subseteq U$ and $\overline{\eta}[U]$ is a 2-MRF. We claim that for any w such that $U' := U \cup \{w\}$ is still connected, $\overline{\eta}[U']$ is also a 2-MRF. Because the subgraph $\mathbb{T}_d[V_1^d]$ has diameter 2, $\overline{\eta}[V_1^d]$ is trivially a 2-MRF. Therefore, if the claim holds, then by induction, $\overline{\eta}[U]$ is a 2-MRF for any finite, connected $U \subset V^d$. By Remark 6.9, this implies that $\overline{\eta}$ is a 2-SGMRF. Thus, it suffices to prove the claim.

Fix any finite, connected $U \subset V^d$ and suppose that $\overline{\eta}[U]$ is a 2-MRF. Let $v \in U$ and $k \in \{1, \ldots, d-1\}$ be such that $w := vk \notin U$. Then $U' := U \cup \{w\}$ is also finite and connected. Let $A, B, C \subseteq U'$ partition U' such that C doubly separates A and B in $\mathbb{T}_d[U']$. Note that $C \cap U$ also doubly separates $A \cap U$ and $B \cap U$ in $\mathbb{T}_d[U]$.

Let $O_1 = \{w\}$ be the new vertex of U'. Let $O_2 = U \setminus W$ where $W = \{v, \pi_v, c_v(\mathbb{T}_d[U])\}$. Then $W = \mathcal{N}^2_{O_1}(\mathbb{T}_d[U'])$. By Definition 8.16, $\overline{Z}_{O_1} \perp\!\!\!\perp \overline{Z}_{O_2} | \overline{Z}_W$. Note that $W \cup O_1$ has diameter 2, so because C doubly separates A and B, at least one of A or B must be a subset of O_2. We assume without loss of generality that $B \subseteq O_2$. Define the following random elements:

	A	B	C
O_1	$\overline{Z}_1$		$\overline{Z}_2$
O_2	$\overline{Z}_3$	$\overline{Z}_4$	$\overline{Z}_5$
W	$\overline{Z}_6$		$\overline{Z}_7$

Here the table indicates intersection. For example, $\overline{Z}_1 = \overline{Z}_{A \cap O_1}$. Note that we let $\overline{Z}_\emptyset := \emptyset$ be the empty vector. Because O_1 is a singleton, either $\overline{Z}_1$ or $\overline{Z}_2$ is empty. This implies that

$$\overline{Z}_i \perp\!\!\!\perp \overline{Z}_{\{3,4,5\}} | \overline{Z}_{\{6,7\}} \text{ where } i = 1, 2, \text{ is defined s.t. } \overline{Z}_i = O_1.$$

Applying the above display and Lemma C.3(d) with $\mathcal{G}_1 = \sigma(\overline{Z}_i)$, $\mathcal{G}_2 = \sigma(\overline{Z}_{\{3,4,5\}}$, $\mathcal{G}_3 = \sigma(\overline{Z}_{\{6,7\}})$, $\mathcal{G}_4 = \sigma(\overline{Z}_4)$ and $\mathcal{G}_5 = \sigma(\overline{Z}_{\{3,5\}})$ yields,

$$\overline{Z}_i \perp\!\!\!\perp \overline{Z}_4 | \overline{Z}_{\{3,5,6,7\}} \text{ where } i = 1, 2, \text{ is defined s.t. } \overline{Z}_i = O_1. \tag{8.3.5}$$

Lastly, because $\overline{Z}_U$ forms a 2-MRF (by assumption), Lemma 6.8 implies that $\overline{Z}_{A \setminus \{w\}} \perp\!\!\!\perp \overline{Z}_B | \overline{Z}_C$ or,

$$\overline{Z}_{\{3,6\}} \perp\!\!\!\perp \overline{Z}_4 | \overline{Z}_{\{5,7\}}. \tag{8.3.6}$$

Case 1: $\overline{Z}_1 \neq \emptyset$

Because O_1 is a singleton, this implies $\overline{Z}_2 = \emptyset$. Applying (8.3.5) with $i = 1$, (8.3.6) and Lemma C.3(e) for $\mathcal{G}_1 = \sigma(\overline{Z}_{\{3,6\}})$, $\mathcal{G}_2 = \sigma(\overline{Z}_4)$, $\mathcal{G}_3 = \sigma(\overline{Z}_{\{5,7\}})$ and $\mathcal{G}_4 = \sigma(\overline{Z}_1)$ implies,

$$\overline{Z}_{\{1,3,6\}} \perp\!\!\!\perp \overline{Z}_4 | \overline{Z}_{\{5,7\}},$$

or equivalently, $\overline{Z}_A \perp\!\!\!\perp \overline{Z}_B | \overline{Z}_C$ as desired.

Case 2: $\overline{Z}_2 \neq \emptyset$

This implies that $\overline{Z}_1 = \emptyset$. Combining (8.3.6), (8.3.5) for $i = 2$ and Lemma C.3(g) for $\mathcal{G}_1 = \sigma(\overline{Z}_4)$, $\mathcal{G}_2 = \sigma(\overline{Z}_{3,6})$, $\mathcal{G}_3 = \sigma(\overline{Z}_{5,7})$ and $\mathcal{G}_4 = \sigma(\overline{Z}_2)$ yields,

$$\overline{Z}_{\{3,6\}} \perp\!\!\!\perp \overline{Z}_4 | \overline{Z}_{\{2,5,7\}},$$

or equivalently, $\overline{Z}_A \perp\!\!\!\perp \overline{Z}_B | \overline{Z}_C$ as desired.

In both cases, $\overline{Z}_A \perp\!\!\!\perp \overline{Z}_B | \overline{Z}_C$, so $\overline{\eta}[V']$ is a 2-MRF, completing the proof of the claim and therefore the lemma. $\square$

8.3.5 Proof of Proposition 6.12

Proof of Proposition 6.12. First let $\overline{\eta} \in \mathcal{P}(\mathcal{Z}^{V^d})$ be any automorphism invariant 2-SGMRF. Then by Lemma 8.15, $\eta := \overline{\eta}[V_1^d]$ is strongly symmetric, proving the first assertion of the proposition. Furthermore, by the same lemma, $\overline{\eta}$ is the unique automorphism invariant 2-SGMRF whose marginal is given by η which proves uniqueness in the second assertion of the proposition. Next, let $\eta' \in \mathcal{P}(\mathcal{Z}^{V_1^d})$ be strongly symmetric. Define $\overline{\eta}' \in \mathcal{P}(\mathcal{Z}^{V^d})$ to be the tiled extension of η' in the sense of Definition 8.16 (which exists and is well defined by Lemma 8.17). Then by Lemma 8.18, $\overline{\eta}'$ is automorphism invariant. By Lemma 8.21, $\overline{\eta}'$ is a 2-SGMRF. This proves existence in the second assertion of the proposition completing the proof. $\square$

CHAPTER 9

Proof of Well-Posedness of the Local Equation

9.1 Projection Results

The first step required to prove the well-posedness of the local equation is to establish a basic projection result. Simply put, we show X_U can be described by an SDE of the form of (6.1.1) where for each $(j, v) \in \mathcal{J} \times U$, the jump rate function r_j^v is replaced by a sufficiently regular function $\rho_j^v[U] : \mathbb{R}_+ \times (\mathcal{K} \times \mathcal{D})^U \to \mathbb{R}_+$ such that $t \mapsto \rho_j^v[U](t, \kappa_U^G, X_U^G)$ is a predictable, càglàd modification of $t \mapsto \mathbb{E}\left[r_j^v(t, \kappa^G, X^G) | \kappa_U^G, X_U^G[t)\right]$. This SDE is semi-autonomous in the sense that the form of the jump rate functions depends on μ^G, but the inputs to the jump rate functions depend only on the particles in U. Lemma 9.1 establishes the projection SDE and states the regularity conditions which it satisfies. Proposition 9.2 states that the solution to the projection SDE has law $\mu^G[U]$ as desired. The proofs of both results are deferred to Appendix B.

Lemma 9.1. *Suppose Assumption 4 holds and $G = (V, E)$ is a (possibly improper) subgraph of $\mathbb{T}_d$. Then for any finite $U \subseteq V$ and every $v \in U$ and $j \in \mathcal{J}$, there exists a measurable function $\rho_j^v[U] : \mathbb{R}_+ \times (\mathcal{K} \times \mathcal{D})^U \to \mathbb{R}_+$ such that for all $t \in \mathbb{R}_+$, $\rho_j^v[U](t, \kappa_U^G, X_U^G) = \mathbb{E}\left[r_j^v(t, \kappa^G, X^G) | \kappa_U^G, X_U^G[t)\right]$ a.s., and the SDE*

$$\overline{X}_v^G(t) = \varsigma(\kappa_v^G) + \int_{(0,t] \times \mathbb{R}_+ \times \mathcal{J}} j \mathbb{I}_{\left\{r \leq \rho_j^v[U](t, \kappa_U^G, \overline{X}^G)\right\}} \mathbf{N}_v(ds, dr, dj) \quad v \in U, t \in \mathbb{R}_+, \tag{9.1.1}$$

has jump rate functions that satisfy conditions (1), (2), (3) and (4) of Assumption 4 for the complete graph on U, K_U.

Proposition 9.2. *Suppose Assumption 4 holds. Then the SDE* (9.1.1) *is well-posed and possesses a pathwise unique solution $\overline{X}^G$ for the initial data κ_U^G. Furthermore, $(\kappa_U^G, \overline{X}^G) \sim \mu^G[U]$.*

Remark 9.3. We highlight here that the pathwise uniqueness of Proposition 9.2 is defined in the sense of Definition 6.14. In other words, fix two different sets of jump rate functions $\{\rho_j^{1,v}[U]\}$ and $\{\rho_j^{2,v}[U]\}$ satisfying the conditions placed on $\{\rho_j^v[U]\}$ in Lemma 9.1. Then the solutions $\overline{X}^{G,i}$, $i = 1, 2$, to (9.1.1) with the respective jump rate functions $\{\rho_j^{i,v}[U]\}$, $i = 1, 2$, for the initial data κ_U^G and driven by the Poisson processes $\mathbf{N}$ are indistinguishable.

9.2 Proof of Existence

Lemma 9.4. *Under the conditions of Theorem 6.17, there exists a strong solution Y to the local equation for the initial data $\widehat{\kappa}$ such that Law$(Y) = \mu[V_1^d]$.*

Proof. Let $\mu_\kappa := \overline{\nu}_\kappa$ be the tiled extension of $\nu_\kappa := \text{Law}(\widehat{\kappa})$, and let $\kappa \sim \mu_\kappa$ so that $\widehat{\kappa} = \kappa_{V_1^d}$. By Proposition 6.12, μ_κ is an automorphism invariant 2-SGMRF. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and $\mathbf{N}$ be the filtered probability space and driving Poisson processes defined above Definition 6.14. Recall that $V_1^d := \text{cl}_\varnothing$ and let the functions $\{\rho_j^v[V_1^d]\}_{j \in \mathcal{J}, v \in V_1^d}$ be as defined in Lemma 9.1 with $G = \mathbb{T}_d$. Let $\overline{X}$ solve,

$$(9.2.1) \quad \overline{X}_v(t) = \varsigma(\kappa_v) + \int_{(0,t] \times \mathbb{R}_+ \times \mathcal{J}} j \mathbb{I}_{\left\{ r \le \rho_j^v[V_1^d](t, \widehat{\kappa}_{V_1^d}, \overline{X}_{V_1^d}) \right\}} \mathbf{N}_v(ds, dr, dj) \text{ for } (t, v) \in \mathbb{R}_+ \times V_1^d.$$

Again note that by Lemma 9.1, the jump rate functions of (9.2.1) satisfy conditions (1), (2), (3) and (4) of Assumption 4 with respect to the complete graph $K_{V_1^d}$. Therefore, Lemma 6.3 holds, so $\overline{X}$ is well defined up to indistinguishability, $\mathbb{F}$-adapted and a.s. proper. By Proposition 9.2, $(\widehat{\kappa}, \overline{X}) \stackrel{(\text{d})}{=} (\kappa_{V_1^d}, X_{V_1^d}) \sim \mu[V_1^d]$.

Let (κ, X) be the solution to (6.1.1) (where $G = \mathbb{T}_d$). By Assumption 4, $r_j^\varnothing(\cdot)$ satisfies conditions (1), (2), (3) and (4) of Assumption 4. Thus, for any $(t, \vartheta, x) \in \mathbb{R}_+ \times (\mathcal{K} \times \mathcal{D})^{V^d}$ we may set

$$\rho_j^{\varnothing}[V_1^d](t,\vartheta_{V_1^d},x_{V_1^d}) = r_j^{\varnothing}(t,\vartheta,x) = r_j(t,\vartheta_{V_1^d},x_{V_1^d}).$$

We start by defining the functions $\gamma := \{\gamma_j\}_{j\in\mathcal{J}}$. For each $j \in \mathcal{J}$, the function $\gamma_j : \mathbb{R}_+ \times (\mathcal{K} \times \mathcal{D})^{\{\varnothing,1\}} \to \mathbb{R}_+$ defined by

$$\gamma_j(t,\vartheta_{\{1,\varnothing\}},x_{\{1,\varnothing\}}) := \rho_j^1[\{\varnothing,1\}](t,\vartheta_{\{\varnothing,1\}},x_{\{\varnothing,1\}}), \quad (t,\vartheta,x) \in \mathbb{R}_+ \times (\mathcal{K}\times\mathcal{D})^{\{\varnothing,1\}}$$

satisfies conditions (2), (3) and (4) of Assumption 4 by Lemma 9.1. Let $\varphi : V^d \to V^d$ be any automorphism such that $\varphi(\varnothing) = 1$, $\varphi(1) = \varnothing$ and φ^2 is the identity. By the automorphism invariance of μ, $\varphi_*(\kappa, X) \overset{(\mathrm{d})}{=} (\kappa, X)$. Then for each $(t,j) \in \mathbb{R}_+ \times \mathcal{J}$, we apply the following computation justified below.

$$\begin{aligned}
\gamma_j(t,\kappa_{\{\varnothing,1\}},X_{\{\varnothing,1\}}) &= \gamma_j(t,\varphi_*(\kappa_{\{1,\varnothing\}},X_{\{1,\varnothing\}})) \\
&= \rho_j^1[\{\varnothing,1\}](t,\varphi_*(\kappa_{\{\varnothing,1\}},X_{\{\varnothing,1\}})) \\
&= \mathbb{E}\left[r_j^1(t,\varphi_*(\kappa,X))|\varphi_*(\kappa_{\{\varnothing,1\}},X_{\{\varnothing,1\}}[t])\right] \\
&= \mathbb{E}\left[r_j^1(t,\varphi_*(\kappa,X))|\kappa_{\{\varnothing,1\}},X_{\{\varnothing,1\}}[t]\right] \\
&= \mathbb{E}\left[r_j(t,\kappa_{V_1^d},X_{V_1^d})|\kappa_{\{\varnothing,1\}},X_{\{\varnothing,1\}}[t]\right] \text{ a.s..}
\end{aligned}$$

The first two equalities above follow from the definitions of φ and γ_j. The third equality combines the definition of $\rho_j^1[\{\varnothing,1\}]$ with the observation that $\varphi_*(\kappa,X) \overset{(\mathrm{d})}{=} (\kappa,X)$, and the fourth equality above follows from the observation that $\varphi^{-1}(\mathrm{cl}_1) = V_1^d$. The final equality follows from Remark 6.2 and the fact that φ is its own inverse. The computation proves that if the solution Y to (6.3.1) for the initial data $\widehat{\kappa}$ satisfies $(\widehat{\kappa},Y) \overset{(\mathrm{d})}{=} (\widehat{\kappa},\overline{X})$, then $\gamma := \{\gamma_j\}_{j\in\mathcal{J}}$ satisfies condition (2) of Definition 6.14.

Next, fix any $(j,v) \in \mathcal{J} \times \mathcal{N}_\varnothing$. Let $\varphi^v : V^d \to V^d$ be any rooted automorphism such that $\varphi^v(v) = 1$, $\varphi^v(1) = v$ and $(\varphi^v)^2$ is the identity. Again, by the automorphism-invariance of μ, $\varphi^v_*(\kappa,X) \overset{(\mathrm{d})}{=} (\kappa,X)$. Then note that for every $t \in \mathbb{R}_+$, we can make the following computation (justified below):

$$
\begin{aligned}
\gamma_j(t,\kappa_{\{v,\varnothing\}},X_{\{v,\varnothing\}}) &= \gamma_j(t,\varphi^v_*(\kappa_{\{1,\varnothing\}},X_{\{1,\varnothing\}}))\\
&= \rho^1_j[\{\varnothing,1\}](t,\varphi^v_*(\kappa_{\{\varnothing,1\}},X_{\{\varnothing,1\}}))\\
&= \mathbb{E}\left[r^1_j(t,\varphi^v_*(\kappa,X))|\varphi^v_*(\kappa_{\{\varnothing,1\}},X_{\{\varnothing,1\}}[t))\right]\\
&= \mathbb{E}\left[r^1_j(t,\varphi^v_*(\kappa,X))|\varphi^v_*(\kappa_{V_1^d},X_{V_1^d}[t))\right]\\
&= \mathbb{E}\left[r^v_j(t,\kappa,X)|\kappa_{V_1^d},X_{V_1^d}[t)\right]\\
&= \rho^v_j[V_1^d](t,\kappa_{V_1^d},X_{V_1^d}) \text{ a.s..}
\end{aligned}
$$

The first three equalities above again apply the definitions of φ,γ_j and $\rho^1_j[\{\varnothing,1\}]$ as well as the automorphism-invariance of μ (which holds by Lemma 6.3). The fourth equality makes use of the fact that $\mu_{t-}=\mathrm{Law}(\varphi^v_*(\kappa,X[t)))$ is a 2-SGMRF by Proposition 6.10. The fifth equality applies condition (5) of Assumption 4 and the final equality follows from the definition of $\rho^v_j[V_1^d]$.

The display above shows that $\gamma(\cdot,\kappa_{\{v,\varnothing\}},X_{\{v,\varnothing\}})$ and $\rho^v_j[V_1^d](\cdot,\kappa_{V_1^d},X_{V_1^d})$ are càglàd modifications of one another and therefore indistinguishable. Since $(\widehat{\kappa},\overline{X})\overset{(\mathrm{d})}{=}(\kappa_{V_1^d},X_{V_1^d})$, this implies that $\gamma(\cdot,\widehat{\kappa}_{\{v,\varnothing\}},\overline{X}_{\{v,\varnothing\}})$ and $\rho^v_j[V_1^d](t,\widehat{\kappa},\overline{X})$ are indistinguishable. It immediately follows that,

$$
\begin{aligned}
\overline{X}_v(t) &= \varsigma(\kappa_v)+\int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r\le\rho^v_j[V_1^d](t,\widehat{\kappa},\overline{X})\right\}}\mathbf{N}_v(ds,dr,dj)\quad v\in V_1^d, t\in\mathbb{R}_+,\\
&= \begin{cases}\varsigma(\kappa_v)+\int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r\le\gamma_j(s,\widehat{\kappa}_{\{v,\varnothing\}},\overline{X}_{\{v,\varnothing\}})\right\}}\mathbf{N}_v(ds,dr,dj) & v\in\mathcal{N}_\varnothing, t\in\mathbb{R}_+,\\
\varsigma(\kappa_\varnothing)+\int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r\le r_j(s,\widehat{\kappa},\overline{X})\right\}}\mathbf{N}_\varnothing(ds,dr,dj) & v=\varnothing, t\in\mathbb{R}_+.\end{cases}
\end{aligned}
$$

Thus, $Y:=\overline{X}$ is the solution to (6.3.1) for the initial data $\widehat{\kappa}$, so $(\widehat{\kappa},Y)\sim\mu[V_1^d]$. By Remark 6.16, Y is $\mathcal{H}^{\widehat{\kappa}}\vee\mathbb{H}^{\mathbf{N}}$-adapted. Therefore conditions (1), (2) and (3) of Definition 6.14 are satisfied, so Y is a strong solution to the local equation and $\mathrm{Law}(Y)=\mu[V_1^d]$. □

9.3 Proof of Uniqueness

Lemma 9.5. *Suppose the conditions of Theorem 6.17 hold. Let Y be a strong solution to the local equation for the initial data $\widehat{\kappa}$ driven by the Poisson processes $\mathbf{N}$. Then it is the pathwise unique strong solution in the sense that for any other strong solution Y' to the local equation for the initial* *data $\widehat{\kappa}$ driven by $\mathbf{N}$, $Y = Y'$ almost surely. Moreover, μ is the tiled extension of Law$(\widehat{\kappa}, Y)$.*

Proof Outline: We use a *tiling* argument to prove our result. Recall that for each $n \in \mathbb{N}$, $V_n^d = \{v \in V^d : d_{\mathbb{T}_d}(v, \varnothing) \leq n\}$ and $\mathcal{L}(V_n^d) = V_n^d \setminus V_{n-1}^d$ is the collection of leaves in $\mathbb{T}_d[V_n^d]$. Given a strong solution Y with law ν to the local equation, we construct a sequence of measures $\overline{\nu}^n \in \mathcal{P}((\mathcal{K}\times\mathcal{D})^{V_n^d})$, $n \in \mathbb{N}$, analogously to $\overline{\eta}^n$ in (8.3.2). For each $n \in \mathbb{N}$, $\overline{\nu}^n$ is characterized in terms of its Radon-Nikodym derivative with respect to $\widehat{\mu}_{t-}^{\mathbb{T}_d[V_n^d]}$ for each $t > 0$. Then using Kolmogorov's extension theorem, there exists a unique measure $\overline{\nu}$ such that $\overline{\nu}[V_n^d] = \overline{\nu}^n$ for every $n \in \mathbb{N}$, and we find that it is precisely the law of the solution to (6.1.1). By well-posedness (Lemma 6.3), it follows that $\overline{\nu} = \mu$ in which case $\nu = \overline{\nu}^1 = \overline{\nu}[V_1^d] = \mu[V_1^d]$. Lastly, the pathwise uniqueness of the strong solution follows from Remark 6.16 and Lemma 6.3. □

Proof of Theorem 6.17. Lemmas 9.4 and 9.5 together directly imply Theorem 6.17. □

Proof. Suppose that the local equation is unique in law and the law of its solution ν is strongly symmetric. Let (Y, γ) and (Y', γ') be two strong solutions to the local equation for the initial data $\widehat{\kappa}$. By Remark 6.16, (6.3.1) is equivalent to (6.1.1) with jump rate functions that satisfy Assumption 4. Then by Lemma 6.3, for fixed functions γ, (6.3.1) has a unique (up to indistinguishability) solution Y, so if $\gamma = \gamma'$, then $Y = Y'$ a.s.. Now suppose that $\gamma \neq \gamma'$. By uniqueness in law, $(\widehat{\kappa}_{\{\varnothing,1\}}, Y_{\{\varnothing,1\}}) \overset{(d)}{=} (\widehat{\kappa}_{\{\varnothing,1\}}, Y'_{\{\varnothing,1\}})$. Then by condition (2) of Definition 6.14, for any $t \in \mathbb{R}_+$ and $j \in \mathcal{J}$, $\gamma'_j(t, \widehat{\kappa}_{\{\varnothing,1\}}, Y_{\{\varnothing,1\}}) = \mathbb{E}\left[r_j(t, \widehat{\kappa}, Y) | \widehat{\kappa}_{\{\varnothing,1\}}, Y_{\{\varnothing,1\}}[t]\right] = \gamma_j(t, \widehat{\kappa}_{\{\varnothing,1\}}, Y_{\{\varnothing,1\}})$ a.s.. Thus, condition (2) of Definition 6.14 implies that $\gamma_j(\cdot, \widehat{\kappa}_{\{\varnothing,1\}}, Y_{\{\varnothing,1\}})$ and $\gamma'_j(\cdot, \widehat{\kappa}_{\{\varnothing,1\}}, Y_{\{\varnothing,1\}})$ are càglàd modifications of one another, and are therefore indistinguishable. Combined with the strong symmetry of ν, this means that for every $j \in \mathcal{J}$, $v \in \mathcal{N}_\varnothing$ and ν-a.s. $(\widehat{\vartheta}, y) \in (\mathcal{K} \times \mathcal{D})^{V_1^d}$, $\gamma_j(t, \widehat{\vartheta}_{\{v,\varnothing\}}, y_{\{v,\varnothing\}}) = \gamma'_j(t, \widehat{\vartheta}_{\{v,\varnothing\}}, y_{\{v,\varnothing\}})$ for all $t \in \mathbb{R}_+$. Thus,

$$
\begin{aligned}
Y_v(t) &= \begin{cases} \varsigma(\widehat{\kappa}_v) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r \leq \gamma_j(s, \widehat{\kappa}_{\{v,\varnothing\}}, Y_{\{v,\varnothing\}})\right\}} \mathbf{N}_v(ds, dr, dj) & \text{if } v \in \mathcal{N}_\varnothing, t \geq 0, \\ \varsigma(\widehat{\kappa}_v) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r \leq r_j(s, \widehat{\kappa}, Y)\right\}} \mathbf{N}_v(ds, dr, dj) & \text{if } v = \varnothing, t \geq 0. \end{cases} \\
&= \begin{cases} \varsigma(\widehat{\kappa}_v) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r \leq \gamma'_j(s, \widehat{\kappa}_{\{v,\varnothing\}}, Y_{\{v,\varnothing\}})\right\}} \mathbf{N}_v(ds, dr, dj) & \text{if } v \in \mathcal{N}_\varnothing, t \geq 0, \\ \varsigma(\widehat{\kappa}_v) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r \leq r_j(s, \widehat{\kappa}, Y)\right\}} \mathbf{N}_v(ds, dr, dj) & \text{if } v = \varnothing, t \geq 0, \end{cases}
\end{aligned}
$$

where the equalities hold a.s.. Thus, (Y, γ') is also a solution to the local equation, so $Y = Y'$ a.s..

For the remainder of the proof we establish the claim by proving that the solution to the local equation is unique in law. Let μ_κ be the tiled extension of ν_κ (so that $\mu_\kappa[V_1^d] = \nu_\kappa$ and μ_κ is an automorphism invariant 2-SGMRF by Proposition 6.12) and assume without loss of generality (by extending the probability space if necessary) that $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ supports a $\kappa \sim \mu_\kappa$ such that $\kappa_{V_1^d} = \widehat{\kappa}$ and a collection of $\mathbb{F}$-driving Poisson processes $\overline{\mathbf{N}} := \{\mathbf{N}_v\}_{v\in V^d}$ such that $\mathbf{N} = \{\mathbf{N}_v\}_{v\in V_1^d}$. Let (Y, γ) be the strong solution to the local equation with law ν. For any connected subgraph $G = (V, E, \varnothing)$ of $\mathbb{T}_d$, recall the definition of $\widehat{\mu}^G := (\kappa^G, \widehat{X}^G)$ from Section 9.1, where $\kappa^G = \kappa_V$. Recall that $\widehat{\mu}^G := \mathrm{Law}(\kappa^G, \widehat{X}^G)$ is well defined by Lemma 6.3. Let ν be a weak solution to the local equation. Recall that the elements of V_1^d are assigned a specific labeling: $V_1^d = \{\varnothing, 1, \dots, d\}$. Lastly, define the respective measurable functions $\Upsilon : \mathbb{R}_+ \times (\mathcal{K}\times\mathcal{D})^{V_1^d} \to \mathbb{R}_+$ and $\overline{\Upsilon} : \mathbb{R}_+ \times (\mathcal{K}\times\mathcal{D})^{\{\varnothing,1\}} \to \mathbb{R}_+$ by

$$\Upsilon(t,\vartheta,y) := \left(\prod_{t_k(y_\varnothing)\in(0,t]} r_{j_k(y_\varnothing)}(t_k(y_\varnothing),\vartheta,y)\right)\exp\left(-\sum_{j\in\mathcal{J}}\int_0^t \left(r_j(s,\vartheta,y)-1\right)\,ds\right),$$

$$\overline{\Upsilon}(t,\vartheta_{\{\varnothing,1\}},y_{\{\varnothing,1\}}) := \left(\prod_{t_k(y_\varnothing)\in(0,t]} \gamma_{j_k(y_\varnothing)}(t_k(y_\varnothing),\vartheta_{\{\varnothing,1\}},y_{\{\varnothing,1\}})\right)$$
$$\exp\left(\sum_{j\in\mathcal{J}}\int_0^t \left(\gamma_j(s,\vartheta_{\{\varnothing,1\}},y_{\{\varnothing,1\}})-1\right)\,ds\right),$$

for $(t,\vartheta,y)\in\mathbb{R}_+\times(\mathcal{K}\times\mathcal{D})^{V_1^d}$. Note that if $(\vartheta,y)\in(\mathcal{K}\times\mathcal{D})^{V^d}$, then for any $\{u,v\}\in E$ and any automorphism $\varphi : V^d\to V^d$ such that $\varphi(u)=\varnothing$ and $\varphi(v)=1$ (so that $\varphi(\mathrm{cl}_u)=V_1^d$ and $\varphi(\{u,v\})=\{\varnothing,1\}$), we may write

$$\Upsilon(t,\vartheta_{\mathrm{cl}_u},y_{\mathrm{cl}_u}) := \Upsilon(t,\varphi_*(\vartheta_{V_1^d},y_{V_1^d})),\quad \overline{\Upsilon}(t,\vartheta_{\{u,v\}},y_{\{u,v\}}) = \overline{\Upsilon}(t,\varphi_*(\vartheta_{\{\varnothing,1\}},y_{\{\varnothing,1\}})).$$

Note that by Remark 6.2, the specific choice of automorphism does not matter. By Remark 6.16 (6.3.1) is equivalent to (6.1.1) and its jump rate functions satisfy Assumption 4 with respect to graph $G := G_1 := \mathbb{T}_d[V_1^d]$ and the jump rate functions $\mathbf{r}^1_\gamma := \{r_j^\varnothing\}_{j\in\mathcal{J}}\cup\{\gamma_j(\cdot_v,\cdot_\varnothing)\}_{v\in V_1^d\setminus\{\varnothing\},j\in\mathcal{J}}$. Define $\widehat{\mu}^{G_1} := \widehat{\mu}^1$ where $\widehat{\mu}^1$ was defined in Definition 8.1 with respect to the jump rate functions $\mathbf{r}^1_\gamma$ with $G = G_1$. Then for any $t\in\mathbb{R}_+$, it immediately follows from Proposition 8.9 applied to (6.3.1) that,

$$\text{(9.3.1)} \qquad \frac{d\nu_{t-}}{d\widehat{\mu}^{G_1}_{t-}} = \Upsilon(t-, \kappa_{V_1^d}, \widehat{X}^{G_1}) \prod_{v \in \mathcal{L}(V_1^d)} \overline{\Upsilon}(t-, \kappa_{\{v,\varnothing\}}, \widehat{X}^{G_1}_{\{v,\varnothing\}}).$$

Recall that for each $v \in V^d \setminus \{\varnothing\}$, π_v is the parent of v in $\mathbb{T}_d$. Now, for each $n \in \mathbb{N}$, define the process,

$$\text{(9.3.2)} \qquad L_n(t) = \left(\prod_{v \in V_{n-1}^d} \Upsilon(t-, \kappa_{\mathrm{cl}_v}, \widehat{X}^{G_n}_{\mathrm{cl}_v}) \right) \left(\prod_{v \in \mathcal{L}(V_n^d)} \overline{\Upsilon}(t-, \kappa_{\{v,\pi_v\}}, \widehat{X}^{G_n}_{\{v,\pi_v\}}) \right).$$

For each $n \in \mathbb{N}$, we use L_n to define a measure $\overline{\nu}^n$, and we show that $\nu = \overline{\nu}^n[V_1^d]$. Let $\widehat{\mu}^{G_n} := \widehat{\mu}^n[V_n^d]$, where $\widehat{\mu}^n$ satisfies Definition 8.1 with respect to the jump rate functions $\{r_j^v\}_{v \in V^d, j \in \mathcal{J}}$ and the graph $G = \mathbb{T}_d$. Note that in the $n = 1$ case this yields an equivalent definition of $\widehat{\mu}^{G_1}$ to the definition given above.

Lemma 9.6. *For every $n \in \mathbb{N}$ and $t > 0$, there exists a measure $\overline{\nu}^n_{t-} \in \mathcal{P}((\mathcal{K} \times \mathcal{D}_{t-})^{V_n^d})$ such that,*

$$\frac{d\overline{\nu}^n_{t-}}{d\widehat{\mu}^{G_n}_{t-}} = L_n(t),$$

and for $n > 1$, $\overline{\nu}^n_{t-}[V_{n-1}^d] = \overline{\nu}^{n-1}_{t-}$.

Proof. Consider the following SDE:

$$\begin{aligned} Y'_\varnothing(t) &= \varsigma(\widehat{\kappa}_\varnothing) + \int_{(0,t] \times \mathbb{R}_+ \times \mathcal{J}} j \mathbb{I}_{\left\{r \le \gamma_j(s, \widehat{\kappa}_{\{\varnothing,1\}}, Y'_{\{\varnothing,1\}})\right\}} \mathbf{N}_\varnothing(ds, dr, dj), \\ Y'_1(t) &= \varsigma(\widehat{\kappa}_1) + \int_{(0,t] \times \mathbb{R}_+ \times \mathcal{J}} j \mathbb{I}_{\left\{r \le \gamma_j(s, \widehat{\kappa}_{\{1,\varnothing\}}, Y'_{\{1,\varnothing\}})\right\}} \mathbf{N}_1(ds, dr, dj). \end{aligned} \qquad \text{(9.3.3)}$$

Note that this SDE is of the form (6.1.1) for $G = \mathbb{T}_d[\{\varnothing, 1\}]$ and its jump rate functions clearly satisfy conditions (1) and (5) of Assumption 4. Furthermore, by condition (2) of Definition 6.14, its jump rate functions also satisfy conditions (2), (3) and (4) of Assumption 4. Therefore, the jump rate functions of the SDE above satisfy Assumption 4. By condition (2) of Definition 6.14, γ_j satisfies condition (2) of Assumption 4 so that for any $(t, j) \in \mathbb{R}_+ \times \mathcal{J}$, $\gamma_j(t, \widehat{\kappa}_{\{1,\varnothing\}}, Y_{\{1,\varnothing\}})$ is $\sigma(\widehat{\kappa}_{\{1,\varnothing\}}, Y_{\{1,\varnothing\}}[t])$-measurable. This implies that $\gamma_j(t, \widehat{\kappa}_{\{1,\varnothing\}}, Y_{\{1,\varnothing\}}) = \mathbb{E}\left[\gamma_j(t, \widehat{\kappa}_{\{1,\varnothing\}}, Y_{\{1,\varnothing\}}) | \widehat{\kappa}_{\{\varnothing,1\}}, Y_{\{\varnothing,1\}}[t]\right]$ a.s..

Furthermore, $\gamma_j(t, \widehat{\kappa}_{\{\varnothing,1\}}, Y_{\{\varnothing,1\}}) = \mathbb{E}\left[r_j(t, \widehat{\kappa}, Y)|\widehat{\kappa}_{\{\varnothing,1\}}, Y_{\{\varnothing,1\}}[t]\right]$ a.s. by condition (2) of Definition 6.14. Thus, the jump rate functions in (9.3.3) are equivalent to (9.1.1) for $G = G_1$, $U = \{\varnothing, 1\}$, $\rho_j^{\varnothing}[U] := \gamma_j$, $j \in \mathcal{J}$ and $\rho_j^1[U](t, \widehat{\vartheta}_U, y_U) := \gamma_j(t, \widehat{\vartheta}_{\{1,\varnothing\}}, y_{\{1,\varnothing\}})$, $j \in \mathcal{J}$ for $(t, \widehat{\vartheta}, y) \in \mathbb{R}_+ \times (\mathcal{K} \times \mathcal{D})^U$ and satisfy conditions (1)-(4) of Assumption 4. Then by Proposition 8.9, for any $t > 0$ and for $\widehat{\mu}^{\mathbb{T}_d[\{\varnothing,1\}]} := \widehat{\mu}^{G_1}[\{\varnothing, 1\}]$,

$$D_1(t) := \frac{d\nu_{t-}[\{\varnothing, 1\}]}{d\widehat{\mu}_{t-}^{\mathbb{T}_d[\{\varnothing,1\}]}} = \overline{\Upsilon}(t-, \kappa_{\{1,\varnothing\}}, \widehat{X}^{G_1}_{\{1,\varnothing\}})\overline{\Upsilon}(t-, \kappa_{\{\varnothing,1\}}, \widehat{X}^{G_1}_{\{\varnothing,1\}}) \tag{9.3.4}$$

Moreover, note that by (9.3.1) and (9.3.2), $L_1(t) = \frac{d\nu_{t-}}{d\widehat{\mu}^{G_1}_{t-}}$. Therefore,

$$F_1(t) := \frac{L_1(t)}{D_1(t)} = \frac{d\nu_{t-}[\{2, \dots, d\}|\{\varnothing, 1\}]}{d\widehat{\mu}^{G_1}[\{2, \dots, d\}|\{\varnothing, 1\}]}(\kappa_{\{\varnothing,1\}}, \widehat{X}^{G_1}_{\{\varnothing,1\}}), \tag{9.3.5}$$

where we apply the convention that $\frac{0}{0} = 0$. Fix $n > 1$. For each $v \in \mathcal{L}(V^d_{n-1})$, define the following process:

$$F_v(t) := \frac{\Upsilon(t-, \kappa_{\mathrm{cl}_v}, \widehat{X}^{G_n}_{\mathrm{cl}_v}) \prod_{u \in c_v} \overline{\Upsilon}(t-, \kappa_{\{u,v\}}, \widehat{X}^{G_n}_{\{u,v\}})}{\overline{\Upsilon}(t-, \kappa_{\{v,\pi_v\}}, \widehat{X}^{G_n}_{\{v,\pi_v\}}[t])},$$

where by convention we let $F_v(t) = 0$ under the same convention of $\frac{0}{0} = 0$. First note that because $\widehat{X}^{G_n}_{V^d_{n-1}} = \widehat{X}^{G_{n-1}}$ and by (9.3.2),

$$L_n(t) = L_{n-1}(t) \prod_{v \in \mathcal{L}(V^d_{n-1})} F_v(t). \tag{9.3.6}$$

Furthermore, because $\widehat{\mu}^{G_n} = \widehat{\mu}[V^d_n]$, the automorphism-invariance of $\widehat{\mu}$ (which follows from the fact that $\widehat{\mu}_\kappa = \mu_\kappa$ is automorphism invariant and by Lemma 6.3) implies that $(\kappa_{\mathrm{cl}_v}, \widehat{X}^{G_n}_{\mathrm{cl}_v}) \stackrel{(\mathrm{d})}{=} (\kappa_{V^d_1}, \widehat{X}^{G_1})$. Given that $F_v(t)$ is $\mathcal{H}^{\kappa_{\mathrm{cl}_v}} \vee \mathcal{H}_{t-}^{\widehat{X}^{G_n}_{\mathrm{cl}_v}}$-measurable, it immediately follows that $F_v(t) \stackrel{(\mathrm{d})}{=} F_1(t)$. Lastly, by Proposition 6.10, $\widehat{\mu}^{G_n}$ is a 2-SGMRF. Then for each $v \in \mathcal{L}(V^d_{n-1})$, applying the 2-SGMRF property, the fact that $F_v(t) \stackrel{(\mathrm{d})}{=} F_1(t)$ and (9.3.5),

$$\mathbb{E}\left[F_v(t)|\kappa_{V^d_{n-1}}, \widehat{X}^{G_n}_{V^d_{n-1}}[t]\right] = \mathbb{E}\left[F_v(t)|\kappa_{\{v,\pi_v\}}, \widehat{X}^{G_n}_{\{v,\pi_v\}}[t]\right] \stackrel{(\mathrm{d})}{=} \mathbb{E}\left[F_1(t)|\kappa_{\{\varnothing,1\}}, \widehat{X}^{G_1}_{\{\varnothing,1\}}[t]\right] = 1 \text{ a.s..}$$

By the 2-SGMRF property, the $\mathcal{H}^{\kappa_{\mathrm{cl}_v}} \vee \mathcal{H}_{t-}^{\widehat{X}^{G_n}_{\mathrm{cl}_v}}$-measurability of $F_v(t)$ for each $v \in \mathcal{L}(V^d_{n-1})$ implies that $\{F_v(t)\}_{v \in \mathcal{L}(V^d_{n-1})}$ are mutually conditionally independent given $\mathcal{H}^{\kappa_{V^d_{n-1}}} \vee \mathcal{H}_{t-}^{\widehat{X}^{G_n}_{V^d_{n-1}}}$. This combined with (9.3.6) and the display above yields,

$$\begin{aligned}
\mathbb{E}\left[L_n(t)|\kappa_{V^d_{n-1}}, \widehat{X}^{G_n}_{V^d_{n-1}}[t]\right] &= \mathbb{E}\left[L_{n-1}(t)\prod_{v\in\mathcal{L}(V^d_{n-1})} F_v(t)\middle|\kappa_{V^d_{n-1}}, \widehat{X}^{G_n}_{V^d_{n-1}}[t]\right] \\
&= L_{n-1}(t)\prod_{v\in\mathcal{L}(V^d_{n-1})} \mathbb{E}\left[F_v(t)|\kappa_{V^d_{n-1}}, \widehat{X}^{G_n}_{V^d_{n-1}}[t]\right] \\
&= L_{n-1}(t) \text{ a.s..} \qquad (9.3.7)
\end{aligned}$$

We finish with a brief inductive argument. We have already shown that $\frac{d\overline{\nu}^1_{t-}}{d\widehat{\mu}^{G_1}_{t-}} = \frac{d\nu_{t-}}{d\widehat{\mu}^{G_1}_{t-}} = L_1(t)$. Suppose that $\frac{d\overline{\nu}^{n-1}_{t-}}{d\widehat{\mu}^{G_{n-1}}_{t-}} = L_{n-1}(t)$. Then $\mathbb{E}\left[L_n(t)\right] = \mathbb{E}\left[\mathbb{E}\left[L_n(t)|\kappa_{V^d_n}, \widehat{X}^{G_n}_{V^d_n}[t]\right]\right] = \mathbb{E}\left[L_{n-1}(t)\right] = 1$, so $L_n(t)$ is a Radon-Nikodym derivative. If we define $\overline{\nu}^n_{t-}$ so that $L_n(t) = \frac{d\overline{\nu}^n_{t-}}{d\widehat{\mu}^{G_n}_{t-}}$, then (9.3.7) immediately implies that $\overline{\nu}^n_{t-}[V^d_{n-1}] = \overline{\nu}^{n-1}_{t-}$. This completes the proof. □

We finish the proof by constructing an $\overline{\nu}$ such that $\overline{\nu}[V^d_n] = \overline{\nu}^n$ for every $n \in \mathbb{N}$. Note that we can trivially define for each finite $U \subseteq V^d$ the measure $\overline{\nu}[U] := \overline{\nu}^n[U]$ for any n such that $U \subseteq V^d_n$, and by Lemma 9.6, this definition does not depend upon the choice of n. Thus, this family of measures satisfies the Kolmogorov consistency conditions. Therefore, by the Kolmogorov extension theorem (see [86, Theorem 2.4.3]), noting that each $\overline{\nu}[U]$ is Borel and therefore inner regular, there exists a uniqe measure $\overline{\nu} \in \mathcal{P}((\mathcal{K}\times\mathcal{D})^{V^d})$ such that $\overline{\nu}[V^d_n] = \overline{\nu}^n$ for each $n \in \mathbb{N}$. Consider the following SDE.

$$\begin{aligned}
\overline{X}^n_v(t) &= \varsigma(\kappa_v) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r\le r_j(s,\kappa_{\mathrm{cl}_v},\overline{X}^n_{\mathrm{cl}_v})\right\}}\mathbf{N}_v(ds,dr,dj) \text{ if } v\in V^d_{n-1}, \\
\overline{X}^n_v(t) &= \varsigma(\kappa_v) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r\le\gamma_j(s,\kappa_{\{v,\pi_v\}},\overline{X}^n_{\{v,\pi_v\}})\right\}}\mathbf{N}_v(ds,dr,dj) \text{ if } v\in\mathcal{L}(V^d_n).
\end{aligned}$$

It is easily verifiable that this SDE is equivalent to (6.1.1) with jump rate functions satisfying Assumption 4. Then by Proposition 8.9, for any $t > 0$, the Radon-Nikodym derivative of $\mathrm{Law}(\kappa_{V^d_n}, \overline{X}^n[t])$ with respect to the law $\widehat{\mu}^{G_n}_{t-}$ is given by $L_n(t)$, so $\overline{\nu}^n_{t-} = \mathrm{Law}(\kappa_{V^d_n}, \overline{X}^n[t])$. It follows by Proposition 8.12(a) that the dual $\overline{P}^n_{V^d_{n-1}}$ of $\overline{X}^n_{V^d_{n-1}}$ (defined in Definition 8.11) has $\mathcal{H}^{\kappa_{V^d_n}}\vee\mathbb{H}^{\overline{X}^n}$-intensity,

$$\Gamma^n(t,j,v) = r_j(t,\kappa_{\mathrm{cl}_v},\overline{X}^n_{\mathrm{cl}_v}), \quad (t,j,v)\in\mathbb{R}_+\times\mathcal{J}\times V^d_{n-1}.$$

Let $(\kappa, \overline{X}) \sim \overline{\nu}$. For any $m \in \mathbb{N}$, let $\overline{P}_{V_m^d}$ be the dual of $\overline{X}_{V_m^d}$. Note that for any $n > m+1$, $(\kappa_{V_m^d}, \overline{X}_{V_m^d}, \overline{P}_{V_m^d}) \overset{(\mathrm{d})}{=} (\kappa_{V_m^d}, \overline{X}^n_{V_m^d}, \overline{P}^n_{V_m^d})$, which immediately implies that the dual of $\overline{X}_{V_m^d}$ has $\mathcal{H}^{\kappa_{V_n^d}} \vee \mathbb{H}^{\overline{X}_{V_n^d}}$-intensity

$$\overline{\Gamma}(t, j, v)|_{v \in V_m^d} := r_j(t, \kappa_{\mathrm{cl}_v}, \overline{X}_{\mathrm{cl}_v}), \quad (t, j, v) \in \mathbb{R}_+ \times \mathcal{J} \times V_m^d.$$

Since this is true for all n sufficiently large, $\overline{\Gamma}|_{v \in V_m^d}$ is also the $\mathcal{H}^{\kappa} \vee \mathbb{H}^{\overline{X}}$-intensity of $\overline{P}_{V_m^d}$. Thus, the dual $\overline{P}$ of $\overline{X}$ has $\mathcal{H}^{\overline{\kappa}} \vee \mathbb{H}^{\overline{X}}$-intensity,

$$\overline{\Gamma}(t, j, v) = r_j(t, \kappa_{\mathrm{cl}_v}, \overline{X}_{\mathrm{cl}_v}), \quad (t, j, v) \in \mathbb{R}_+ \times \mathcal{J} \times V^d.$$

Then by Proposition 8.12(b), by extending the probability space if necessary, there exists a filtration $\mathbb{G}$ satisfying the usual conditions and a collection of i.i.d. $\mathbb{G}$-driving Poisson processes $\{\overline{\mathbf{N}}_v\}_{v \in V^d}$ such that,

$$\overline{X}_v(t) = \varsigma(\kappa_v) + \int_{(0,t] \times \mathbb{R}+ \times \mathcal{J}} j \mathbb{I}_{\{r \leq r_j(s, \kappa_{\mathrm{cl}_v}, \overline{X}_{\mathrm{cl}_v})\}} \overline{\mathbf{N}}_v(ds, dr, dj) \text{ for } v \in V^d, t \geq 0.$$

Note by Remark 6.2 that this is precisely (6.1.1), so by Lemma 6.3, $\overline{\nu} = \mathrm{Law}(\kappa, \overline{X}) = \mathrm{Law}(\kappa, X) = \mu$, so $\overline{\nu} = \mu$. By definition of $\overline{\nu}$, $\nu = \overline{\nu}[V_1^d] = \mu[V_1^d]$. By Proposition 6.10, μ is an automorphism invariant 2-SGMRF, so by Proposition 6.12, μ is the tiled extension of ν and ν is strongly symmetric. This completes the proof. $\square$

CHAPTER 10

Proofs of Other Properties Relating to the Local Equation

10.1 Proofs of Flow and Stationary Properties of the Local Equation

Recall that Corollary 6.21 established the continuity of the continuity of both the local flow Ψ and the flow generated by the IPS $\overline{\Psi}$ (see Definition 6.20) and described the relation between the two flows.

Proof of Corollary 6.21. By Theorem 6.17, whenever ν_κ is strongly symmetric there exists a unique two-sided solution $[\nu]$ to the local equation such that $[\nu]_- = \nu_\kappa$, and this solution is strongly symmetric. Likewise, by Lemma 6.3, for any μ_κ there exists a unique two-sided solution $[\mu]$ to (6.1.1) such that $[\mu]_- = \mu_\kappa$. Thus, both Ψ and $\overline{\Psi}$ are well defined. By Theorem 6.17, if ν_κ is strongly symmetric with tiled extension $\mu_\kappa := \overline{\nu}_\kappa$, then $[\mu] = \overline{[\nu]}$ is the tiled extension of $[\nu]$. It follows that for $0 \leq s \leq t < \infty$, the tiled extension of $\Psi_{s,t}(\nu_\kappa)$ is the tiled extension of $(\theta_{t-s}[\nu])_-$. By Proposition 6.10, μ_{t-s} is an automorphism invariant, 2-SGMRF such that $\mu_{t-s}[V_1^d] = \nu_{t-s}$, so it is the tiled extension of ν_{t-s}. It immediately follows that $\overline{\Psi}_{s,t}(\mu_\kappa) = (\theta_{t-s}[\mu])_-$ is the tiled extension of $(\theta_{t-s}[\nu])_- = \Psi_{s,t}(\nu_\kappa)$. By definition of the tiled extension, it follows that $\Psi_{s,t}(\nu_\kappa) = \overline{\Psi}_{s,t}(\overline{\nu}_\kappa)[V_1^d]$, so the final assertion of the corollary also holds.

The continuity of Ψ when ν_κ is not strongly symmetric follows from continuity of the identity function. When ν_κ is strongly symmetric, then continuity of Ψ follows from the fact that the shift operator is continuous and that for $Y \sim [\nu]$, there exists an a.s. positive random variable $\tau > 0$ such that $Y[0,\tau)$ is constant whenever $Y \sim [\nu]$. Thus, $(\theta_s Y)(-\infty, 0] \to Y(-\infty, 0]$ almost surely in $\overline{\mathcal{D}}_0^{V_1^d}$ as $s \searrow 0$. It follows that $\Psi_{0,s}(\nu_\kappa) = (\theta_s[\nu])_- \Rightarrow \nu_\kappa$ as $s \searrow 0$. The continuity of $\overline{\Psi}$ follows the same argument. For $X \sim [\mu]$ and any $v \in V^d$, $(\theta_s X_v)(-\infty, 0] \to X_v(-\infty, 0]$ as $s \searrow 0$. Therefore, $(\theta_s X)(-\infty, 0] \to X(-\infty, 0]$ as $s \searrow 0$, so $\overline{\Psi}_{0,s}(\mu_\kappa) = (\theta_s[\mu])_- \to \mu_\kappa$ as $s \searrow 0$.

Note that both $\overline{\Psi}$ and Ψ are time-homogeneous (candidate) flows by definition. All that remains is to show that $\overline{\Psi}$ and Ψ satisfy the semi-group property. We start with $\overline{\Psi}$ and note that by time-homogeneity, it suffices to prove that for any $s, t \geq 0$, $\overline{\Psi}_{0,t} \circ \overline{\Psi}_{0,s} = \overline{\Psi}_{0,s+t}$. Fix $s \geq 0$ and for each $v \in V^d$ let $(\theta_s \mathbf{N}_v)(dt, dr, dj) := \mathbf{N}_v(dt + s, dr, dj)$ noting that $\{\theta_s \mathbf{N}_v\}_{v \in V^d}$ is an i.i.d. collection of $\theta_s \mathbb{F} := \{\mathcal{F}_{\cdot+s}\}$-Poisson processes on $\mathbb{R}_+^2 \times \mathcal{J}$ with intensity $\mathrm{Leb}^2 \otimes \#_{\mathcal{J}}$ and that $\theta_s X(-\infty, 0]$ is $\theta_s \mathcal{F}_0$-measurable. Then by Assumption 5 note that for any $s' > s$, $v \in V^d$, $j \in \mathcal{J}$ and $x \in \overline{\mathcal{D}}^{V^d}$,

$$r_j^v(s', x[0,\infty), x(-\infty, 0]) = \breve{r}_j^v((\theta_{s'} x(-\infty, 0)) = r_j^v(s' - s, \theta_s x[0, \infty), \theta_s x(-\infty, 0]).$$

Then by the above display, for any $v \in V^d$ and $t \geq 0$,

$$\begin{aligned}
(\theta_s X_v)(t) &= X(s+t) \\
&= X(s) + \int_{(s,s+t] \times \mathbb{R}_+ \times \mathcal{J}} j \mathbb{I}_{\left\{r \leq r_j^v(s', X[0,\infty), X(-\infty,0])\right\}} \mathbf{N}_v(ds', dr, dj) \\
&= \theta_s X(0) + \int_{(0,t] \times \mathbb{R}_+ \times \mathcal{J}} j \mathbb{I}_{\left\{r \leq r_j^v(s', \theta_s X[0,\infty), \theta_s X(-\infty,0])\right\}}, \theta_s \mathbf{N}_v(ds', dr, dj).
\end{aligned}$$

Thus, by Lemma 6.3, $\mu' := \theta_s[\mu]$ is the unique two-sided solution to (6.1.1) such that $\mu'_- = (\theta_s[\mu])_-$. It immediately follows that for any $t \geq 0$, $\overline{\Psi}_{0,t} \circ \overline{\Psi}_{0,s}(\mu_\kappa) = \overline{\Psi}_{0,t}((\theta_s[\mu])_-) = (\theta_{s+t}[\mu])_- = \overline{\Psi}_{0,s+t}(\mu_\kappa)$. Given that we have already shown $\overline{\Psi}$ is time-homogeneous, it follows that $\overline{\Psi}$ is a time-homogeneous, continuous flow. Now note that Ψ trivially satisfies the semi-group property when applied to a measure that is not strongly symmetric. Let $\nu_\kappa \in \mathcal{P}(\overline{\mathcal{D}}_0^{V_1^d})$ be strongly continuous. Fix $0 \leq s, t < \infty$. Note that by the final assertion of the corollary, which we have already proven, $\overline{\Psi}_{0,s}(\overline{\nu}_\kappa)[V_1^d] = \Psi_{0,s}(\nu_\kappa)$ and its tiled extension is simply $\overline{\Psi}_{0,s}(\overline{\nu}_\kappa)$. Below, in the first equality we apply the final assertion of the corollary twice. Then we apply the argument described above that shows that

$\overline{\overline{\Psi}_{0,s}(\overline{\nu}_\kappa)[V_1^d]} = \overline{\Psi}_{0,s}(\overline{\nu}_\kappa)$. Then we use the fact that $\overline{\Psi}$ satisfies the semi-group property and apply the final assertion of the corollary one last time:

$$\Psi_{0,t} \circ \Psi_{0,s}(\nu_\kappa) = \overline{\Psi}_{0,t}(\overline{\overline{\Psi}_{0,s}(\overline{\nu}_\kappa)[V_1^d]})[V_1^d] = \overline{\Psi}_{0,t}(\overline{\Psi}_{0,s}(\overline{\nu}_\kappa))[V_1^d] = \overline{\Psi}_{0,s+t}(\overline{\nu}_\kappa)[V_1^d] = \Psi_{0,s+t}(\nu_\kappa).$$

Therefore, Ψ is also a time-homogeneous, continuous flow. □

Now we prove Corollary 6.24, which establishes that every symmetric limiting IPS has a marginal that is a symmetric limiting solution to the local equation.

Proof of Corollary 6.24. Let $\widetilde{\mu}$ be a symmetric limiting solution to (6.1.1), so that there exists an automorphism invariant, 2-SGMRF two-sided solution to (6.1.1) $[\mu]$ such that $\lim_{t\to\infty} \theta_t[\mu] = \widetilde{\mu}$. By Theorem 6.17, $[\mu][V_1^d]$ is a two-sided solution to the local equation. Thus, $\lim_{t\to\infty} \theta_t[\mu][V_1^d] = \widetilde{\mu}[V_1^d]$ is a symmetric limiting solution to the local equation. □

Finally, we prove Corollary 6.25, which established a bijection between stationary solutions to the local equation and stationary, automorphism invariant 2-SGMRF IPS.

Proof of Corollary 6.25. Let $[\nu]$ be any strongly symmetric stationary solution to the local equation. Then ν is the law of the solution to the local equation for initial data with law ν_κ. By Theorem 6.17, this implies that $\overline{\nu}$ is the law of the solution to (6.1.1) for initial data with law $\overline{\nu}_\kappa$, where $\overline{\nu}$ and $\overline{\nu}_\kappa$ are the tiled extensions of ν and ν_κ respectively. For any $t > 0$, let $\nu^t \in \mathcal{P}((\mathcal{D} \times \overline{\mathcal{D}}_0)^{V_1^d})$ and $\overline{\nu}^t \in \mathcal{P}((\mathcal{D} \times \overline{\mathcal{D}}_0)^{V_1^d})$ be the unique measures such that $[\nu^t] = \theta_t[\nu] = [\nu]$ and $[\overline{\nu}^t] = \theta_t[\overline{\nu}]$. Then by Corollary 6.21, $\overline{\nu}^t_\kappa = (\theta_t[\overline{\nu}])_- = \overline{\Psi}_{0,t}(\overline{\nu}_\kappa)$, which is the tiled extension of $\Psi_{0,t}(\nu_\kappa) = (\theta_t[\nu])_- = ([\nu])_- = \nu_\kappa$. Thus, $\overline{\nu}^t_\kappa = \overline{\nu}_\kappa$. Since this holds for all $t > 0$, $[\overline{\nu}]$ must therefore be stationary, so $[\overline{\nu}]$ is a stationary solution to (6.1.1).

In the reverse direction, if $[\mu]$ is a stationary solution to (6.1.1), then by Theorem 6.17, $[\mu][V_1^d]$ is a two-sided solution to the local equation. Moreover, $[\mu][V_1^d]$ is stationary, and therefore a stationary solution to the local equation. □

10.2 Forward Equation and Stationarity of the Markov Local Equation

We prove Proposition 6.27 in two steps. We first establish that (6.6.3) is well-posed. Then we construct a solution to the Markov local equation whose law is described by the unique solution to (6.6.3).

We begin with a general extension of the Picard-Lindelöf theorem, the proof of which is deferred to Appendix D.4. Fix $n > 1$. Let $\Delta_n := \{b \in \mathbb{R}_+^n : \sum_{i=1}^n b_i = 1\}$ be the unit simplex in $\mathbb{R}^n$, let $\Delta_n^\circ := \{b \in (0,1)^n : \sum_{i=1}^n b_i = 1\}$ be the interior of Δ_n (with respect to the plane containing Δ_n) and let $\partial\Delta_n^\circ := \Delta_n \setminus \Delta_n^\circ$ be the corresponding boundary of the unit simplex.

Lemma 10.1. *Fix any $n > 1$ and $a \in \Delta_n^\circ$. For each $i, k \in \{1, \dots, n\}$, $i \neq k$, define the bounded, continuously differentiable functions $f_{i,k} : \Delta_n^\circ \to \mathbb{R}_+$ and $g_i : \Delta_n^\circ \to \mathbb{R}_+$. Then the initial value problem*

$$\frac{dx_i}{dt} = \sum_{k \neq i} f_{i,k}(x)x_j - g_i(x)x_i,$$

$$x(0) = a \tag{10.2.1}$$

has a unique continuously differentiable solution $x : \mathbb{R}_+ \to \Delta_n^\circ$ whenever

$$\sum_{\substack{i,k=1 \\ i \neq k}}^n f_{i,k}(b)b_k - \sum_{i=1}^n g_i(b)b_i = 0 \tag{10.2.2}$$

for all $b \in \Delta_n^\circ$.

Lemma 10.2. *Suppose Assumptions 4 and 6 hold, that $|\mathcal{X}| < \infty$ and that $p_0(a) > 0$ for all $a \in \mathcal{X}^{V_1^d}$. Then there exists an $s' > 0$ such that the initial value problem* (6.6.3) *has a unique, continuously differentiable solution $p : \mathbb{R}_+ \times \mathcal{X}^{V_1^d} \to (0,1)$.*

Proof. Let $n := |\mathcal{X}^{V_1^d}| < \infty$ and let $\psi : \{1, \dots, n\} \to \mathcal{X}^{V_1^d}$ be a bijection. Throughout the proof, given the function $p : I \times \mathcal{X}^{V_1^d} \to [0,1]$ for some open interval $I \subseteq \mathbb{R}$, define $q : I \to \Delta_n$ to be the function given by $q(t) = (p(t, \psi(i)))_{i=1}^n$. Furthermore let $q_0 = (p_0(\psi(i)))_{i=1}^n \in \Delta_n^\circ$. Then it immediately follows that p is a continuously differentiable solution to (6.6.3) if and only if q is a continuously differentiable solution to (10.2.1) where $a = q_0$ and for each $i, k \in \{1, \dots, n\}$ with

$i \neq k$,

$$f_{i,k}(c) = \begin{cases} \widetilde{r}_j(\psi(k)) & \text{if } \psi(k) = S_{\varnothing,-j}(\psi(i)) \text{ for some } j \in \mathcal{J}, \\ \dfrac{\sum_{\substack{b \in \mathcal{X}^{V_1^d} \\ b_{\{\varnothing,1\}} = (\psi(i)_v - j, \psi(i)_\varnothing)}} \widetilde{r}_j(b) c_{\psi^{-1}(b)}}{\sum_{\substack{b \in \mathcal{X}^{V_1^d} \\ b_{\{\varnothing,1\}} = (\psi(i)_v - j, \psi(i)_\varnothing)}} c_{\psi^{-1}(b)}} & \text{if } \psi(k) = S_{v,-j}(\psi(i)) \text{ for some } v \in \mathcal{N}_\varnothing, j \in \mathcal{J}, \\ 0 & \text{if } \psi(k) \neq S_{v,-j}(\psi(i)) \text{ for all } v \in V_1^d, j \in \mathcal{J}, \end{cases}$$

and

$$g_i(c) = \sum_{j \in \mathcal{J}} \left(\widetilde{r}_j(\psi(i)) + \sum_{v \in \mathcal{N}_\varnothing} \frac{\sum_{\substack{b \in \mathcal{X}^{V_1^d} \\ b_{\{\varnothing,1\}} = \psi(i)_{\{v,\varnothing\}}}} \widetilde{r}_j(b) c_{\psi^{-1}(b)}}{\sum_{\substack{b \in \mathcal{X}^{V_1^d} \\ b_{\{\varnothing,1\}} = \psi(i)_{\{v,\varnothing\}}}} c_{\psi^{-1}(b)}} \right),$$

for all $c \in \Delta_n^\circ$. Then $f_{i,k}$ and g_i are clearly continuously differentiable in Δ_n° and bounded from above by $(d+1)C$ where $C := C(1)$, and where $t \mapsto C(t)$ is the constant function defined in condition (3) of Assumption 4. We now verify (10.2.2). Note that, with some justification provided below, for each $i = 1, \ldots, n$,

$$\begin{aligned} &\sum_{j \in \mathcal{J}} \sum_{v \in \mathcal{N}_\varnothing} \sum_{k=1}^n \mathbb{I}_{\{\psi(i) = S_{v,-j}(\psi(k))\}} f_{k,i}(c) c_i \\ &\quad = \sum_{j \in \mathcal{J}} \sum_{v \in \mathcal{N}_\varnothing} \sum_{k=1}^n \mathbb{I}_{\{\psi(i) = S_{v,-j}(\psi(k))\}} \frac{\sum_{\substack{b \in \mathcal{X}^{V_1^d} \\ b_{\{\varnothing,1\}} = (\psi(k)_v - j, \psi(k)_\varnothing)}} \widetilde{r}_j(b) c_{\psi^{-1}(b)}}{\sum_{\substack{b \in \mathcal{X}^{V_1^d} \\ b_{\{\varnothing,1\}} = (\psi(k)_v - j, \psi(k)_\varnothing)}} c_{\psi^{-1}(b)}} c_i \\ &\quad = \sum_{j \in \mathcal{J}} \sum_{v \in \mathcal{N}_\varnothing} \sum_{k=1}^n \mathbb{I}_{\{\psi(i) = S_{v,-j}(\psi(k))\}} \frac{\sum_{\substack{b \in \mathcal{X}^{V_1^d} \\ b_{\{\varnothing,1\}} = \psi(i)_{\{v,\varnothing\}}}} \widetilde{r}_j(b) c_{\psi^{-1}(b)}}{\sum_{\substack{b \in \mathcal{X}^{V_1^d} \\ b_{\{\varnothing,1\}} = \psi(i)_{\{v,\varnothing\}}}} c_{\psi^{-1}(b)}} c_i \\ &\quad = \sum_{j \in \mathcal{J}} \sum_{v \in \mathcal{N}_\varnothing} \frac{\sum_{\substack{b \in \mathcal{X}^{V_1^d} \\ b_{\{\varnothing,1\}} = \psi(i)_{\{v,\varnothing\}}}} \widetilde{r}_j(b) c_{\psi^{-1}(b)}}{\sum_{\substack{b \in \mathcal{X}^{V_1^d} \\ b_{\{\varnothing,1\}} = \psi(i)_{\{v,\varnothing\}}}} c_{\psi^{-1}(b)}} c_i \\ &\quad = g_i(c) c_i - c_i \sum_{j \in \mathcal{J}} \widetilde{r}_j(\psi(i)). \end{aligned}$$

In the second equality above, we use the fact that $\psi(i) = S_{v,-j}(\psi(k))$, and in the third equality we note that the term in the summand no longer depends on k. Furthermore,

$$\sum_{j\in\mathcal{J}} \mathbb{I}_{\{\psi(i)=S_{\varnothing,-j}(\psi(k))\}} f_{k,i}(c)c_i = c_i \sum_{j\in\mathcal{J}} \widetilde{r}_j(\psi(i)),$$

so for each $i \in \{1,\ldots,n\}$,

$$\sum_{k\neq i} f_{k,i}(c)c_i + g_i(c)c_i = 0.$$

Therefore,

$$\sum_{\substack{i,k=1\\ i\neq k}}^{n} f_{i,k}(c)c_k + \sum_{i=1}^{n} g_i(c)c_i = 0.$$

The result follows by Lemma 10.1. □

Before proving Proposition 6.27, we state one more technical lemma. Recall the definition of $S_{v,j}(\cdot)$ from Section 6.6. The proof of this result is relegated to Appendix D.4

Lemma 10.3. *Suppose Assumption 4 holds and let $G = (V, E)$ be a (possibly improper) subgraph of $\mathbb{T}_d$. Suppose also that there exist uniformly bounded functions $\widehat{r}^v_j : \mathbb{R}_+ \times \mathcal{X}^{cl_v} \to \mathbb{R}_+$, $v \in V, j \in \mathcal{J}$ so that for all $(t, \vartheta, x) \in \mathbb{R}_+ \times (\mathcal{K} \times \mathcal{D})^V$,*

$$r^v_j(t, \vartheta, x) = \widehat{r}^v_j(t, x_{cl_v}(t-)).$$

Then for any finite $W \subseteq V$, $U \subseteq W$, $v \in U$, $j \in \mathcal{J}$, $a \in \mathcal{X}^U$, $b \in \mathcal{X}^W$, $t \in \mathbb{R}_+$ and $\Delta t > 0$, if $cl_v \subseteq W$ and $a \neq b_U$, then the solution X^G to (6.1.1) satisfies

$$\mathbb{P}(X^G_U(t+\Delta t) = a | X^G_W(t) = b) = \sum_{v\in U}\sum_{j\in\mathcal{J}} \widehat{r}^v_j(t+, b_{cl_v})\Delta t \mathbb{I}_{\{a=S_{v,j}(b_U)\}} + O\left((\Delta t)^2\right).$$

We now prove Proposition 6.27.

Proof of Proposition 6.27. The first assertion of the Proposition holds by Lemma 10.2. Let $p : \mathbb{R}_+ \times \mathcal{X}^{V_1^d} \to (0, 1)$ be any continuous function such that $\sum_{a\in\mathcal{X}^{V_1^d}} p(t, a) = 1$ for all $t \in \mathbb{R}_+$ and $p(0, \cdot) = p_0(\cdot)$. For each $a \in \mathcal{X}^{\{\varnothing,1\}}$, $t \in \mathbb{R}_+$ and $j \in \mathcal{J}$, define

$$\widetilde{\gamma}_j(t, a) := \mathbb{E}^{Z\sim p(t,\cdot)}\left[\widetilde{r}_j(Z) | Z_{\{\varnothing,1\}} = a\right] = \frac{\sum_{\substack{b\in\mathcal{X}^{V_1^d}\\ b_{\{\varnothing,1\}}=a_{\{\varnothing,1\}}}} \widetilde{r}_j(b)p(t, b)}{\sum_{\substack{b\in\mathcal{X}^{V_1^d}\\ b_{\{\varnothing,1\}}=a_{\{\varnothing,1\}}}} p(t, b)}, \tag{10.2.3}$$

and note that for any $a \in \mathcal{X}^{\{\varnothing,1\}}$ and $j \in \mathcal{J}$, $t \mapsto \widetilde{\gamma}_j(t,a)$ is continuous. Therefore, for any $x \in \mathcal{D}^{\{\varnothing,1\}}$ and $j \in \mathcal{J}$, $t \mapsto \widetilde{\gamma}_j(t, x(t-))$ is càglàd. Therefore, $\widetilde{\gamma}$ satisfies all parts of condition 2 of Definition 6.6.2 except that it may fail to satisfy (6.6.1).

Now let $\widetilde{Y}$ solve (6.6.2) where $\widetilde{\gamma}$ is defined in (10.2.3). The equation (6.6.2) is equivalent to (6.1.1) on the graph $G = \mathbb{T}_d[V_1^d]$ and where the jump rate functions $\{r_j^v\}_{v\in V_1^d, j\in\mathcal{J}}$ are given by $r_j^{\varnothing} = r_j$ and $r_j^v(t, \vartheta_{\{v,\varnothing\}}, x_{\{v,\varnothing\}}) := \widetilde{\gamma}_j(t, x_{\{v,\varnothing\}}(t-))$. Moreover, it is easily verified that $\{r_j^v\}$ satisfy Assumption 4, so by Lemma 6.3, $\widetilde{Y}$ is uniquely defined up to indistinguishability and by Lemma 3.11, it is $\mathcal{H}^{\widehat{\kappa}} \vee \mathbb{H}^{\mathbf{N}_{V_1^d}}$-adapted. Therefore, $(\widetilde{Y}, \widetilde{\gamma}, \mathbf{N}_{V_1^d})$ satisfies conditions 1 and 3 of Definition 6.26, so $\widetilde{Y}$ is a strong solution to the Markov local equation if and only if $\widetilde{\gamma}$ satisfies (6.6.1). Moreover, if there exists a unique function $\widetilde{\gamma}$ such that both (10.2.3) and (6.6.1) are satisfied then $\widetilde{Y}$ is the only strong solution to the Markov local equation up to indistinguishability, so the Markov local equation is pathwise unique. Note that, unlike the case of the local equation, if $\widetilde{Y}$ is a solution to the Markov local equation, then $\widetilde{\gamma}$ is completely determined by $\mathrm{Law}(\widetilde{Y})$ via (6.6.1), so it suffices to prove that $\widetilde{Y}$ is the unique solution to the Markov equation in law.

Define $p' : \mathbb{R}_+ \times \mathcal{X}^{V_1^d} \to [0,1]$ by $p'(t,a) := \mathcal{P}(\widetilde{Y}(t) = a)$. Then for any $t \in \mathbb{R}_+$, $\Delta t > 0$ and $a \in \mathcal{X}^{V_1^d}$, by Lemma 10.3,

$$
\begin{aligned}
&p'(t+\Delta t, a) - p'(t,a) \\
&= \sum_{v\in V_1^d}\sum_{j\in\mathcal{J}} \mathbb{P}(a = \widetilde{Y}(t+\Delta t) = S_{v,j}(\widetilde{Y}(t))) - \mathbb{P}(\widetilde{Y}(t+\Delta t) = S_{v,j}(\widetilde{Y}(t)) = S_{v,j}(a)) + O\left((\Delta t)^2\right) \\
&= \sum_{v\in V_1^d}\sum_{j\in\mathcal{J}} \Big(\mathbb{P}(\widetilde{Y}(t+\Delta t) = S_{v,j}(\widetilde{Y}(t)) | \widetilde{Y}(t) = S_{\varnothing,-j}(a)) p'(t, S_{v,-j}(a)) \\
&\qquad - \mathbb{P}(\widetilde{Y}(t+\Delta t) = S_{v,j}(\widetilde{Y}(t)) | \widetilde{Y}(t) = a) p'(t,a)\Big) + O\left((\Delta t)^2\right) \\
&= \Delta t \sum_{j\in\mathcal{J}} \left(\widetilde{r}_j(S_{\varnothing,-j}(a)) p'(t, S_{\varnothing,-j}(a)) - \widetilde{r}_j(a) p'(t,a)\right) \\
&\qquad + \Delta t \sum_{v\in\mathcal{N}_\varnothing}\sum_{j\in\mathcal{J}} \left(\widetilde{\gamma}_j(t, (a_v - j, a_\varnothing)) p'(t, S_{v,-j}(a)) - \widetilde{\gamma}_j(t, a_{\{v,\varnothing\}}) p'(t,a)\right) + O\left((\Delta t)^2\right).
\end{aligned}
$$

Substituting the value of $\widetilde{\gamma}_j$ derived in (10.2.3), dividing by Δt and taking a limit as $\Delta t \to 0$ yields,

$$
\frac{d}{dt} p'(t,a) = \sum_{j\in\mathcal{J}} \left(\widetilde{r}_j(S_{\varnothing,-j}(a)) p'(t, S_{\varnothing,-j}(a)) - \widetilde{r}_j(a) p'(t,a)\right)
$$

(10.2.4)

$$+\sum_{v\in\mathcal{N}_\varnothing}\sum_{j\in\mathcal{J}}\left(\frac{\sum_{\substack{b\in\mathcal{X}^{V_1^d}\\ b_{\{\varnothing,1\}}=(a_v-j,a_\varnothing)}}\widetilde{r}_j(b)p(t,b)}{\sum_{\substack{b\in\mathcal{X}^{V_1^d}\\ b_{\{\varnothing,1\}}=(a_v-j,a_\varnothing)}}p(t,b)}p'(t,S_{v,-j}(a))-\frac{\sum_{\substack{b\in\mathcal{X}^{V_1^d}\\ b_{\{\varnothing,1\}}=a_{\{v,\varnothing\}}}}\widetilde{r}_j(b)p(t,b)}{\sum_{\substack{b\in\mathcal{X}^{V_1^d}\\ b_{\{\varnothing,1\}}=a_{\{v,\varnothing\}}}}p(t,b)}p'(t,a)\right).$$

Given the initial condition $p'(0,a)=p_0(a)$, (10.2.4) is a linear ODE whose coefficients vary continuously in time, which therefore has a unique solution for any p defined on the entire interval $\mathbb{R}_+$ [6, Theorem 2.25].

Note that $\widetilde{\gamma}$ satisfies both (10.2.3) and (6.6.1) if and only if for all $t\in\mathbb{R}_+$, $j\in\mathcal{J}$ and $a\in\mathcal{X}^{\{\varnothing,1\}}$,

$$\widetilde{\gamma}_j(t,a)=\frac{\sum_{\substack{b\in\mathcal{X}^{V_1^d}\\ b_{\{\varnothing,1\}}=a_{\{\varnothing,1\}}}}\widetilde{r}_j(b)p'(t,b)}{\sum_{\substack{b\in\mathcal{X}^{V_1^d}\\ b_{\{\varnothing,1\}}=a_{\{\varnothing,1\}}}}p'(t,b)}=\frac{\sum_{\substack{b\in\mathcal{X}^{V_1^d}\\ b_{\{\varnothing,1\}}=a_{\{\varnothing,1\}}}}\widetilde{r}_j(b)p(t,b)}{\sum_{\substack{b\in\mathcal{X}^{V_1^d}\\ b_{\{\varnothing,1\}}=a_{\{\varnothing,1\}}}}p(t,b)}.$$

However, if the above display holds, then (10.2.4) holds when p is replaced by p'. Given that (10.2.4) and (6.6.3) are equivalent under the additional condition that $p=p'$, it follows that p' must be the unique solution to (6.6.3). That is, $\widetilde{Y}$ solves the Markov local equation if and only if p' solves (6.6.3), so we have shown that the solution $\widetilde{Y}$ is unique in law which completes the proof. □

To prove Proposition 6.28, we first derive a forward equation describing the dynamics of a solution to (6.1.1) in the neighborhood of the root, then we apply the 2-SGMRF property and translation invariance show that π must satisfy (6.6.3).

Proof of Proposition 6.28: Let π be an automorphism invariant, 2-SGMRF stationary solution to (6.1.1). Fix a probability space $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$, driving Poisson processes $\mathbf{N}$ and an associated weak solution X to (6.1.1) for the initial condition $X(0)\sim\pi$. By stationarity, $\mathrm{Law}(X(t))=\pi$ for all $t\in\mathbb{R}_+$. Define the function $p:\mathbb{R}_+\times\mathcal{X}^{V_1^d}$ by $p(t,a)=\mathbb{P}(X_{V_1^d}(t)=a)$ for all $t\in\mathbb{R}_+$ and $a\in\mathcal{X}^{V_1^d}$.

By Lemma 10.3, for any $a,b\in\mathcal{X}^{V_1^d}$ such that $a\neq b$, $\mathbb{P}(X_{V_1^d}(t)=a|X_{V_1^d}(0)=b)=O(t^2)$. This implies that for every $t\in\mathbb{R}_+$ and $a\in\mathcal{X}^{V_1^d}$,

$$\begin{aligned}p(t,a)-p(0,a)=\sum_{v\in V_1^d}\sum_{j\in\mathcal{J}}&\left(\mathbb{P}\left(X_{V_1^d}(t)=a\middle|X_{V_1^d}(0)=S_{v,-j}(a)\right)p(0,S_{v,-j}(a)),\right.\\&\left.-\mathbb{P}\left(X_{V_1^d}(t)=S_{v,j}(a)\middle|X_{V_1^d}(0)=a\right)p(0,a)\right)+O\left(t^2\right)\end{aligned}$$

Dividing by t and taking a limit as $t \to 0$ yields,

$$\begin{aligned}\frac{d}{dt}p(0,a) &= \lim_{t\to 0}\frac{1}{t}\sum_{v\in V_1^d}\sum_{j\in\mathcal{J}}\mathbb{P}\left(X_{V_1^d}(t)=a\middle|X_{V_1^d}(0)=S_{v,-j}(a)\right)p(0,S_{v,-j}(a)) \\ &\quad -\lim_{t\to\infty}\frac{1}{t}\sum_{v\in V_1^d}\sum_{j\in\mathcal{J}}\mathbb{P}\left(X_{V_1^d}(t)=S_{v,j}(a)\middle|X_{V_1^d}(0)=a\right)p(0,a).\end{aligned} \tag{10.2.5}$$

Recall that by Lemma 10.3 (with $G=\mathbb{T}_d$, $W=V_2^d$ and $U=V_1^d$), for any $v\in V_1^d$, $j\in\mathcal{J}$, $b\in\mathcal{X}^{V_2^d}$ and $t>0$,

$$\mathbb{P}\left(X_{V_1^d}(t)=S_{v,j}(X_{V_1^d}(0))\middle|X_{V_2^d}(0)=b\right)=\widetilde{r}_j(b_{\mathrm{cl}_v})t+O\left(t^2\right). \tag{10.2.6}$$

To evaluate the terms in (10.2.5), we first apply (10.2.6). For $v\in V_1^d$, $j\in\mathcal{J}$ and $a\in\mathcal{X}^{V_1^d}$, note that $\mathrm{cl}_v\subset V_2^d$. Then,

$$\begin{aligned}\frac{1}{t}\mathbb{P}\left(X_{V_1^d}(t)=S_{v,j}(a)\middle|X_{V_1^d}(0)=a\right) &= \frac{1}{t}\sum_{\substack{b\in\mathcal{X}^{V_2^d}\\ b_{V_1^d}=a}}\mathbb{P}\left(X_{V_1^d}(t)=S_{v,j}(a)\middle|X_{V_2^d}(0)=b\right)\mathbb{P}\left(X_{V_2^d}(0)=b\middle|X_{V_1^d}(0)=a\right) \\ &\to \sum_{\substack{b\in\mathcal{X}^{V_2^d}\\ b_{V_1^d}=a}}\widetilde{r}_j(b_{\mathrm{cl}_v})\mathbb{P}\left(X_{V_2^d}(0)=b\middle|X_{V_1^d}(0)=a\right).\end{aligned} \tag{10.2.7}$$

When $v=\varnothing$, we can further simplify the above expression by noting that $\mathrm{cl}_v=\mathrm{cl}_\varnothing=V_1^d$, so

$$\sum_{\substack{b\in\mathcal{X}^{V_2^d}\\ b_{V_1^d}=a}}\widetilde{r}_j(b_{\mathrm{cl}_v})\mathbb{P}\left(X_{V_2^d}(0)=b\middle|X_{V_1^d}(0)=a\right)=\widetilde{r}_j(a)\sum_{\substack{b\in\mathcal{X}^{V_2^d}\\ b_{V_1^d}=a}}\mathbb{P}\left(X_{V_2^d}(0)=b\middle|X_{V_1^d}(0)=a\right)=\widetilde{r}_j(a). \tag{10.2.8}$$

When $v\neq\varnothing$, we apply the 2-SGMRF property, Assumption 6 and automorphism-invariance of π with respect to an automorphism that maps v to $\varnothing$ and $\varnothing$ to 1 to get,

$$\begin{aligned}
\sum_{\substack{b\in\mathcal{X}^{V_2^d}\\ b_{V_1^d}=a}} \widetilde{r}_j(b_{\mathrm{cl}_v})\mathbb{P}\left(X_{V_2^d}(0)=b\middle|X_{V_1^d}(0)=a\right) &= \sum_{\substack{b\in\mathcal{X}^{\mathrm{cl}_v}\\ b_{\{v,\varnothing\}}=a_{\{v,\varnothing\}}}} \widetilde{r}_j(b)\mathbb{P}\left(X_{\mathrm{cl}_v}(0)=b\middle|X_{V_1^d}(0)=a\right)\\
&= \sum_{\substack{b\in\mathcal{X}^{\mathrm{cl}_v}\\ b_{\{v,\varnothing\}}=a_{\{v,\varnothing\}}}} \widetilde{r}_j(b)\mathbb{P}\left(X_{\mathrm{cl}_v}(0)=b\middle|X_{\{v,\varnothing\}}(0)=a_{\{v,\varnothing\}}\right)\\
&= \sum_{\substack{b\in\mathcal{X}^{V_1^d}\\ b_{\{\varnothing,1\}}=a_{\{v,\varnothing\}}}} \widetilde{r}_j(b)\mathbb{P}\left(X_{V_1^d}(0)=b\middle|X_{\{\varnothing,1\}}(0)=a_{\{v,\varnothing\}}\right)\\
&= \frac{\sum_{\substack{b\in\mathcal{X}^{V_1^d}\\ b_{\{\varnothing,1\}}=a_{\{v,\varnothing\}}}} \widetilde{r}_j(b)p(0,b)}{\sum_{\substack{b\in\mathcal{X}^{V_1^d}\\ b_{\{\varnothing,1\}}=a_{\{v,\varnothing\}}}} p(0,b)} \qquad (10.2.9)
\end{aligned}$$

By applying stationarity of π, and combining (10.2.5) with (10.2.7),(10.2.8) and (10.2.9), it follows that p is a stationary solution to (6.6.3) as desired. $\square$

Part IV

Appendix

APPENDIX A

Measurable Representatives of Graph Isomorphism Classes

In this section, we establish several properties of the space $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ and the measurability of marked graph representatives of random isomorphism classes, culminating in the proofs of Theorem 2.29. All results are stated in Sections 2.2 and 2.3 and proven here.

A.1 Proofs of Results in Section 2.2

Let $\odot$ be an arbitrary point not lying in $\overline{\mathcal{Z}} \cup \mathcal{Z}$, define the spaces $\overline{\mathcal{Z}}_\odot := \overline{\mathcal{Z}} \cup \{\odot\}$ and $\mathcal{Z}_\odot := \mathcal{Z} \cup \{\odot\}$, and endow them with the corresponding Polish topologies from $\overline{\mathcal{Z}}$ and $\mathcal{Z}$, respectively, with $\odot$ being an isolated point. Throughout the section, we often implicitly denote the components of (possibly random) $[\overline{\mathcal{Z}}, \mathcal{Z}]$-graphs by $G := (V, E, \varnothing, \overline{\vartheta}, \vartheta)$ and $G_n := (V_n, E_n, \varnothing_n, \overline{\vartheta}^n, \vartheta^n)$.

We first begin with the proof that convergence in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z} \times \mathcal{Z}']$ is equivalent to componentwise convergence when the underlying graph converge in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$.

Proof of Lemma 2.21. Because $G_n \to G$ in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, the marked graphs (G_n, z^n), $n \in \mathbb{N}$ and (G, z), satisfy conditions 1-4 and 6 of Definition 2.19. Moreover the vertex marks of G_n converge componentwise to the vertex marks of G in the sense of condition 5 of Definition 2.19. Thus, by condition 5 of Definition 2.19, $(G_n, z^n) \to (G, z)$ if and only if the display above holds. $\square$

Next, we show that truncations of the underlying graph converge in finite time if and only if conditions 1,2,3 and 6 of Definition 2.19 hold.

Proof of Lemma 2.22. For convenience of notation, let $[n] := \{1, \ldots, n\}$. We start by proving the first statement. Fix any $m \in \mathbb{N}$, and let $M_m := \max\{v \in B_m(G)\}$. Then conditions 1 and 2 of Definition 2.19 imply that for sufficiently large n, $V \cap [M_m] = V_n \cap [M_m]$ and $E \cap [M_m]^2 = E_n \cap [M_m]^2$. Condition 3 implies that $\varnothing_n = \varnothing$ for n sufficiently large. Thus, the subgraphs of $B_m([G_*])$ and $B_m([G_{n,*}])$ induced by the set $[M_m]$ exactly match for sufficiently large n. Then by condition 6, $\max\{\mathrm{cl}_v(G_n)\} = \max\{\mathrm{cl}_v(G_n)\}$ for all $v \in V \cap [M_m]$ and n sufficiently large. Because $B_m(G) \subseteq [M_m]$, these statements imply that all of the vertices in $B_m(G_n)$ also fall inside $[M_m]$, and hence, $B_m([G_*]) = B_m([G_{n,*}])$ for all sufficiently large n.

Now suppose that $V = \{1, \ldots, |V|\}$ and $B_m([G_{n,*}]) = B_m([G_*])$ for all $m \in \mathbb{N}$ and n sufficiently large. Then for n sufficiently large, $\varnothing_n = \varnothing$ which completes the verification of condition 3. For any $v \in V$, there must exist an $m \in \mathbb{N}$ such that $v \in B_m(G)$. Then for n sufficiently large, $B_{m+1}([G_{n,*}]) = B_{m+1}([G_*])$, so $v \in V_n$ and $\mathrm{cl}_v(G_n) = \mathrm{cl}_v(G)$ so $\max\{u \in \mathrm{cl}_v(G_n)\} = \max\{u \in \mathrm{cl}_v(G_n)\}$ which completes the proof of condition 6. If $k \in \mathbb{N} \setminus V$, then $|V| < \infty$ so there must exist an M such that $B_M(G) = G$. Then for n sufficiently large that $B_{M+1}([G_{n,*}]) = B_{M+1}([G_*])$, $V_n = V$ so $k \notin V_n$. Thus condition 1 holds. Likewise, for any $e \in E$ there must exist an m such that $e \in E_{B_m(G)}$ in which case $e \in E_{B_m(G_n)}$ for sufficiently large n. Suppose $\{k, k'\} \notin E$. If either $k \notin V$ or $k' \notin V$ then we are done as we've already verified condition 1. Otherwise there exists an m such that $k, k' \in B_m(G)$ so for sufficiently large n, $e \notin E_{B_m(G)} = E_{B_m(G_n)}$. Thus, $E_n \to E$ so condition 2 also holds. □

We now prove the final useful lemma for establishing convergence in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$.

Proof of Lemma 2.23. Fix the sequences $m_n, N_n \to \infty$ such that $B_{m_n}(G_n) \to G$ and $G_n[\{1, \ldots, N_n\}] \to G$. Then conditions 1-5 of Definition 2.19 hold directly because $G_n[\{1, \ldots, N_n\}] \to G$. All that remains is to verify condition 6. However, for each $v \in V$, there must exist a $n \in \mathbb{N}$ such that $v \in V_{B_{m-1}(G)}$ so that $\mathrm{cl}_v(G) \subseteq V_{B_m(G)}$. Then because $B_{m_n}(G_n) \to G$, condition 6 of Definition 2.19 implies that,

$$\lim_{\substack{n\to\infty \\ v\in V_n}} \max\{u \in \mathrm{cl}_v(G_n)\} = \lim_{\substack{n\to\infty \\ v\in V_n \\ m_n\geq m}} \max\{u \in \mathrm{cl}_v(B_{m_n}(G_n))\} = \max\{u \in \mathrm{cl}_v(G)\},$$

so condition 6 of Definition 2.19 holds as well. $\square$

The next proof shows that $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ is Polish. Note that although this proof is long, the idea of the proof is relatively simple. We define a Polish space $\mathcal{R}$ and construct an injective map $\psi : \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}] \to \mathcal{R}$ such that the image of ψ is a closed subspace of $\mathcal{R}$. We then show that this map is continuous and that its inverse is also continuous. This is enough to show that $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ must also be Polish.

Proof of Proposition 2.25. Define the Polish space $\mathcal{R} := \left(\mathbb{N} \times \mathcal{Z}_\odot \times (\overline{\mathcal{Z}}_\odot)^{\mathbb{N}}\right)^{\mathbb{N}} \times \mathbb{N}$ equipped with the product topology, and consider the map $\psi : \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}] \to \mathcal{R}$ defined by $\psi(G) := ((\psi_k(G))_{k\in\mathbb{N}}, \varnothing)$, where for each $k \in \mathbb{N}$, $\psi_k(G) = (c'_k, \vartheta'_k, (\overline{\vartheta}'_{\{k,k'\}})_{k'\in\mathbb{N}})$, with

$$c'_k := \begin{cases} \max\{\mathrm{cl}_k(G)\} & \text{if } k \in V, \\ k & \text{otherwise,} \end{cases} \qquad \vartheta'_k := \begin{cases} \vartheta_k & \text{if } k \in V, \\ \odot & \text{otherwise,} \end{cases} \qquad \overline{\vartheta}'_{\{k,k'\}} := \begin{cases} \overline{\vartheta}_{\{k,k'\}} & \text{if } \{k,k'\} \in E, \\ \odot & \text{otherwise.} \end{cases}$$

We then have the following observations:

(i) The map ψ is a bijection from $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ to $\widehat{\mathcal{R}}$, where $\widehat{\mathcal{R}}$ is the subset of elements $\zeta = ((c'_k, \vartheta'_k, \{\overline{\vartheta}'_{\{k,k'\}}\}_{k'\in\mathbb{N}})_{k\in\mathbb{N}}, \varnothing)$ in $\mathcal{R}$ that satisfy the following constraints:

  (a) $\varnothing \in V_\zeta := \{k \in \mathbb{N} : \vartheta'_k \neq \odot\}$;

  (b) for every $k \notin V_\zeta$, $c'_k = k$;

  (c) One has $\{k, k'\} \subseteq V_\zeta$ for every $\{k, k'\} \in E_\zeta := \cup_{k\in\mathbb{N}} E_\zeta(k)$, where $E_\zeta(k) := \{\{k, k'\} \subset \mathbb{N} : k \neq k', \overline{\vartheta}'_{\{k,k'\}} \neq \odot\}$;

  (d) for every $k \in \mathbb{N}$, $c'_k = \max[\{k\} \cup \{k' \in \mathbb{N} : \{k, k'\} \in E_\zeta(k)\}]$.

  It is trivial to check that for every $G \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, $\psi(G)$ must satisfy conditions (a)–(d) above. Thus, the image of $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ under ψ is contained in $\widehat{\mathcal{R}}$. On the other hand, note that any $\zeta = ((c'_k, \vartheta'_k, \{\overline{\vartheta}'_{\{k,k'\}}\}_{k'\in\mathbb{N}})_{k\in\mathbb{N}}, \varnothing) \in \widehat{\mathcal{R}}$ has a unique inverse under ψ, described by $\psi^{-1}(\zeta) = G_\zeta := (V_\zeta, E_\zeta, \varnothing_\zeta, \overline{\vartheta}_k(\zeta), \overline{\vartheta}'_{\{k,k'\}}(\zeta))$, where V_ζ and E_ζ are respectively defined as in (a) and (c) above, $\varnothing_\zeta = \varnothing$, $\vartheta_k(\zeta) = \vartheta'_k \in \mathcal{Z}$ for $k \in V_\zeta$, and $\overline{\vartheta}_{\{k,k'\}}(\zeta) = \overline{\vartheta}'_{\{k,k'\}} \in \bar{\mathcal{Z}}$ for $\{k,k'\} \in E_\zeta$. condition (c) also ensures that $\zeta \in \widehat{\mathcal{R}}$ implies $|E_\zeta(k)| < \infty$ for every $k \in V_\zeta$, and so the resulting graph G_ζ is locally finite. Moreover, conditions (a) and (c) together ensure that the edge and vertex marks lie in $\overline{\mathcal{Z}}$ and $\mathcal{Z}$, respectively, thus showing that G_ζ is a $[\overline{\mathcal{Z}}, \mathcal{Z}]$-graph. Lastly, it is easy to see that for any $G \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ and $\zeta \in \widehat{\mathcal{R}}$, $\psi \circ \psi^{-1}(\zeta) = \zeta$ and $\psi^{-1} \circ \psi(G) = G$, thus proving ψ is a bijection between $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ and $\widehat{\mathcal{R}}$.

(ii) $\widehat{\mathcal{R}}$ is a closed subset of $\mathcal{R}$ under componentwise convergence: suppose the sequence $\zeta^n = ((c_k^{\prime,n}, \vartheta_k^{\prime,n}, \{\overline{\vartheta}^{\prime,n}_{\{k,k'\}}\}_{k' \in \mathbb{N}})_{k \in \mathbb{N}}, \varnothing^n) \in \widehat{\mathcal{R}}, n \in \mathbb{N}$, converges to $\zeta = ((c'_k, \vartheta'_k, \{\overline{\vartheta}'_{\{k,k'\}}\}_{k' \in \mathbb{N}})_{k \in \mathbb{N}}, \varnothing)$ pointwise. Then $\zeta \in \mathcal{R}$ because $\mathcal{R}$ is Polish. To show $\zeta \in \widehat{\mathcal{R}}$, it suffices to show that ζ satisfies the constraints (a)–(d) in (i) above. To prove condition (a) we argue by contradiction. Suppose $\vartheta'_\varnothing = \odot$. Then since $\varnothing^n \to \varnothing$ it follows that there exists $N < \infty$ such that $\varnothing^n = \varnothing$ for all $n \geq N$. Since $\vartheta^{\prime,n}_\varnothing \to \vartheta'_\varnothing$ and $\odot$ is isolated, this implies that for all sufficiently large n, $\vartheta^{\prime,n}_{\varnothing_n} = \vartheta^{\prime,n}_\varnothing = \odot$, which implies $\varnothing_n \notin V_{\zeta^n}$ and thus contradicts the assumption that $\zeta^n \in \widehat{\mathcal{R}}$. Thus, this proves that ζ satisfies condition (a). Condition (c) can be established in an exactly analogous fashion. Next, suppose $k \notin V_\zeta$. Then since ζ satisfies (a) as shown above, $\vartheta_k = \odot$ and since $\odot$ is isolated and $\vartheta_k^{\prime,n} \to \vartheta_k$, it follows that $\vartheta_k^{\prime,n} = \odot$ for all sufficiently large n. In turn, since $\zeta^n \in \widehat{\mathcal{R}}$, this implies $k \notin V_{\zeta^n}$ and hence that $c_k^{\prime,n} = k$. Since $c_k^{\prime,n} \to c'_k$, this implies $c'_k = k$ and condition (b) follows for ζ. Finally, condition (d) for ζ, which implies each $E_\zeta(k)$, $k \in \mathbb{N}$, is a finite set, can be deduced by similarly observing that $c_k^{\prime,n} = c_k$ and therefore $E_{\zeta^n}(k) = E_\zeta(k)$ for all $k \in \mathbb{N}$ and all sufficiently large n, and the fact that each ζ^n satisfies condition (d).

(iii) The map ψ is a homeomorphism from $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ to $\widehat{\mathcal{R}}$: For each $n \in \mathbb{N}$, let $G_n = (V_n, E_n, \varnothing_n, \overline{\vartheta}_n, \vartheta_n) \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ and $\psi(G_n) = ((c_k^{\prime,n}, \vartheta_k^{\prime,n}, \{\overline{\vartheta}^{\prime,n}_{\{k,k'\}}\}_{k' \in \mathbb{N}})_{k \in \mathbb{N}}, \varnothing^n) \in \widehat{\mathcal{R}}$. Then we wish to show that $G_n \to G = (V, E, \varnothing, \vartheta, \overline{\vartheta})$ in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ if and only if $\psi(G_n) \to \psi(G) := ((c'_k, \vartheta'_k, \{\overline{\vartheta}'_{\{k,k'\}}\}_{k' \in \mathbb{N}})_{k \in \mathbb{N}}, \varnothing)$ in $\widehat{\mathcal{R}}$.

(a) If $G_n \to G$, then $\lim_{n\to\infty} \psi(G_n) = \psi(G)$: for any $k, k' \in \mathbb{N}$, conditions 1 and 2 of Definition 2.19 imply that if $k \in V$ and $\{k,k'\} \in E$, then $k \in V_n$ and $\{k,k'\} \in E_n$ for n sufficiently large. Then for any $k \in V$, condition 6 implies that $\lim_{n\to\infty} c_k^{\prime,n} \to \lim_{\substack{n\to\infty \\ k \in V_n}} \max\{\mathrm{cl}_k(G_n)\} = \max\{\mathrm{cl}_k(G)\} = c'_k$ and for $k \notin V$, $k \notin V_n$

for n sufficiently large so $\lim_{n\to\infty} c_k^{\prime,n} = k = c_k^\prime$. Conditions 4 and 5 imply that for $k \in V$ and $\{k,k'\} \in E$, $\lim_{n\to\infty}(\vartheta_k^{\prime,n}, \overline{\vartheta}_{\{k,k'\}}^{\prime,n}) = \lim_{n\to\infty,(\vartheta_k^{\prime,n},\overline{\vartheta}_{\{k,k'\}}^{\prime,n})\in V_n\times E_n}(\vartheta_k^n, \overline{\vartheta}_{\{k,k'\}}^n) = (\vartheta_k, \overline{\vartheta}_{\{k,k'\}}) = (\vartheta_k^\prime, \overline{\vartheta}_{\{k,k'\}}^\prime)$. If $k \notin V$, then $k \notin V_n$ for n sufficiently large and because $\odot$ is an isolated point, this implies that $\lim_{n\to\infty} \vartheta_k^{\prime,n} = \odot = \vartheta_k^\prime$. By the same argument, if $\{k,k'\} \notin E$, then $\lim_{n\to\infty} \overline{\vartheta}_{\{k,k'\}}^{\prime,n} = \odot = \overline{\vartheta}_{\{k,k'\}}^\prime$. Lastly, $\varnothing^n \to \varnothing$ by condition 3 of Definition 2.19. Thus, $\psi(G_n) \to \psi(G)$, which establishes the continuity of ψ.

(b) If $\psi(G_n) \to \psi(G)$, then $G_n \to G$: conditions 1 and 2 of Definition 2.19 follow from the convergence of $\vartheta_k^{\prime,n}$ to $\vartheta_k^\prime$ and the convergence of $\overline{\vartheta}_{\{k,k'\}}^{\prime,n}$ to $\overline{\vartheta}_{\{k,k'\}}^\prime$ and the fact that $\odot$ is isolated in $\mathcal{Z}_\odot$ and $\overline{\mathcal{Z}}_\odot$. This directly implies that every vertex $k \in V$ and edge $\{k,k'\} \in E$ is in V_n and E_n respectively for all sufficiently large n, and likewise every non-vertex $k \in \mathbb{N} \setminus V$ and non-edge $\{k,k'\} \in \mathbb{N} \setminus E$ is not in V_n or E_n respectively for all sufficiently large n. Condition 3 follows from the convergence of $\varnothing^n$ to $\varnothing$. To prove condition 4, recall that we have already shown that any $k \in V$ is in V_n for n sufficiently large. Thus, $\lim_{\substack{n\to\infty \\ k\in V_n}} \vartheta_k^n = \lim_{\substack{n\to\infty \\ k\in V_n}} \vartheta_k^{\prime,n} = \lim_{n\to\infty} \vartheta_k^{\prime,n} = \vartheta_k^\prime = \vartheta_k$. The proof of condition 5 is exactly analogous except we apply the fact that $\{k,k'\} \in E$ implies $\{k,k'\} \in E_n$ for n sufficiently large and then apply the convergence of $\overline{\vartheta}_{\{k,k'\}}^{\prime,n}$ to $\overline{\vartheta}_{\{k,k'\}}^\prime$. Lastly, for each $k \in V$, because $c_k^{\prime,n} = c_k^\prime$ for sufficiently large n, this implies that $E_\zeta(k) \cap \{c_k^\prime + 1, \dots\} = \emptyset$ for n sufficiently large so $c_k^\prime \geq \max \mathrm{cl}_k(G)$. However, $\{k, c_k^\prime\} \in E$ if and only if $\{k, c_k^\prime\} \in E_n$ for all n sufficiently large, and because $\{k, c_k^{\prime,n}\} \in E_n$ and $c_k^{\prime,n} = c_k^\prime$ for n sufficiently large, this implies that $\{k, c_k^\prime\} \in E$ so $c_k^\prime = \max \mathrm{cl}_k(G)$. Thus, condition 6 holds.

Because ψ is a one-to-one map onto $\widehat{\mathcal{R}}$, this proves the claim.

Since $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ is homeomorphic to the closed subset $\widehat{\mathcal{R}}$ of the Polish space $\mathcal{R}$, it is also Polish. $\square$

A.2 Proofs of Results in Section 2.3

We begin the section by proving that the topology on $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ is compatible with the local topology. By compatible, we meant that the map $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}] \ni G \mapsto \langle G \rangle \in \mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ is continuous and its inverse is lower semi-continuous in the sense of [5, Definition 1.4.2].

Proof of Proposition 2.26. We start by proving the first statement. By Lemma 2.22, conditions 1,2,3 and 6 of Definition 2.19 imply that $B_m([G_*]) = B_m([G_{n,*}])$ for all $m \in \mathbb{N}$ and sufficiently large n, that is $N_m := \min\{n \in \mathbb{N} : B_m([G_{n',*}]) = B_m([G_*])$ for all $n' \geq n\}$ is finite (where the minimum of an empty set is taken to be ∞). For $n > N_m$, let $\varphi_{n,m} : B_m(G) \to B_m(G_n)$ be the identity isomorphism. Then for each $v \in V_{B_m(G)}$ and $e \in E_{B_m(G)}$, conditions 4 and 5 of Definition 2.19 imply $\lim_{\substack{n\to\infty \\ n>N_m}} \vartheta^n_{\varphi_{n,m}(v)} = \lim_{\substack{n\to\infty \\ v\in V_n}} \vartheta^n_v = \vartheta_v$, and $\lim_{\substack{n\to\infty \\ n>N_m}} \overline{\vartheta}^n_{\varphi_{n,m}(e)} = \lim_{\substack{n\to\infty \\ e\in E_n}} \overline{\vartheta}^n_e = \overline{\vartheta}_e$. By Definition 2.1, this proves that $\langle G_n\rangle \to \langle G\rangle$.

We now prove the second statement. For convenience of notation, let $[n] := \{1, \ldots, n\}$. First consider the case when the representative $G \in \langle G\rangle$ is connected and has a vertex set that is canonical in the sense that $V = [|V|] = \{1, \ldots, |V|\}$. It is easy to see that one can always choose connected representatives $G'_n = (V'_n, E'_n, \varnothing'_n, \overline{\vartheta}^{\prime,n}, \vartheta^{\prime,n})$ of $\langle G_n\rangle$, $n \in \mathbb{N}$, such that

$$V'_n = [|V'_n|] = \{1, \ldots, |V'_n|\} \text{ for } n \in \mathbb{N}, \tag{A.2.1}$$

where we interpret $\{1, \ldots, \infty\}$ as $\mathbb{N}$. By construction, each V'_n is also in canonical form. Then by Definition 2.1, for each $m \in \mathbb{N}_0$, there exist $n_m < \infty$ and a collection of isomorphisms $\varphi_{n,m} \in I(B_m([G_*]), B_m([G'_{n,*}]))$, $m \in \mathbb{N}_0, n > n_m$, such that for each $m \in \mathbb{N}$, the inclusions $v \in B_m(G)$ and $e \in E_{B_m(G)}$ imply

$$\lim_{\substack{n\to\infty \\ n>n_m}} d_{\overline{\mathcal{Z}}}(\overline{\vartheta}^{\prime,n}_{\varphi_{n,m}(e)}, \overline{\vartheta}_e) = 0 \quad \text{ and } \quad \lim_{\substack{n\to\infty \\ n>n_m}} d_{\mathcal{Z}}(\vartheta^{\prime,n}_{\varphi_{n,m}(v)}, \vartheta_v) = 0. \tag{A.2.2}$$

Hence, there exists a sequence of non-decreasing integers $M_n, n \in \mathbb{N}$, converging to infinity such that for each $n \in \mathbb{N}$, the inclusions $v \in B_{M_n}(G)$ and $e \in E_{B_{M_n}(G)}$ imply $d_{\mathcal{Z}}(\vartheta^{\prime,n}_{\varphi_{n,M_n}(v)}, \vartheta_v) < 2^{-M_n}$ and $d_{\overline{\mathcal{Z}}}(\overline{\vartheta}^{\prime,n}_{\varphi_{n,M_n}(e)}, \overline{\vartheta}_e) < 2^{-M_n}$ when $M_n > 0$. Set $\varphi_n := \varphi_{n,M_n}$, and for each $n \in \mathbb{N}$, define $\overline{\varphi}_n : \mathbb{N} \to \mathbb{N}$ by

$$\overline{\varphi}_n(v) = \begin{cases} \varphi_n(v) & \text{if } v \in B_{M_n}(G), \\ w^n_v & \text{otherwise,} \end{cases}$$

where if v is the kth smallest element of $\mathbb{N} \setminus B_{M_n}(G)$, then w^n_v is the kth smallest element of $\mathbb{N} \setminus B_{M_n}(G'_n)$. Now for each $n \in \mathbb{N}$, define

$$G_n := (V_n, E_n, \varnothing_n, \overline{\vartheta}^n, \vartheta^n) := \left(\overline{\varphi}_n^{-1}(V'_n), \overline{\varphi}_n^{-1}(E'_n), \overline{\varphi}_n^{-1}(\varnothing'_n), \overline{\varphi}_{n,*}^{-1}\overline{\vartheta}^{\prime,n}, \overline{\varphi}_{n,*}^{-1}\vartheta^{\prime,n}\right).$$

By construction, for each $n \in \mathbb{N}$, $V_n = V'_n = \{1, \dots, |V_n|\}$, just as in (A.2.1). Then by definition, for each $m \in \mathbb{N}$ and noting that (i) when $m \le M_n$, $\overline{\varphi}_n|_{B_m(G)} = \varphi_{n,M_n}|_{B_m(G)}$ (which implies $\overline{\varphi}_n^{-1}|_{B_m(G_n)} = \varphi_{n,M_n}^{-1}|_{B_m(G_n)}$) and (ii) M_n increases to infinity, one has

$$B_m([G_{n,*}]) = \varphi_{n,M_n}^{-1}(B_m([G'_{n,*}])) = B_m([G_*]) \text{ for } n \text{ sufficiently large.} \tag{A.2.3}$$

Then by Lemma 2.22, conditions 1-3 and 6 of Definition 2.19 hold. Now for each $v \in V$ and $e \in E$, there must exist $m \in \mathbb{N}$ such that $v \in B_m(G)$ and $e \in E_{B_m(G)}$. Then by the definition of M_n, it follows that

$$\begin{aligned}
\lim_{n\to\infty} d_{\mathcal{Z}}(\vartheta_v^n, \vartheta_v) &= \lim_{n\to\infty} d_{\mathcal{Z}}(\vartheta'^{,n}_{\varphi_{n,M_n}(v)}, \vartheta_v) = \lim_{n\to\infty} 2^{-M_n} = 0\\
\lim_{n\to\infty} d_{\mathcal{Z}}(\overline{\vartheta}_e^n, \overline{\vartheta}_e) &= \lim_{n\to\infty} d_{\mathcal{Z}}(\overline{\vartheta}'^{,n}_{\varphi_{n,M_n}(e)}, \overline{\vartheta}_e) = \lim_{n\to\infty} 2^{-M_n} = 0.
\end{aligned}$$

Thus, conditions 4 and 5 of Definition 2.19 also hold, proving that $G_n \to G$ in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$.

Now consider the more general case in which G is an arbitrary connected representative of $\langle G \rangle$ with no restriction on the vertex set. In this case, there exists a connected graph $G' := (V', E', \varnothing', \overline{\vartheta}', \vartheta') \cong G$ whose vertex set $V' := V_{G'}$ is in canonical form. Using the argument above, construct the sequence $\{G'_n\}$ whose vertex sets $V'_n := V_{G'_n}$ are in canonical form and such that $G'_n \to G'$ in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. Given any isomorphism $\varphi \in I(G', G)$, define

$$\overline{\varphi}(v) := \begin{cases} \varphi(v) & \text{if } v \in V_{G'}, \\ \\ w_v & \text{otherwise,} \end{cases}$$

where if v is the kth smallest element of $\mathbb{N} \setminus V'$, then w_v is the kth smallest element of $\mathbb{N} \setminus V$. Then $\overline{\varphi} : \mathbb{N} \to \mathbb{N}$ is a bijection, so for each $n \in \mathbb{N}$,

$$G_n := \overline{\varphi}(G'_n) := (\overline{\varphi}(V'_n), \overline{\varphi}(E'_n), \overline{\varphi}(\varnothing'_n), \overline{\varphi}_* \overline{\vartheta}'^{,n}, \overline{\varphi}_* \vartheta'^{,n}),$$

is isomorphic to G'_n and $G_n \to G$.

We finish the proof of the second assertion by considering the case in which G may not be connected. Let $G'_n := (V'_n, E'_n, \varnothing'_n, \overline{\vartheta}'^{,n}, \vartheta'^{,n})$, $n \in \mathbb{N}$, be a collection of connected graphs such that $G'_n \to \mathscr{C}_\varnothing(G)$ and let $U = \{v \in V : d_G(v, \varnothing) = \infty\}$. For each $n \in \mathbb{N}$, let $G''_n = G[U \setminus V'_n]$ and let G_n be the disjoint union of the graphs G'_n and G''_n. Because $G'_n \to \mathscr{C}_\varnothing(G)$ so that $V'_n \to V \setminus U$ and $V_{G''_n} = U \setminus V'_n$, it follows that $V_n := V_{G_n} = V'_n \cup V_{G''_n} = V'_n \cup U \to V_{\mathscr{C}_\varnothing(G)} \cup (V \setminus V_{\mathscr{C}_\varnothing(G)}) = V$ so condition 1 of Definition 2.19 holds. Likewise, $E_n := E_{G_n} = E'_n \cup E_{G''_n}$. By assumption, $E'_n \to E_{\mathscr{C}_\varnothing(G)}$ and $E_{G''_n} \to E_{G[U]}$, so $E_n = E'_n \cup E_{G''_n} \to E_{\mathscr{C}_\varnothing(G)} \cup E_{G[U]} = E$. Thus condition 2 of Definition 2.19 holds. Condition 3 of Definition 2.19 holds because $G'_n \to \mathscr{C}_\varnothing(G)$. To verify condition 6, suppose $v \in V_{\mathscr{C}_\varnothing(G)}$ in which case $v \in V'_n$ for all sufficiently large n so that $\lim_{\substack{n\to\infty \\ v \in V_n}} \max\{u \in \mathrm{cl}_{G_n}(v)\} = \lim_{\substack{n\to\infty \\ v\in V'_n}} \max\{u \in \mathrm{cl}_{G'_n}(v)\} = \max\{u \in \mathrm{cl}_{\mathscr{C}_\varnothing(G)}(v)\} = \max\{u \in \mathrm{cl}_G(v)\}$. Now suppose $v \in U$. Then $v \notin V_{\mathscr{C}_\varnothing(G)}$ so for sufficiently large n, $v \in V_{G''_n}$. Furthermore, by definition of G''_n, for any $u \in \mathrm{cl}_G(v)$ and n sufficiently large that $u \in V_{G''_n}$, $u \in \mathrm{cl}_{G''_n}(v) = \mathrm{cl}_{G_n}(v)$. Furthermore, if $u \notin \mathrm{cl}_G(v)$ and $v \in V_{G''_n}$ then by definition of G''_n, $u \notin \mathrm{cl}_{G''_n}(v) = \mathrm{cl}_{G_n}(v)$. Thus, $\lim_{\substack{n\to\infty \\ v\in V_n}} \max\{u \in \mathrm{cl}_{G_n}(v)\} = \lim_{\substack{n\to\infty \\ v\in V_{G''_n}}} \max\{u \in \mathrm{cl}_{G''_n}(v)\} = \max\{u \in \mathrm{cl}_{G[U]}(v)\} = \max\{u \in \mathrm{cl}_G(v)\}$, so condition 6 holds. Lastly, for any $v \in V$ (respy. $e \in E$) there are two cases to consider: first suppose that $v \in V_{\mathscr{C}_\varnothing(G)}$ (respy. $e \in E_{\mathscr{C}_\varnothing(G)}$). Then by the convergence of G'_n to $\mathscr{C}_\varnothing(G)$ and condition 4 (respy. condition 5) of Definition 2.19, $\lim_{\substack{n\to\infty \\ v\in V_n}} \vartheta^n_v = \lim_{\substack{n\to\infty \\ v\in V'_n}} \vartheta^n_v = \vartheta_v$ (respy. $\lim_{\substack{n\to\infty \\ e\in E_n}} \overline{\vartheta}^n_e = \lim_{\substack{n\to\infty \\ e\in E'_n}} \overline{\vartheta}^n_e = \overline{\vartheta}_e$). For the second case, suppose $v \in U$ (respy. $e \in E_{G[U]}$). Then for sufficiently large n, $\vartheta^n_v = \vartheta_v$ (respy. $\overline{\vartheta}^n_e = \overline{\vartheta}_e$) by definition of G''_n. Thus, conditions 4 and 5 of Definition 2.19 hold, so $G_n \to G$ and $\mathscr{C}_{\varnothing n}(G_n) = G'_n$ which implies $G_n \in \langle G_n \rangle$. Thus, the correspondence $\langle G \rangle \mapsto \{G' \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}] : G' \in \langle G \rangle\}$ is lower semicontinuous by [5, Definition 1.4.2]. $\square$

We split the proof of Theorem 2.29 into multiple parts. We begin with a statement of the Kuratowski and Ryll-Nardzewski measurable selection theorem which we use to prove the result. The theorem makes use of the notion of a *random closed set*.

Definition A.1. Given a Polish space $\mathcal{Z}'$ and a measurable space $(\Omega, \mathcal{F})$, a mapping $F : \Omega \mapsto \mathrm{clo}(\mathcal{Z}')$, where $\mathrm{clo}(\mathcal{Z}')$ denotes the set of closed subsets of $\mathcal{Z}'$, is said to be an $\mathcal{F}$-random closed subset of $\mathcal{Z}'$ if it is weakly measurable in the following sense: for every open set $U \subseteq \mathcal{Z}'$, the set $\{\omega \in \Omega : U \cap F(\omega) \neq \emptyset\}$ lies in $\mathcal{F}$. Moreover, F is said to be non-empty if $F(\omega) \neq \emptyset$ for every $\omega \in \Omega$.

Theorem A.2 (Kuratowski & Ryll-Nardzewski Measurable Selection Theorem)**.** *Suppose $\mathcal{Z}'$ is a Polish space, $(\Omega, \mathcal{F})$ is a measurable space and $F : \Omega \mapsto \mathrm{clo}(\mathcal{Z}')$ is a non-empty $\mathcal{F}$-random closed subset of $\mathcal{Z}'$. Then there exists an $\mathcal{F}$-measurable function $Z : \Omega \to \mathcal{Z}'$ such that $Z(\omega) \in F(\omega)$ for all $\omega \in \Omega$.*

Proof. This is simply a restatement of [16, Theorem 6.9.3] in our notation, in particular with $X, \mathcal{B}$ and Ψ in [16] replaced by $\mathcal{Z}'$, $\mathcal{F}$ and F, respectively. $\square$

We now state and prove the first assertion of Theorem 2.29.

Lemma A.3. *[36, Lemma B.7] Given any $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$-random element $\langle G\rangle$, there exists a connected measurable representative G of $\langle G\rangle$.*

Proof. For every $\langle H\rangle \in \mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, define

$$\Psi_*(\langle H\rangle) := \{H' \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}] : H' \in \langle H\rangle\}. \tag{A.2.4}$$

Then the set $\Psi_*(\langle H\rangle)$ is non-empty since it contains H, and is also closed due to the continuity of the map $H \mapsto \langle H\rangle$ established in Proposition 2.26. Thus Ψ_* maps any isomorphism class in $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ to the closed set in $\mathrm{clo}(\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}])$ composed of all graphs in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ that lie in that isomorphism class. Given the $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$-random element $\langle G\rangle$, it follows from [5, Proposition 1.4.4] that for every open set $U \subseteq \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, the following set is open in $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$:

$$\Lambda_*^U := \{\langle H\rangle \in \mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}] : \Psi_*(\langle H\rangle) \cap U \neq \emptyset\} = \bigcup_{H\in U} \{\langle H\rangle\}.$$

Since $\langle G\rangle$ is a $\mathscr{G}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$-random element, this implies that $\Psi_*(\langle G\rangle)$ is a non-empty $\sigma(\langle G\rangle)$-random closed subset of $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ in the sense of Definition A.1, and the lemma follows on applying Theorem A.2 with $\mathcal{Z}' = \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, $F = \Psi_*(\langle G\rangle)$ and $\mathcal{F} = \sigma(\langle G\rangle)$ and by noting that if the resulting $\mathcal{F}$-measurable random graph $G \in \Psi_*(\langle G\rangle)$ is disconnected, then $\mathscr{C}_\varnothing(G) \in \langle G\rangle$ is connected and it is also $\mathcal{F}$-measurable. $\square$

Before proving the rest of Theorem 2.29, we need to introduce a notion of a measurable random isomorphism. We then show that measurable random isomorphisms between isomorphic a.s. finite $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$-random elements always exist.

Remark A.4. [36, Remark B.8] For $W_1, W_2 \subseteq \mathbb{N}$, let $\mathcal{M}(W_1, W_2)$ be the space of mappings from W_1 to W_2, and define $\mathcal{M} := \bigcup_{W_1,W_2\subseteq\mathbb{N}} \mathcal{M}(W_1, W_2)$. Then any map $f \in \mathcal{M}(W_1, W_2) \subset \mathcal{M}$ can be embedded in $\mathbb{N}_0^{\mathbb{N}}$ (equipped with the discrete product topology) via the bijective map

$$\beta(f) := (\beta_n(f))_{n\in\mathbb{N}}, \qquad \text{where} \quad \beta_n(f) = \begin{cases} f(n) & \text{if } n \in W_1, \\ 0 & \text{otherwise.} \end{cases} \qquad \text{for } f \in \mathcal{M}(W_1, W_2). \tag{A.2.5}$$

We equip $\mathcal{M}$ with the topology induced by β, that is, we define a subset $U \subseteq \mathcal{M}$ to be open if and only if $\beta(U) \subseteq \mathbb{N}_0^{\mathbb{N}}$ is open. With this definition, β is automatically a homeomorphism and $\mathcal{M}$ is a Polish space.

Lemma A.5. *[36, Lemma B.9] Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that supports a.s. finite, $\mathcal{F}$-measurable $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$-random elements G_i, $i = 1, 2$, that satisfy $\mathbb{P}(G_1 \cong G_2) > 0$. Then $I(G_1, G_2)$ is an $\mathcal{F}$-random closed set. Furthermore, given any $\mathcal{F}$-random closed subset A of $\mathcal{M}$ such that A is a non-empty subset of $I(G_1, G_2)$ on the event $\{G_1 \cong G_2\}$, there exists an $\mathcal{F}$-measurable map $\varphi \in \mathcal{M}(V_1, V_2)$ such that $\varphi \in A$ on the event $\{G_1 \cong G_2\}$.*

Proof. Given any $\mathcal{F}$-random closed subset A of $\mathcal{M}$ such that $A \subseteq I(G_1, G_2)$, consider the set-valued mapping $B(\omega) := \mathcal{M}(V_1, V_2)$ if $G_1 \ncong G_2$ and $B(\omega) := A$ otherwise. Note that by finiteness of G_1 and G_2, B is also finite because $B \subseteq \mathcal{M}(V_1, V_2)$, and thus $B(\omega)$ is closed for all $\omega \in \Omega$. Let $\overline{B} = \{\beta(f) : f \in B\}$, with β as in (A.2.5), and note that $\overline{B}(\omega)$ is similarly finite and therefore closed and non-empty for all $\omega \in \Omega$. We claim that $\overline{B}(\omega)$ is weakly $\mathcal{F}$-measurable. That is, for every open $U \subseteq \mathbb{N}_0^{\mathbb{N}}$, the set $\{\overline{B} \cap U \neq \emptyset\}$ is $\mathcal{F}$-measurable. If the claim holds, then $\overline{B}$ is a non-empty $\mathcal{F}$-random closed subset, so by Theorem A.2 with $\mathcal{Z}' = \mathbb{N}_0^{\mathbb{N}}$ and $F = \overline{B}$, there exists an $\mathcal{F}$-measurable random variable b such that $b(\omega) \in \overline{B}(\omega)$ for every $\omega \in \Omega$. The lemma follows on setting $\varphi = \beta^{-1}(b)$.

To prove the claim, first note that the set $\{\omega : G_1 \cong G_2\}$ is $\mathcal{F}$-measurable by assumption. Moreover, fixing any open set $U \subseteq \mathbb{N}_0^{\mathbb{N}}$, note that $\{G_1 \cong G_2\} \cap \{\overline{B} \cap U \neq \emptyset\} = \{G_1 \cong G_2\} \cap \{A \cap \beta^{-1}(U) \neq \emptyset\}$, and also that $U' := \beta^{-1}(U) \subset \mathcal{M}$ is open since β is continuous. Because A is an $\mathcal{F}$-random closed subset, this implies $\{G_1 \cong G_2\} \cap \{\overline{B} \cap U \neq \emptyset\}$ is $\mathcal{F}$-measurable. Next, observe that $\{G_1 \ncong G_2\} \cap \{\overline{B} \cap U \neq \emptyset\} = \{G_1 \ncong G_2\} \cap \{\mathcal{M}(V_1, V_2) \cap \beta^{-1}(U) \neq \emptyset\} = \{G_1 \ncong G_2\} \cap \{\mathcal{M}(V_1, V_2) \cap U' \neq \emptyset\}$. However, note that

$$\{\mathcal{M}(V_1, V_2) \cap U' \neq \emptyset\} = \bigcup_{\text{finite } W_1, W_2 \subset \mathbb{N}} \{V_i(\omega) = W_i, i = 1, 2\} \cap \{U' \cap \mathcal{M}(W_1, W_2) \neq \emptyset\},$$

which is a countable union of $\mathcal{F}$-measurable sets since U' is open. Thus, $\{\overline{B} \cap U \neq \emptyset\}$ is $\mathcal{F}$-measurable for every open U, and the claim follows. $\square$

We may now prove Theorem 2.29.

Proof of Theorem 2.29. Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting $\langle G_n \rangle$, $n \in \mathbb{N}$ and $\langle G \rangle$, and for each $n \in \mathbb{N}$ let $\mathcal{F}_n := \sigma(\langle G \rangle) \vee \sigma(\langle G_n \rangle)$. The first assertion holds by Lemma A.3, so let G be any connected measurable representative of $\langle G\rangle$. For each $n \in \mathbb{N}$, let G_n be a connected, $\mathcal{F}_n$-measurable representative of $\langle G_n\rangle$. Then for any $\varphi \in \mathcal{M}$, define the following random variable:

$$\Psi_{n,m}(\varphi) := \begin{cases} \sum_{v \in V_{B_m(G)}} d_{\mathcal{Z}}(\vartheta^n_{\varphi(v)}, \vartheta_v) + \sum_{e \in E_{B_m(G)}} d_{\overline{\mathcal{Z}}}(\overline{\vartheta}^n_{\varphi(e)}, \overline{\vartheta}_e) & \text{if } \varphi \in I_{n,m}, \\ \infty & \text{otherwise,} \end{cases}$$

where $I_{n,m} := I(B_m([G_*]), B_m([G_{n,*}]))$. Let $\mathcal{M}^{\Psi}_{n,m} = \operatorname{argmin}_{\varphi \in \mathcal{M}} \Psi_{n,m}(\varphi)$, where we define $\mathcal{M}^{\Psi}_{n,m}$ to be empty if $\Psi_{n,m}(\varphi) = \infty$ for all $\varphi \in \mathcal{M}$. We first show that it suffices to prove the following claim:

Key claim: For each $n, m \in \mathbb{N}$, $\mathcal{M}^{\Psi}_{n,m}$ is a $\mathcal{F}_n$-random closed subset of $\mathcal{M}$ that is non-empty on the event $\{I_{n,m} \neq \emptyset\}$.

Deferring the proof of the claim, first note that given the claim, applying Lemma A.5, with $G_1 = B_m([G_*])$, $G_2 = B_m([G_{n,*}])$ and $A = \mathcal{M}^{\Psi}_{n,m}$, there exists a sequence of $\mathcal{F}_n$-measurable maps $\{\overline{\varphi}_{n,m}\}_{n,m \in \mathbb{N}}$ such that $\overline{\varphi}_{n,m} \in \mathcal{M}^{\Psi}_{n,m}$ for every $n, m \in \mathbb{N}$ on the event $\{I_{n,m} \neq \emptyset\}$. Hence, by the definition of $\mathcal{M}^{\Psi}_{n,m}$, the fact that $\overline{\varphi}_{n,m} \in I_{n.m}$ and by Definition 2.1, for any $m \in \mathbb{N}$ (where the minimum over an the empty set is assumed to be ∞),

$$\lim_{n\to\infty} \min_{\varphi \in I_{n,m}} \Psi_{n,m}(\varphi) = \lim_{n\to\infty} \Psi_{n,m}(\overline{\varphi}_{n,m}) = 0.$$

For each $n \in \mathbb{N}$, consider the $\mathcal{F}_n$-measurable random variable $M_n = \max\{m \in \mathbb{N} : \Psi_{n,m}(\overline{\varphi}_{n,m}) < 2^{-m}\}$ where the maximum of the empty set is assumed to be 0. Then the above display implies that if $\langle G_n\rangle \to \langle G\rangle$ a.s., then $M_n \to \infty$ a.s. as $n \to \infty$.

Now for each $n \in \mathbb{N}$, define the following $\mathcal{F}_n$-measurable $\mathcal{M}$-random element:

$$\overline{\varphi}_n(v) = \begin{cases} \overline{\varphi}^{-1}_{n,M_n}(v) & \text{if } v \in V_{B_{M_n}(G_n)}, \\ w_v & \text{otherwise.} \end{cases}$$

where if v is the kth smallest element of $V_n \setminus V_{B_{M_n}(G_n)}$, then w_v is the kth smallest element of $\{k \in \mathbb{N} : k > \max\{V_{B_m(G)}\}\}$. Because $\overline{\varphi}_{n,M_n} \in I(B_{M_n}([G]), B_{M_n}([G_{n,*}]))$, $\overline{\varphi}_n(v) \in V_{B_{M_n}(G)}$ if and only if $v \in V_{B_{M_n}(G_n)}$. Thus, $\overline{\varphi}_n \in \mathcal{M}(V_n, \mathbb{N})$ is a injective map. Then we can define,

$$G'_n := (V'_n, E'_n, \varnothing'_n, \overline{\vartheta}'^{,n}, \vartheta'^{,n}) := (\overline{\varphi}_n(V_n), \overline{\varphi}_n(E_n), \overline{\varphi}_n(\varnothing_n), \overline{\varphi}_{n,*}\overline{\vartheta}^n, \overline{\varphi}_{n,*}\vartheta^n).$$

It follows that $G'_n \cong G_n$ with $\overline{\varphi}_n \in I(G_n, G'_n)$ so $G'_n \in \langle G_n \rangle$ and G'_n is connected. Furthermore, by the $\mathcal{F}_n$-measurability of $\overline{\varphi}_n$ and G_n, G'_n is also $\mathcal{F}_n$-measurable. We now assume that $\langle G_n \rangle \to \langle G \rangle$ a.s. and show that $G'_n \to G$ a.s. in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. Because $M_n \to \infty$ a.s., for every $m \in \mathbb{N}$, there exists a $\mathcal{F}$-measurable random variable N_m such that $m \leq M_n$ a.s. on the event $\{n \geq N_m\}$. For $m > 0$, this implies that $B_m(\langle G'_{n,*} \rangle) = B_m \langle G_* \rangle)$ a.s. on the event $\{n \geq N_m\}$. Because N_m is a.s. finite for every $m \in \mathbb{N}$, this implies that conditions 1,2,3 and 6 of Definition 2.19 hold a.s. by Lemma 2.22. Lastly, for any $v \in V$ and $e \in E$ and for any m such that $v \in V_{B_m(G)}$ and $e \in E_{B_m(G)}$, if $n \geq N_m$, then $m \leq M_n$, so given that $\Psi_{n,m}$ is monotonically increasing in m, $d_{\mathcal{Z}}(\vartheta'^{,n}_v, \vartheta_v) = d_{\mathcal{Z}}(\vartheta^n_{\overline{\varphi}_{n,M_n}(v)}, \vartheta_v) \leq \Psi_{n,M_n}(\overline{\varphi}_{n,M_n}) < 2^{-m}$. Likewise, $d_{\overline{\mathcal{Z}}}(\overline{\vartheta}'^{,n}_e, \overline{\vartheta}_e) = d_{\overline{\mathcal{Z}}}(\overline{\vartheta}^n_{\overline{\varphi}_{n,M_n}(e)}, \overline{\vartheta}_e) \leq \Psi_{n,M_n}(\overline{\varphi}_{n,M_n}) < 2^{-m}$. Since m can be arbitrarily large, this proves conditions 4 and 5 of the definition. Thus, by Definition 2.19, $G'_n \to G$ a.s., which concludes the proof of the theorem given the key claim.

We now turn to the proof of the key claim. Fix $n, m \in \mathbb{N}$. We first prove the following:

Sub-Claim 1: For each $\varphi \in \mathcal{M}$, $\Psi_{n,m}(\varphi)$ is $\mathcal{F}_n$-measurable.

Proof of Sub-Claim 1: Fix $\varphi \in \mathcal{M}$. Define $\mathcal{A}^{\varphi}_m := \mathcal{A}_{m,1} \cap \mathcal{A}^{\varphi}_{m,2}$, where

$$\begin{aligned}
\mathcal{A}_{m,1} &:= \left\{ (H_1, H_2) \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]^2 : \max_{i=1,2, v \in V_{H_i}} d_{H_i}(\varnothing, v) \leq m \right\}, \\
\mathcal{A}^{\varphi}_{m,2} &:= \left\{ (H_1, H_2) \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]^2 : \varphi \in I(\theta(H_1), \theta(H_2)) \right\},
\end{aligned}$$

with $\theta : \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}] \to \widehat{\mathscr{G}}_*$ being the mapping that takes $H = (V_H, E_H, \varnothing_H, \overline{\vartheta}^H, \vartheta^H) \in \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$ to $(V_H, E_H, \varnothing_H) \in \widehat{\mathscr{G}}_*$, which is the rooted representative graph H with all marks removed. Then consider the mapping $\Theta^{\varphi}_m : (\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}])^2 \to \mathbb{R}_+ \cup \{\infty\}$ given by

$$\Theta^{\varphi}_m(H_1, H_2) := \begin{cases} \sum_{v \in V_{H_1}} d_{\mathcal{Z}}(\vartheta^{H_2}_{\varphi(v)}, \vartheta^{H_1}_v) + \sum_{e \in E_{H_1}} d_{\overline{\mathcal{Z}}}(\overline{\vartheta}^{H_2}_{\varphi(e)}, \overline{\vartheta}^{H_1}_e) & \text{if } (H_1, H_2) \in \mathcal{A}^{\varphi}_m, \\ \infty & \text{otherwise.} \end{cases}$$

Since $\Psi_{n,m}(\varphi) = \Theta^{\varphi}_m(B_m(G), B_m(G_n))$, to prove Sub-Claim 1, it suffices to show the following:

Sub-Claim 2: The map Θ^{φ}_m is continuous.

Proof of Sub-Claim 2: Fix $\varphi \in \mathcal{M}(W_1, W_2)$ and $m \in \mathbb{N}$. If W_1 is infinite, $\Theta_m^\varphi \equiv \infty$, and so is trivially continuous. Next, suppose $|W_1| < \infty$. Then, since θ and $H \mapsto \max_{v \in V_H} d_H(\varnothing, v)$ are continuous and $\{(H_1, H_2) \in \widehat{\mathscr{G}}_*^2 : \varphi \in I(H_1, H_2)\}$ is closed, it follows that $\mathcal{A}_{m,1}$ and $\mathcal{A}_{m,2}^\varphi$ are closed, so $\mathcal{A}_m^\varphi$ is likewise closed. Let $\{(H_1^n, H_2^n)\}_{n \in \mathbb{N}} \subseteq \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}] \subset \mathcal{A}_m^\varphi$ be any sequence that is convergent in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$, and let (H_1^∞, H_2^∞) denote its limit. Then $(H_1^\infty, H_2^\infty) \in \mathcal{A}_m^\varphi$ since $\mathcal{A}_m^\varphi$ is closed. Moreover, by conditions 1 and 2 of Definition 2.19, $V_{H_1^n} = V_{H_1^\infty} = W_1$ and $E_{H_1^n} = E_{H_1^\infty}$ for all sufficiently large n. Then the convergence of $\Theta_m^\varphi(H_1^n, H_2^n)$ to $\Theta_m^\varphi(H_1^\infty, H_2^\infty)$ is an immediate consequence of the definition of Θ_m^φ and conditions 4 and 5 of Definition 2.19. Lastly note that the map $\theta_2 : (\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}])^2 \to \widehat{\mathscr{G}}_*^2$ defined by $\theta_2(H_1, H_2) := (\theta(H_1), \theta(H_1))$ is continuous and that $A_m^\varphi = \theta_2^{-1}(\overline{A}_m^\varphi)$ where $\overline{A}^\varphi \subseteq \widehat{\mathscr{G}}_*^2$ consists entirely of pairs of graphs with radius at most m. However, it is easily verified by Definition 2.19 that any finite graph is an isolated point in $\widehat{\mathscr{G}}_*$, which implies that $\overline{A}_m^\varphi$, and therefore A_m^φ, must also be open. Sub-Claim 2 then follows on noting that Θ_m^φ is identically equal to infinity and thus trivially continuous on the closed set $(\mathcal{A}_m^\varphi)^c$.

Next, define $\Psi_{n,m}^{\min} := \min_{\varphi \in \mathcal{M}} \Psi_{n,m}(\varphi)$ where $\Psi_{n,m}^{\min} = \infty$ when $I_{n,m} = \emptyset$. Then $\Psi_{n,m}^{\min}$ always exists because $I_{n,m}$ is finite. Note that since $|B_m(G)| + |B_m(G_n)| < \infty$, it follows that

$$\Psi_{n,m}^{\min} = \min_{\substack{\varphi \in \mathcal{M}(W_1, W_2): \\ |W_1| + |W_2| < \infty}} \Psi_{n,m}(\varphi),$$

is a minimum over a countable collection of $\mathcal{F}_n$-measurable random variables. Sub-Claim 1 then shows that $\Psi_{n,m}^{\min}$ is also $\mathcal{F}_n$-measurable. For any open $U \subseteq \mathcal{M}$,

$$\{\mathcal{M}_{n,m}^\Psi \cap U \neq \emptyset\} = \bigcup_{\substack{\varphi \in \mathcal{M}(W_1, W_2) \cap U: \\ |W_1| + |W_2| < \infty}} \{\Psi_{n,m}(\varphi) = \Psi_{n,m}^{\min}\} \cap \{\Psi_{n,m}^{\min} < \infty\}.$$

Since this is a countable union of $\mathcal{F}_n$-measurable sets, the key claim follows from Definition A.1. $\square$

We conclude the appendix with a proof of Corollary 2.30. The proof is slightly involved and involves the construction of several intermediate graphs. We'll briefly summarize them here to make the proof easier to follow. Fix any $n \in \mathbb{N}$ and $i \in \{1, 2\}$.

- $(G_n, \varnothing_n^i)$ is the finite graph in the statement of the lemma such that $\langle \mathscr{C}_{\varnothing_n^i}(G_n) \rangle \to \langle G^i \rangle$.

- $(G_n'^{,i}, \widetilde{\varnothing}_n^i) \in \langle G_n, \varnothing_n^i \rangle$ and $(G^i, \varnothing^i) \in \langle G^i \rangle$ are connected, measurable representative graphs such that $(G_n^{\prime,i}, \widetilde{\varnothing}_n^i) \to G^i$ in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$.

- $(G_n^i, \widetilde{\varnothing}_n^i)$ are (possibly disconnected) extensions of $(G_n^{\prime,i}, \widetilde{\varnothing}_n^i)$ defined such that $(G_n^i, \widetilde{\varnothing}_n^i) \cong (G_n, \varnothing_n^i)$ and such that $G_n^i[\{1,\dots,2n\}] = G_n^{\prime,i}[\{1,\dots,2n\}]$.
- Letting M_n denote $1/3$ of the distance between the roots $\varnothing_n^1$ and $\varnothing_n^2$ in G_n (adjusted so M_n is an integer), $\widetilde{G}_n$ is an unrooted graph such that $\widetilde{G}_n[U_n] = G_n^2[U_n]$ where U_n is the collection of vertices within distance M_n of $\widetilde{\varnothing}_n^2$ in $\widetilde{G}_n^2$, and $\widetilde{G}_n[\mathbb{N} \setminus U_n] = G_n^1[\mathbb{N} \setminus U_n]$. Thus, $\widetilde{G}_n$ is identical to G_n^2 near $\widetilde{\varnothing}_n^2$ and it is identical to G_n^1 everywhere else. Lastly, $(\widetilde{G}_n, \widetilde{\varnothing}_n^1, \widetilde{\varnothing}_n^2) \cong (G_n, \varnothing_n^1, \varnothing_n^2)$.

For convenience, we will use the notation $[n] = \{1,\dots,n\}$ in the following proof.

Proof of Corollary 2.30. Fix $i \in \{1,2\}$. By Theorem 2.29 there exists a sequence of connected, $\mathcal{F}_n := \sigma(\langle G^i \rangle, (G_n, \varnothing_n^i) : i = 1,2) \supseteq \sigma(\langle G^i \rangle) \vee \sigma(\langle \mathscr{C}_{\varnothing_n^i}(G_n) \rangle : i = 1,2)$-measurable representatives $(G_n^{\prime,i}, \widetilde{\varnothing}_n^i) := (V_n^{\prime,i}, E_n^{\prime,i}, \widetilde{\varnothing}_n^i, \overline{\vartheta}^{\prime,i,n}, \vartheta^{\prime,i,n})$ of $\langle \mathscr{C}_{\varnothing_n^i}(G_n) \rangle$, such that $(G_n^{\prime,i}, \widetilde{\varnothing}_n^i)$ converges to the connected, $\sigma(\langle G^i \rangle : i = 1,2)$-measurable $(G^i, \varnothing^i) = (V^i, E^i, \varnothing^i, \overline{\vartheta}^i, \vartheta^i) \in \langle G^i \rangle$ a.s. in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. By mapping the vertices v to $2v + 1 - i$ for $i = 1,2$, we may assume without loss of generality that the vertices of $G_n^{\prime,1}$, $n \in \mathbb{N}$, and G^1 are restricted to the even numbers while the vertices of $G_n^{\prime,2}$, $n \in \mathbb{N}$, and G^2 are restricted to the odd numbers.

Note that for $n \in \mathbb{N}$, the graph G_n may not be connected but $G_n^{\prime,i}$ is connected, so it is possible that $(G_n^{\prime,i}, \widetilde{\varnothing}_n^i) \not\cong (G_n, \varnothing_n^i)$, and we need to do a little more work to extend $(G_n^{\prime,i}, \widetilde{\varnothing}_n^i)$ to a disconnected graph that is isomorphic to $(G_n, \varnothing_n^i)$. Recall from Lemma A.5 (where $G_1 = \mathscr{C}_{\varnothing_n^i}(G_n)$, $G_2 = (G_n^{\prime,i}, \widetilde{\varnothing}_n^i)$ and $A = I(G_1, G_2)$ is a.s. non-empty) that there exists for each $n \in \mathbb{N}$ an $\mathcal{F}_n$-measurable isomorphism $\varphi_n^{\prime,i} \in I(\mathscr{C}_{\varnothing_n^i}(G_n), (G_n^{\prime,i}, \widetilde{\varnothing}_n^i))$. Construct the vertex set V_n^i as follows. For $k \le M_{i,n} := 2n \vee (\max\{v \in V_n^{\prime,i}\} + i - 1)$, $k \in V_n^i$ if and only if $k \in V_n^{\prime,i}$. For $k > M_{i,n}$, $k \in V_{G_n^i}$ if and only if $k - M_{i,n} = 2v + i - 1$ for a vertex v of G_n that is not in the same connected component as $\varnothing_n^i$. Then V_n^1 consists purely of even vertices while V_n^2 is restricted to odd vertices. Because G_n and $G_n^{\prime,i}$ are finite graphs, it follows that $\varphi_n^i : V_n \to V_n^i$ defined by

$$\varphi_n^i(v) := \begin{cases} \varphi_n^{\prime,i}(v) & \text{if } v \in V_{\mathscr{C}_{\varnothing_n^i}(G_n)}, \\ 2v + i - 1 + M_{i,n} & \text{otherwise,} \end{cases}$$

is an $\mathcal{F}_n$-measurable bijection. Let $(G_n^i, \widetilde{\varnothing}_n^i)$ be the $\mathcal{F}_n$-measurable random element in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$

defined by,

$$(G^i_n, \widetilde{\varnothing}^i_n) := (V^i_n, \varphi^i_n(E_n), \widetilde{\varnothing}^i_n, \varphi^i_{n,*}\overline{\vartheta}^n, \varphi^i_{n,*}\vartheta^n).$$

Then $(G^i_n, \widetilde{\varnothing}^i_n) \cong (G_n, \varnothing^i_n)$, $\mathscr{C}_{\widetilde{\varnothing}^i_n}(G^i_n) = (G'^{,i}_n, \widetilde{\varnothing}^i_n)$ and for any $N \in \mathbb{N}$ and $2n \geq N$ (noting that $M_{i,n} \geq N$), the restriction of $(G^i_n, \widetilde{\varnothing}^i_n)$ to the vertices $[N]$ is equal to the restriction of $(G'^{,i}_n, \widetilde{\varnothing}^i_n)$ to the same vertex set. Since $(G'^{,i}_n, \widetilde{\varnothing}^i_n) \to (G^i, \varnothing^i)$ a.s., this implies that for any two sequences of positive integers $m_n, N_n \to \infty$ such that $N_n < 2n$ for all $n \in \mathbb{N}$, $B_{m_n}(G^i_n, \widetilde{\varnothing}^i_n) \to (G^i, \varnothing^i)$ and $(G^i_n, \widetilde{\varnothing}^i_n)[[N_n]] \to (G^i, \varnothing^i)$ a.s.. Thus, by Lemma 2.23, $(G^i_n, \widetilde{\varnothing}^i_n) \to (G^i, \varnothing^i)$ a.s. in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$.

Let $L_n := d_{G_n}(\varnothing^1_n, \varnothing^2_n)$ and $M_n := \left\lfloor \frac{1}{3} L_n \right\rfloor$, and note that L_n and $M_n \to \infty$ in probability by assumption. Recall that $\varphi^1_n \in I((G_n, \varnothing^1_n), (G^1_n, \widetilde{\varnothing}^1_n))$ is $\mathcal{F}_n$-measurable. Let $V^1_n \ni \overline{\varnothing}^2_n := \varphi^1_n(\varnothing^2_n)$. Then by definition, $B_{M_n}(G^1_n, \overline{\varnothing}^2_n) \cong B_{M_n}(G_n, \varnothing^2_n) \cong B_{M_n}(G^2_n, \widetilde{\varnothing}^2_n)$ for all $n \in \mathbb{N}$. Then applying Lemma A.5 once more (with $G_1 = B_{M_n}(G^1_n, \overline{\varnothing}^2_n)$, $G_2 = B_{M_n}(G^2_n, \widetilde{\varnothing}^2_n)$ and $A = I(G_1, G_2)$), there exists an $\mathcal{F}_n$-measurable $\overline{\varphi}_n \in I(B_{M_n}(G^1_n, \overline{\varnothing}^2_n), B_{M_n}(G^2_n, \widetilde{\varnothing}^2_n))$.

We can now construct the $\mathcal{F}_n$-measurable graphs $\widetilde{G}_n$, $n \in \mathbb{N}$, respectively. Define $\widetilde{\varphi}_n \in \mathcal{M}(V^1_n, \mathbb{N})$ by,

$$\widetilde{\varphi}_n(v) = \begin{cases} v & \text{if } d_{G^1_n}(v, \overline{\varnothing}^2_n) > M_n, \\ \overline{\varphi}_n(v) & \text{otherwise.} \end{cases}$$

Then noting that $V^1_n \subseteq \{2k : k \in \mathbb{N}\}$ while the image of $\overline{\varphi}_n$ is contained in $\{2k-1 : k \in \mathbb{N}\}$, it follows that $\widetilde{\varphi}_n$ is an injection and that for $v \in B_{M_n}(G^1_n, \overline{\varnothing}^2_n)$, $\widetilde{\varphi}_n(v) \in B_{M_n}(G^2_n, \widetilde{\varnothing}^2_n)$. Then define,

$$\widetilde{G}_n := (\widetilde{V}_n, \widetilde{E}_n, \overline{\widetilde{\vartheta}}^n, \widetilde{\vartheta}^n) := (\widetilde{\varphi}_n(V^1_n), \widetilde{\varphi}_n(E^1_n), \widetilde{\varphi}_{n,*}\overline{\vartheta}^{1,n}, \widetilde{\varphi}_{n,*}\vartheta^{1,n}).$$

Let $\{n_k\}$ be any subsequence of $\mathbb{N}$. Then because $M_n \to \infty$ in probability, there exists a subsequence $\{n_{k_\ell}\}$ on $\{n_k\}$ such that $\lim_{\ell\to\infty} M_{n_{k_\ell}} = \infty$ a.s.. Then by construction, for each $i = 1, 2$,

$$\text{(A.2.6)} \qquad \lim_{\ell\to\infty} B_{M_{n_{k_\ell}}}(\widetilde{G}_{n_{k_\ell}}, \widetilde{\varnothing}^i_{n_{k_\ell}}) = \lim_{\ell\to\infty} B_{M_{n_{k_\ell}}}(G^i_{n_{k_\ell}}, \widetilde{\varnothing}^i_{n_{k_\ell}}) = (G^i, \varnothing^i) \text{ a.s. in } \widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}].$$

Now let

$$G = G^1 \cup G^2 = (V^1 \cup V^2, E^1 \cup E^2, ((\overline{\vartheta}^1_e)_{e\in E^1}, (\overline{\vartheta}^2_e)_{e\in E^2}), ((\vartheta^1_v)_{v\in V^1}, (\vartheta^2_v)_{v\in V^2})).$$

Then for each $i = 1, 2$, $\mathscr{C}_{\varnothing^i}(G) = (G^i, \varnothing^i) \in \langle G^i \rangle$ by construction. All that remains is to prove that

$(\widetilde{G}_{n_{k_\ell}}, \widetilde{\varnothing}^i_{n_{k_\ell}}) \to (G, \varnothing^i)$ a.s. in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$.

To do that, fix any $N > \varnothing^1 \vee \varnothing^2$. Then a.s. for sufficiently large ℓ, the following hold for $i = 1, 2$:

1. $\varnothing^i_{n_{k_\ell}} = \varnothing^i < N$;

2. $M_{n_{k_\ell}} > M_N$ where for every $v \in V^i \cap [N]$, $d_G(v, \varnothing^i) \leq M_N$;

3. $B_{M_N}([(\widetilde{G}_{n_{k_\ell}}, \widetilde{\varnothing}^i_{n_{k_\ell}})_*]) = B_{M_N}([\mathscr{C}_{\varnothing^i}(G)_*])$.

For each $\ell' \in \mathbb{N}$, let $N_{\ell'}$ be defined such that the four (in)equalities hold for $N = N_{\ell'}$ and $\ell \geq \ell'$. If no such integer exists, then let $N_{\ell'} = -1$. Note that for sufficiently large ℓ, $N_\ell > -1$. Fix any $\ell \in \mathbb{N}$ such that $N_\ell > -1$ and note that by definition of $\widetilde{\varphi}_n$, $n \in \mathbb{N}$, and the fact that $M_{N_\ell} \leq M_{n_{k_\ell}}$, $B_{M_{N_\ell}}(\widetilde{G}_{n_{k_\ell}}, \widetilde{\varnothing}^i_{n_{k_\ell}}) = B_{M_{N_\ell}}(G^i_{n_{k_\ell}}, \widetilde{\varnothing}^i_{n_{k_\ell}})$. Then properties 2 and 3 above imply that for every $v \in V_{\widetilde{G}_{n_{k_\ell}}} \cap [N_\ell]$, $\min_i d_{\widetilde{G}_{n_{k_\ell}}}(v, \widetilde{\varnothing}^i_{n_{k_\ell}}) = \min_i d_G(v, \varnothing^i) \leq M_{N_\ell}$. Thus, $(\widetilde{G}_{n_{k_\ell}})[[N_\ell]] = (G^1_{n_{k_\ell}} \cup G^2_{n_{k_\ell}})[[N_\ell]]$, and by property 1 above, $\varnothing^i_{n_{k_\ell}} < N_\ell$. Lastly, for each i, $(G^i_{n_{k_\ell}}, \widetilde{\varnothing}^i_{n_{k_\ell}})[[N_\ell]] \to (G^i, \varnothing^i)$ a.s. by the a.s. convergence of $(G^i_{n_{k_\ell}}, \widetilde{\varnothing}^i_{n_{k_\ell}})$ to $(G^i, \varnothing^i)$. Noting that the unions below are disjoint, this implies that,

$$\lim_{\ell \to \infty} (\widetilde{G}_{n_{k_\ell}}, \widetilde{\varnothing}^1_{n_{k_\ell}})[[N_\ell]] = \lim_{\ell \to \infty} ((G^1_{n_{k_\ell}}, \widetilde{\varnothing}^1_{n_{k_\ell}}) \cup G^2_{n_{k_\ell}})[[N_\ell]] = ((G^1, \varnothing^1) \cup G^2) = (G, \varnothing^1).$$

By the same argument, $\lim_{\ell \to \infty} (\widetilde{G}_{n_{k_\ell}}, \widetilde{\varnothing}^2_{n_{k_\ell}})[[N_\ell]] = (G, \varnothing^2)$. Combining the above display with (A.2.6) and Lemma 2.23, $(\widetilde{G}_{n_{k_\ell}}, \widetilde{\varnothing}^i_{n_{k_\ell}}) \to (G, \varnothing^i)$ a.s. in $\widehat{\mathscr{G}}_*[\overline{\mathcal{Z}}, \mathcal{Z}]$. Thus, $(\widetilde{G}_n, \widetilde{\varnothing}^i_n) \to (G, \varnothing^i)$ in probability for each $i = 1, 2$, as desired. $\square$

APPENDIX B

The Regularity of Projected Processes

In this section we prove Lemma 9.1 and Proposition 9.2. We first establish the existence of a càglàd modification of the process $t \mapsto \mathbb{E}\left[r_j^v(t, \kappa^G, X^G)|\kappa_U^G, X_U^G[t)\right]$ by modifying a result by [82]. We then apply a few technical arguments to complete the proof of Lemma 9.1. Proposition 9.2 follows immediately by Lemma 6.3, Proposition 8.12 and some standard results from point process theory. For simplicity, given two (possibly random) functions $f, g : \mathbb{R} \to \mathbb{R}$, we say f is bounded by g if $|f(t)| \leq |g(t)|$ for all $t \in \mathbb{R}_+$.

Lemma B.1. *Let $Z := \{Z(t)\}_{t\in\mathbb{R}_+}$ be a real-valued stochastic process with càglàd (left-continuous with right limits) sample paths. Let $\mathbb{G}$ be any filtration satisfying the usual conditions defined on the same probability space as Z. If there exists a non-decreasing function $M : \mathbb{R}_+ \to \mathbb{R}_+$ such that Z is bounded by M a.s. for all $t \in \mathbb{R}_+$, then $\overline{Z} = \{\overline{Z}(t)\}_{t\in\mathbb{R}_+}$ defined by $\overline{Z}(t) = \mathbb{E}\left[Z(t)|\mathcal{G}_{t-}\right]$ has a càglàd modification.*

Proof. Let Z^+ be the stochastic process defined by $Z^+(t) = Z(t+)$ and define $\overline{Z}^+$ by $\overline{Z}^+(t) := \mathbb{E}\left[Z^+(t)|\mathcal{G}_t\right]$. If M is constant, then Z^+ is càdlàg, bounded and $\mathbb{G}$ is an increasing filtration satisfying the usual conditions, so the assumptions of [82, Theorem 6] hold. It follows that $\overline{Z}^+$ has a càdlàg modification. Now suppose M is not bounded. Note that for each $n \in \mathbb{N}$, $Z^{n,+} := \{Z^+(t\wedge n)\}_{t\in\mathbb{R}_+}$ is bounded from above by $M(n+) < \infty$, so $\overline{Z}^{n,+} := \{\mathbb{E}\left[Z^+(t \wedge n)|\mathcal{G}_t\right]\}_{t\in\mathbb{R}_+}$ has càdlàg modifications for each $n \in \mathbb{N}$. However, for each $t < n$, $\overline{Z}^{n,+}(t) = \overline{Z}^+(t)$. Thus, $\overline{Z}^+$ still has a càdlàg modification, which we call g^+.

Note that the process g defined by $g(t) := g^+(t-)$ for all $t \in \mathbb{R}_+$ is càglàd, and we will show that it is, in fact, the càglàd modification we are looking for. Fix $t > 0$ and let $\{t_k\}_{k\in\mathbb{N}}$ be any sequence of real numbers increasing to t. Because Z has càglàd sample paths, it follows that $\lim_{k\to\infty} Z(t_k+) = Z(t)$ almost surely and $\{|Z(t_k+)|\}_{k\in\mathbb{N}}$ are bounded from above by $M(t)$. Lastly, note that $\bigvee_{k\in\mathbb{N}} \mathbb{G}_{t_k} = \mathbb{G}_{t-}$. Then, by definition of g, g^+ and by [26, Theorem 9.4.8],

$$g(t) = g^+(t-) = \lim_{k\to\infty} g^+(t_k) = \lim_{k\to\infty} \mathbb{E}\left[Z^+(t_k)|\mathcal{G}_{t_k}\right] = \mathbb{E}\left[Z(t)|\mathcal{G}_{t-}\right] = \overline{Z}(t) \text{ a.s..}$$

Because this holds for all $t > 0$, g is a càglàd modification of $\overline{Z}$. □

We now prove Lemma 9.1.

Proof of Lemma 9.1. Fix any $(j, v) \in \mathcal{J} \times U$. Note that the (as yet undefined) process $t \mapsto \rho^v_j[U](t, \kappa^G_U, X^G_U)$ will be predictable with respect to the *unaugmented* filtration generated by κ^G_U and X^G_U. Let's call this filtration $\mathbb{G}$ and let $\overline{\mathbb{G}}$ be the augmented filtration associated with $\mathbb{G}$. Once we have established the existence of $\rho^v_j[U]$, we can continue to work on filtrations satisfying the usual conditions as $\rho^v_j[U]$ will also be $\overline{\mathbb{G}}$-predictable.

Let $\rho(t) := r^v_j(t, \kappa^G, X^G)$ and note that $\rho(t)$ is non-negative, a.s. càglàd and bounded by C, where C is the function defined in condition (3) of Assumption 4. For $t > 0$, let $\overline{\rho}(t) := \mathbb{E}\left[\rho(t)|\overline{\mathcal{G}}_{t-}\right] = \mathbb{E}\left[r^v_j(t, \kappa^G, X^G)|\mathcal{H}^{\kappa^G_U} \vee \mathcal{H}^{X^G_U}_{t-}\right]$. Then by Lemma B.1, $\overline{\rho}$ has a càglàd modification. This implies that $\overline{\rho}|_{t\in\mathbb{Q}^+}$, where $\mathbb{Q}^+$ is the set of non-negative rational numbers, has a unique càglàd extension to $\mathcal{D}(\mathbb{R}_+, \mathbb{R}_+)$ a.s.. Note that, although $\mathbb{G}$ may fail to be complete, it is right-continuous because $t \mapsto (\kappa^G, X^G_U(t))$ is a.s. càdlàg, piecewise constant and its points of discontinuity are isolated points (see the discussion on page 357 of [28]). Thus, $\overline{\mathbb{G}}$ is simply the completion of $\mathbb{G}$ so for each $t > 0$, $\overline{\mathcal{G}}_{t-}$ and $\mathcal{G}_{t-}$ only differ by sets of measure 0. Then, for $t \in \mathbb{Q}_+$, define

$$\widehat{\rho}(t) := \mathbb{E}\left[\rho(t)|\mathcal{G}_{t-}\right] = \overline{\rho}(t) \text{ a.s..}$$

It follows that $\widehat{\rho}|_{\mathbb{Q}_+} = \overline{\rho}|_{\mathbb{Q}_+}$ almost surely. Then there must exist a family of measurable functions $\widetilde{\rho}_t : (\mathcal{K} \times \mathcal{D}_{t-})^U \to \mathbb{R}_+$, $t \in \mathbb{Q}_+$, such that

$$\widetilde{\rho}_t(\kappa_U^G, X_U^G[t]) = \widehat{\rho}(t) = \overline{\rho}(t) \text{ for all } t \in \mathbb{Q}_+ \text{ a.s.}.$$

Then we can define $\widetilde{\rho} : \mathbb{R}_+ \times (\mathcal{K} \times \mathcal{D})^U \to \mathbb{R}_+$ by

$$\widetilde{\rho}(t, \vartheta, x) := \liminf_{\substack{s \nearrow t \\ s \in \mathbb{Q}_+}} \widetilde{\rho}_s(\vartheta, x[s]). \tag{B.0.1}$$

Note that $t \mapsto \widetilde{\rho}(t, \kappa_U^G, X_U^G)$ is $\mathbb{G}$-adapted and a.s. càglàd. To account for events of probability 0, we need a predictable method to determine when $t \mapsto \widetilde{\rho}(t, \vartheta, x)$ will fail to be càglàd or fail to be bounded by C for some $(\vartheta, x) \in (\mathcal{K} \times \mathcal{D})^U$. To do this, we take advantage of the fact that $\mathcal{X}^U$ is a countable set so that the discontinuities of x must consist entirely of isolated points. For any $t \in \mathbb{R}_+$ and $x \in \mathcal{D}^U$, define the *look-forward* process $x^t(s) := x(s)$ if $s < t$ and $x^t(s) := x(t-)$ otherwise. For any fixed $\delta > 0$, define

$$z(t, \vartheta, x) = \mathbb{I}_{\left\{s \mapsto \widetilde{\rho}(s, \vartheta, x^t) \text{ is càglàd and bounded by } C \text{ on } [0, t+\delta)\right\}}.$$

Fix any $(\vartheta, x) \in (\mathcal{K} \times \mathcal{D})^U$. We will first show that $t \mapsto z(t, \vartheta, x)$ is càglàd. Then we will show that $z(\cdot, \kappa_U^G, X_U^G)$ is a modification of the constant function $t \mapsto 1$. Fix any t such that $z(t, \vartheta, x) = 1$. Then $\widetilde{\rho}(\cdot, \vartheta, x^t)$ must be bounded and càglàd on $[0, t+\delta)$. Because x is càdlàg on a discrete space, there must exist an ϵ such that $x[t-\epsilon, t)$ is constant. For any $t' \in [t-\epsilon, t)$, note that $x^{t'} = x^t$ so $\widetilde{\rho}(\cdot, \vartheta, x^{t'})$ is also well defined, bounded and càglàd on $[0, t'+\delta) \subset [0, t+\delta)$. Thus, $z(t-, \vartheta, x) = 1$. Likewise, if $z(t, \vartheta, x) = 0$, then there must exist an $\epsilon' < \epsilon$ such that $\widetilde{\rho}(\cdot, \vartheta, x^t)$ is either not bounded by C or not càglàd on the interval $[0, (t-\epsilon')+\delta)$. Then for $t' \in [t-\epsilon', t)$, $\widetilde{\rho}(s, \vartheta, x^{t'}) = \widetilde{\rho}(s, \vartheta, x^t)$ is likewise not bounded by C or not càglàd on $[0, t'+\delta)$, so $z(t-, \vartheta, x) = 0$. Thus, z is left-continuous.

To prove the existence of right limits of $z(\cdot, \vartheta, x)$, define $x^{t+}(s) := x(s)$ if $s \leq t$ and $x^{t+}(s) := x(t)$ otherwise. If $\widetilde{\rho}(s, \vartheta, x^{t+})$ is bounded by C and càglàd on $[0, t+\delta+\epsilon)$ for some $\epsilon > 0$, then $z(t+, \vartheta, x) = 1$ exists. If $\widetilde{\rho}(s, \vartheta, x^{t+})$ is either greater than C or not càglàd on $[0, t+\delta+\epsilon)$ for every $\epsilon > 0$, then there must exist an $s \leq t+\delta$ such that either $\widetilde{\rho}(s-, \vartheta, x^{t+}) \neq \widetilde{\rho}(s, \vartheta, x^{t+})$, $\widetilde{\rho}(s+, \vartheta, x^{t+})$ doesn't exist or $\widetilde{\rho}(s, \vartheta, x^{t+}) > C(s)$. However, in all of these cases, it follows that for each $t' > t$ such that x is constant on $[t, t')$, $x^{t'} = x^{t+}$ so $z(t', \vartheta, x) = 0$, which implies that $z(t+, \vartheta, x) = 0$.

Lastly we show that for every $t \in \mathbb{R}_+$, $z(t, \kappa_U^G, X_U^G) = 1$ almost surely. Fix any $t \in \mathbb{R}_+$ and suppose there exists an $A \in \mathcal{B}((\mathcal{K} \times \mathcal{D})^U)$ such that for every $(\vartheta, x) \in A$, $z(t, \vartheta, x) = 0$ and $\mathbb{P}((\kappa_U^G, X_U^G) \in A) = \mu^G[U](A) > 0$. Then define A^δ as follows:

$$A^\delta = \{(\vartheta, x) \in (\mathcal{K} \times \mathcal{D})^U : x[t+\delta) = (x')^t[t+\delta) \text{ for some } (\vartheta, x') \in A\}.$$

Conditioned on $\{(\kappa_U^G, X_U^G) \in A\}$, $\{(\kappa_U^G, X_U^G) \in A^\delta\} = \{X_U^G[t, t+\delta) \equiv X_U^G(t-)\} \subseteq \{\mathbf{N}_v([t, t+\delta) \times (0, C(t+\delta)) \times \mathcal{J}) = 0, v \in U\}$. Then $\mu^G[U](A^\delta) = \mathbb{P}((\kappa_U^G, X_U^G) \in A^\delta | (\kappa_U^G, X_U^G) \in A)\mu^G[U](A) \geq e^{-\delta|\mathcal{J}|C(t+\delta)}\mu^G[U](A) > 0$, and for $(\vartheta, x) \in A^\delta$, $\widetilde{\rho}(\cdot, \vartheta, x)$ is either unbounded by C or not càglàd on the interval $[0, t+\delta)$. Thus, with positive probability on the interval $[0, t+\delta)$, $\widetilde{\rho}(\cdot, \kappa_U^G, X_U^G)$ is either unbounded by C or it is not càglàd. This is a contradiction, so $z(t, \kappa_U^G, X_U^G) = 1$ a.s.. Thus,

$$\rho_j^v[U](t, \vartheta_U, x_U) = \widetilde{\rho}(t, \vartheta_U, x_U) z(t, \vartheta_U, x_U)$$

is càglàd for all $(\vartheta_U, x_U) \in (\mathcal{K} \times \mathcal{D})^U$ and $\rho_j^v[U](t, \kappa_U^G, X_U^G)$ is a modification of $\overline{\rho}(t)$ that is bounded by C, and therefore satisfies conditions (3) and (4) of Assumption 4.

Note that by (B.0.1) and the display above it, for each $(t, \vartheta, x), (t, \vartheta, y) \in \mathbb{R}_+ \times (\mathcal{K} \times \mathcal{D})^U$ such that $x(s) = y(s)$ for all $s < t$, $\widetilde{\rho}(t, \vartheta, x) = \widetilde{\rho}(t, \vartheta, y)$. Likewise, $x^t = y^t$, so $z(t, \vartheta, x) = z(t, \vartheta, y)$ as well. Therefore, $\rho_j^v[U](t, \vartheta, x) = \rho_j^v[U](t, \vartheta, y)$ so condition (2) of Assumption 4 is also satisfied. Finally, condition (1) holds trivially with respect to the graph K_U. This completes the proof. □

We prove Proposition 9.2 by applying a projection lemma from [28] on point processes, by repeatedly applying duality as described in Proposition 8.12 and by calling on the well-posedness of the SDE.

Proof of Proposition 9.2. Let P^G be the dual of X_U^G. Then by Proposition 8.12(a), P^G is a point process on $\mathbb{R}_+ \times \mathcal{J} \times U$ with $\mathbb{F}$-intensity,

$$\Lambda(t, j, v) := r_j^v(t, \kappa^G, X^G) \text{ for } (t, j, v) \in \mathbb{R}_+ \times \mathcal{J} \times U.$$

Let $\mathbb{G} = \mathcal{H}^{\kappa^G} \vee \mathbb{H}^{X^G}$. By Lemma 9.1, $t \mapsto \overline{\Lambda}(t, j, v) := \rho_j^v[U](t, \kappa_U^G, X_U^G)$ is a càglàd modification of $t \mapsto \mathbb{E}\left[r_j^v(t, \kappa^G, X^G) | \kappa_U^G, X_U^G[t)\right]$ for each $(j, v) \in \mathcal{J} \times U$ and $\rho_j^v[U]$ is bounded by the function C

defined in condition (3) of Assumption 4. Then by [28, Lemma 14.3.III], $\overline{\Lambda}$ is a $\mathbb{G}$-intensity of P^G.

By Proposition 8.12(b), the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ can be expanded to include a filtration $\widehat{\mathbb{G}}$ and a collection of i.i.d. $\widehat{\mathbb{G}}$-Poisson processes $\overline{\mathbf{N}} := (\overline{\mathbf{N}}_v)_{v\in U} \overset{\text{(d)}}{=} (\mathbf{N}_v)_{v\in U}$ such that for all $(t,v) \in \mathbb{R}_+ \times U$,

$$X_v^G(t) = \varsigma(\kappa_v^G) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r\le\rho_j^v[U](t,\kappa_U^G,X_U^G)\right\}}\,\overline{\mathbf{N}}_v(ds,dr,dj) \text{ for } v\in U. \tag{B.0.2}$$

Moreover, by Lemma 9.1, the above equation is equivalent to (6.1.1) on the complete graph with vertices in U, K_U, satisfying conditions (1), (2), (3) and (4) of Assumption 4. Because U is finite, the graph K_U is a graph of bounded maximal degree. Thus, by Lemma 6.3 and Proposition 4.20, (B.0.2) is strongly well-posed. Since (B.0.2) only differs from (9.1.1) in the choice of driving Poisson processes, it follows by well-posedness of the two equations that $(\kappa_U^G, \overline{X}^G) \overset{\text{(d)}}{=} (\kappa_U^G, X_U^G) \sim \mu^G[U]$.

Lastly, let $\{\rho_j^{i,v}[U]\}_{v\in U, j\in\mathcal{J}}$, $i=1,2$, be two sets of jump rate functions satisfying the conditions outlined in Lemma 9.1 and let $\overline{X}^{G,i}$, $i=1,2$, be the respective pathwise unique solutions to (9.1.1) with the jump rate functions $\{\rho_j^{i,v}[U]\}$, $i=1,2$ and driven by the same Poisson processes $\mathbf{N}$. Fix any $(j,v)\in\mathcal{J}\times U$. By Lemma 9.1, $\rho_j^{1,v}[U](\cdot,\kappa_U^G,X_U^G)$ and $\rho_j^{2,v}[U](\cdot,\kappa_U^G,X_U^G)$ are both càglàd modifications of $t\mapsto \mathbb{E}\left[r_j^v(t,\kappa^G,X^G)|\kappa_U^G,X_U^G[t)\right]$ and are therefore indistinguishable. It immediately follows that for $\mu^G[U]$-a.s. $(\vartheta,x)\in(\mathcal{K}\times\mathcal{D})^U$, $\rho_j^{1,v}[U](t,\vartheta,x) = \rho_j^{2,v}[U](t,\vartheta,x)$ for every $t\in\mathbb{R}_+$. However, as we've shown above, $(\kappa_U,\overline{X}^{G,1}) \overset{\text{(d)}}{=} (\kappa_U,\overline{X}^{G,2}) \sim \mu^G[U]$, so

$$\begin{aligned}
\overline{X}_v^{G,1}(t) &= \varsigma(\kappa_v^G) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r\le\rho_j^{1,v}[U](s,\kappa_U,\overline{X}^{G,1})\right\}}\,\mathbf{N}_v(ds,dr,dj), \text{ for } v\in U, t\ge 0,\\
&= \varsigma(\kappa_v^G) + \int_{(0,t]\times\mathbb{R}_+\times\mathcal{J}} j\mathbb{I}_{\left\{r\le\rho_j^{2,v}[U](s,\kappa_U,\overline{X}^{G,1})\right\}}\,\mathbf{N}_v(ds,dr,dj), \text{ a.s. for } v\in U, t\ge 0.
\end{aligned}$$

Then by pathwise uniqueness, $\overline{X}^{G,1}$ and $\overline{X}^{G,2}$ must be indistinguishable. $\square$

APPENDIX C

Properties of Markov Random Fields and Conditional Independence

C.1 Some Supplementary Properties of Conditional Independence

The goal of this section is to prove Lemma 8.3 from Section 8.1.1. The proof relies on several technical properties of conditional independence established in Lemmas C.2 and C.3. We start with a basic measure-theoretic result, which establishes the equivalence of conditional expectations with respect to any σ-algebra $\mathbb{G}$ and its completion $\overline{\mathbb{G}}$.

Lemma C.1. *[37] Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra with $\mathbb{P}$-completion $\overline{\mathcal{G}}$. Then, for any bounded, $\mathcal{F}$-measurable random variable Z,*

$$\mathbb{E}\left[Z|\mathcal{G}\right] = \mathbb{E}\left[Z|\overline{\mathcal{G}}\right] \ \ a.s..$$

Proof. By definition, $\mathbb{E}[Z|\overline{\mathcal{G}}] : \Omega \to \mathbb{R}$ is a bounded, $\overline{\mathcal{G}}$-measurable function. Then by [53, Lemma 1.27], there exists a $\mathbb{G}$-measurable function $g : \Omega \to \mathbb{R}$ such that $g(\omega) = \mathbb{E}[Z|\overline{\mathcal{G}}](\omega)$ a.s.. Then for any $A \in \mathcal{G}$, $A \in \overline{\mathcal{G}}$ so that,

$$\mathbb{E}\left[\mathbb{I}_{\{A\}} g\right] = \mathbb{E}\left[\mathbb{I}_{\{A\}} \mathbb{E}\left[Z|\overline{\mathcal{G}}\right]\right] = \mathbb{E}\left[\mathbb{I}_{\{A\}} Z\right],$$

which implies that $\mathbb{E}[Z|\mathcal{G}] = g = \mathbb{E}[Z|\overline{\mathcal{G}}]$ a.s.. $\square$

Lemma C.1 immediately implies that conditional independence relations of σ-algebras are insensitive to completions.

Lemma C.2. *[37] Let $\mathcal{Z}_i$, $i = 1, 2, 3$, be Polish spaces and suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space. For each $i = 1, 2, 3$, let Z_i be a $\mathcal{F}$-measurable, $\mathcal{Z}_i$-valued random element. Then $Z_1 \perp\!\!\!\perp Z_2|Z_3$ if and only if $\mathcal{H}^{Z_1} \perp\!\!\!\perp \mathcal{H}^{Z_2}|\mathcal{H}^{Z_3}$.*

Proof. Suppose that $Z_1 \perp\!\!\!\perp Z_2|Z_3$ and let $\overline{A} \in \mathcal{H}^{Z_1}$. By [53, Lemma 1.27], there must exist a $\sigma(Z_1)$-measurable function $f : \Omega \to \mathbb{R}$ such that $f = \mathbb{I}_{\{\overline{A}\}}$ a.s.. Then the set $A := f^{-1}(\{1\}) \in \sigma(Z_1)$ and the symmetric difference of A and $\overline{A}$ is null. By repeated application of Lemma C.1,

$$\mathbb{P}(\overline{A}|\mathcal{H}^{Z_{\{2,3\}}}) = \mathbb{P}(A|\mathcal{H}^{Z_{\{2,3\}}}) = \mathbb{P}(A|Z_{\{2,3\}}) = \mathbb{P}(A|Z_3) = \mathbb{P}(A|\mathcal{H}^{Z_3}) = \mathbb{P}(\overline{A}|\mathcal{H}^{Z_3}).$$

Thus, $\mathcal{H}^{Z_1} \perp\!\!\!\perp \mathcal{H}^{Z_2}|\mathcal{H}^{Z_3}$. Now suppose $\mathcal{H}^{Z_1} \perp\!\!\!\perp \mathcal{H}^{Z_2}|\mathcal{H}^{Z_3}$. Then for any $A \in \sigma(Z_1) \subseteq \mathcal{H}^{Z_1}$, Lemma C.1 implies,

$$\mathbb{P}(A|Z_{\{2,3\}}) = \mathbb{P}(A|\mathcal{H}^{Z_{\{2,3\}}}) = \mathbb{P}(A|\mathcal{H}^{Z_3}) = \mathbb{P}(A|Z_3).$$

Thus, $Z_1 \perp\!\!\!\perp Z_2|Z_3$, which completes the proof. $\square$

Before proving Lemma 8.3, we state a number of facts about conditional independence proved in [81].

Lemma C.3. *[37] Let $\mathcal{G}_i$, $i \in \{1, \ldots, 6\}$, be six complete σ-algebras defined on some common measure space. Suppose also that $\mathcal{G}_1 \perp\!\!\!\perp \mathcal{G}_2|\mathcal{G}_3$. Then,*

(a) $\mathcal{G}_2 \perp\!\!\!\perp \mathcal{G}_1|\mathcal{G}_3$ [81, Proposition 2.4(a),(b)];

(b) if $\mathcal{G}_4 \perp\!\!\!\perp (\mathcal{G}_5, \mathcal{G}_6)$, then $\mathcal{G}_4 \perp\!\!\!\perp \mathcal{G}_5|\mathcal{G}_6$ [81, Proposition 2.5(b)];

(c) if $\mathcal{G}_4 \perp\!\!\!\perp \vee_{i=1}^{3}\mathcal{G}_i$, then $\mathcal{G}_1 \vee \mathcal{G}_4 \perp\!\!\!\perp \mathcal{G}_2|\mathcal{G}_3$ [81, Proposition 3.2(d)];

(d) if $\mathcal{G}_2 = \mathcal{G}_4 \vee \mathcal{G}_5$, then $\mathcal{G}_1 \perp\!\!\!\perp \mathcal{G}_4|\mathcal{G}_3 \vee \mathcal{G}_5$ [81, Proposition 3.2(a)];

(e) if $\mathcal{G}_4 \perp\!\!\!\perp \mathcal{G}_2|\mathcal{G}_3 \vee \mathcal{G}_1$, then $\mathcal{G}_1 \vee \mathcal{G}_4 \perp\!\!\!\perp \mathcal{G}_2|\mathcal{G}_3$, also [81, Proposition 3.2(a)];

(f) if $\mathcal{G}_4 \subseteq \mathcal{G}_1$, then $\mathcal{G}_4 \perp\!\!\!\perp \mathcal{G}_2|\mathcal{G}_3$ [81, Theorem 3.1];

(g) if $\mathcal{G}_1 \perp\!\!\!\perp \mathcal{G}_4|\mathcal{G}_2 \vee \mathcal{G}_3$, then $\mathcal{G}_1 \perp\!\!\!\perp \mathcal{G}_2|\mathcal{G}_3 \vee \mathcal{G}_4$ [81, Proposition 3.2(b)];

(h) if $\mathcal{G}_4 \subseteq \mathcal{G}_3$ and for every non-negative, $\mathcal{G}_1$-measurable random variable Z, $\mathbb{E}[Z|\mathcal{G}_3]$ is $\mathcal{G}_4$-measurable, then $\mathcal{G}_1 \perp\!\!\!\perp \mathcal{G}_2|\mathcal{G}_4$ [81, Theorem 3.3].

We may now prove Lemma 8.3.

Proof of Lemma 8.3. For convenience, we will freely apply the symmetry of conditional independence outlined in Lemma C.3(a) without reference. By Lemma C.2, condition 1 and (8.1.5) imply that,

1'. $\mathcal{H}^{Z_i^2}$, $i = 1, 2, 3$, are mutually independent and independent of $\vee_{i=1}^3 \mathcal{H}^{Z_i^1}$;

2'. $\mathcal{H}^{Z_1^1} \perp\!\!\!\perp \mathcal{H}^{Z_2^1}|\mathcal{H}^{Z_3^1}$.

By condition 1', $\mathcal{H}^{Z_3^2} \perp\!\!\!\perp \vee_{i=1}^3 \mathcal{H}^{Z_i^1}$. Then applying Lemma C.3(b), $\mathcal{H}^{Z_1^1} \perp\!\!\!\perp \mathcal{H}^{Z_3^2}| \vee_{i=2}^3 \mathcal{H}^{Z_i^1}$. Combining this with condition 2', Lemma C.3(g) implies that,

$$\mathcal{H}^{Z_1^1} \perp\!\!\!\perp \mathcal{H}^{Z_2^1}| \vee_{i=1}^2 \mathcal{H}^{Z_3^i}. \tag{C.1.1}$$

By condition 1', $\mathcal{H}^{Z_3^2} \perp\!\!\!\perp \mathcal{H}^{Z_1^1} \vee \mathcal{H}^{Z_3^1}$, so by Lemma C.3(b), $\mathcal{H}^{Z_1^1} \perp\!\!\!\perp \mathcal{H}^{Z_3^2}|\mathcal{H}^{Z_3^1}$. Thus, for any non-negative, $\mathcal{H}^{Z_1^1}$-measurable random variable Y, $\mathbb{E}\left[Y|\mathcal{H}^{Z_3^{\{1,2\}}}\right] = \mathbb{E}\left[Y|\mathcal{H}^{Z_3^1}\right]$ is $\mathcal{H}^{Z_3^3}$-measurable by condition 2 of the lemma. Together with (C.1.1), Lemma C.3(h) implies that

$$\mathcal{H}^{Z_1^1} \perp\!\!\!\perp \mathcal{H}^{Z_2^1}|\mathcal{H}^{Z_3^3}. \tag{C.1.2}$$

Next, condition 1' above implies that $\mathcal{H}^{Z_1^2} \perp\!\!\!\perp \mathcal{H}^{Z_3^2} \vee (\vee_{i=1}^3 \mathcal{H}^{Z_i^1}) \supseteq \mathcal{H}^{Z_1^1} \vee \mathcal{H}^{Z_2^1} \vee \mathcal{H}^{Z_3^3}$. Combined with (C.1.2) and Lemma C.3(c), this implies that $\mathcal{H}^{Z_1^1} \vee \mathcal{H}^{Z_1^2} \perp\!\!\!\perp \mathcal{H}^{Z_2^1}|\mathcal{H}^{Z_3^3}$. By condition 1' and the fact that $\mathcal{H}^{Z_3^3} \subseteq \vee_{i=1}^2 \mathcal{H}^{Z_3^i}$, $\mathcal{H}^{Z_2^2} \perp\!\!\!\perp (\vee_{i=1}^2 \mathcal{H}^{Z_i^1}) \vee \mathcal{H}^{Z_1^2} \vee \mathcal{H}^{Z_3^3}$. Applying Lemma C.3(c) again yields,

$$\mathcal{H}^{Z_1^1} \vee \mathcal{H}^{Z_1^2} \perp\!\!\!\perp \mathcal{H}^{Z_2^1} \vee \mathcal{H}^{Z_2^2}|\mathcal{H}^{Z_3^3}. \tag{C.1.3}$$

Finally, Apply Lemma C.3(f) twice given (C.1.3). First apply the result with $\mathcal{G}_1 = \mathcal{H}^{Z_1^1} \vee \mathcal{H}^{Z_1^2}$, $\mathcal{G}_4 = \mathcal{H}^{Z_1^3}$ noting that $\mathcal{G}_4 \subseteq \mathcal{G}_1$ by condition 2 of the lemma to get $\mathcal{H}^{Z_1^3} \perp\!\!\!\perp \mathcal{H}^{Z_2^1} \vee \mathcal{H}^{Z_2^2}|\mathcal{H}^{Z_3^3}$. Then apply the result with $\mathcal{G}_1 = \mathcal{H}^{Z_2^1} \vee \mathcal{H}^{Z_2^2}$, $\mathcal{G}_4 = \mathcal{H}^{Z_2^3}$ noting once again that $\mathcal{G}_4 \subseteq \mathcal{G}_1$ by condition 2 of the lemma. This implies that,

$$\mathcal{H}^{Z_1^3} \perp\!\!\!\perp \mathcal{H}^{Z_2^3}|\mathcal{H}^{Z_3^3}. \tag{C.1.4}$$

Then the result follows by Lemma C.2. $\square$

C.2 Supplementary Properties of SGMRFs

In this appendix we establish an equivalent characterization of α-SGMRFs (Lemma 6.8), then prove that α-SGMRFs on trees generalize Markov chains on trees as explained in Remark 6.9.

Proof of Lemma 6.8. Throughout the proof, fix the probability measure $\eta \in \mathcal{P}(\mathcal{Z}^V)$ and define the random vector $Z \sim \eta$. First, assume that $Z_A \perp\!\!\!\perp Z_B | Z_S$ for all finite, disjoint $A, B, S \subseteq V$ such that S α-separates A and B in G. Then for any $A', B', S' \subset V$ partitioning V so that $S' := \mathcal{N}^{\alpha}_{A'}(G)$ is finite and any finite $A \subseteq A'$ and $B \subseteq B'$, all paths from A to B must contain α consecutive elements of $\mathcal{N}^{\alpha}_{A'}(G)$, and so S' α-separates A and B in G. Thus, $Z_A \perp\!\!\!\perp Z_B | Z_{S'}$, and because A, B are arbitrary finite subsets of A' and B' respectively, it follows that $Z_{A'} \perp\!\!\!\perp Z_{B'} | Z_{S'}$, so η is also an α-SGMRF.

In the other direction, now assume that η is an α-SGMRF and let $A, B, S \subseteq V$ be finite, disjoint vertex sets such that S α-separates A and B. Then we claim there exist A', B', S partitioning V such that

1. $A \subseteq A'$, $B \subseteq B'$ and $S \supseteq \mathcal{N}^{\alpha}_{A'}(G)$;

2. $Z_{A'} \perp\!\!\!\perp Z_{(S \cup B') \setminus \mathcal{N}^{\alpha}_{A'}(G)} | Z_{\mathcal{N}^{\alpha}_{A'}(G)}$.

If the claim holds, by Lemma C.3(d) and applying Lemma C.2 twice, $Z_{A'} \perp\!\!\!\perp Z_{B'} | Z_S$. By condition 1, this directly implies that $Z_A \perp\!\!\!\perp Z_B | Z_C$ as desired, so it suffices to prove the claim.

We now prove the claim. Let A' be the set of elements in $V \setminus S$ that are not α-separated from A in G. Let $B' = V \setminus (S \cup A')$. Then $B \subseteq B'$ because S α-separates A and B. Moreover, by construction, S α-separates A' and B'. Thus, $\mathcal{N}^{\alpha}_{A'}(G) \subseteq S$ which completes the verification of condition 1. Then $A', (S \cup B) \setminus \mathcal{N}^{\alpha}_{A'}(G)$ and $\mathcal{N}^{\alpha}_{A'}(G)$ partition V, and $\mathcal{N}^{\alpha}_{A'}(G) \subseteq S$ is finite by assumption. Thus, condition 2 holds because η is an α-SGMRF. That concludes the proof of the claim and therefore the lemma. $\square$

Proof of Remark 6.9. By uniqueness of paths in trees, note that for the finite, disjoint sets $A, B, S \subseteq V$, S α-separates A and B in G if and only if it α-separates A and B in $G[U]$ for every finite, connected $U \subseteq V$ such that $A \cup B \cup S \subseteq U$. Then by Lemma 6.8, this implies that η is an α-SGMRF with respect to G if and only if $\eta[U]$ is an η-MRF with respect to $G[U]$ for all finite, connected $U \subseteq V$. □

C.2.1 Counterexamples: IPS that do not satisfy stronger MRF Properties

Example C.4. [37] *There exists an IPS on a finite graph $G = (V, E)$ with initial data $\kappa = (\kappa_v)_{v \in V}$ that are mutually independent, but whose trajectories do not form an MRF.*

Let $G = (V, E)$ be the path on three vertices with $V = \{1, 2, 3\}$ and $E = \{\{1, 2\}, \{2, 3\}\}$. Let $\mathcal{K} = \mathcal{X} = \{0, 1\}$, $\mathcal{J} = \{1\}$ and let ς be the identity mapping on $\mathcal{K}$. Set $X_2(0) = \kappa_2 := 0$ and let $(X_1(0), X_3(0)) = (\kappa_1, \kappa_3)$ be i.i.d. Bernoulli(1/2) random variables. Thus, $\{\kappa_v\}_{v=1,2,3}$ are mutually independent.

Now consider the following (Markovian) jump rate functions, $r_1^v : \mathbb{R}_+ \times \mathcal{D}^V \to \mathbb{R}_+$:

$$r_1^1 \equiv r_1^3 \equiv 0 \text{ and } r_1^2(t, x) = \mathbb{I}_{\{x_1(0) \neq x_3(0), x_2(t-)=0\}}.$$

This model satisfies Assumption 4 and the maximum degree of any vertex in G is 2. Also $X_1(0) \neq X_3(0)$ with probability 1/2 and $X_2 \equiv 0$ on the event $\{X_1(0) = X_3(0)\}$. Note that the function $f : \mathcal{D}_{t-} \to \mathbb{R}$ defined by $f(y) = y(0)$ is bounded and measurable. Then it easily follows that on the event $\{X_2(t-) = 1\}$,

$$\begin{aligned} \mathbb{E}\left[f(X_1[t])|X_2[t]\right] &= \mathbb{P}\left(X_1(0) = 1|X_2[t]\right) &&= \tfrac{1}{2}, \\ \mathbb{E}\left[f(X_1[t])|X_{\{2,3\}}[t]\right] &= \mathbb{P}\left(X_1(0) = 1|X_{\{2,3\}}[t]\right) &&= 1 - X_3(0). \end{aligned}$$

Thus, $\mathrm{Law}(X_1[t]|X_{\{2,3\}}[t]) \neq \mathrm{Law}(X_1[t]|X_2[t])$, so the trajectories $X[t]$ are not MRFs for any $t > 0$.

Next we note that there exists a process $X(t) = \{X_v(t)\}_{v \in V}$ that may fail to be a local α-MRF for any $\alpha \in \mathbb{N}$ and some $t > 0$.

Example C.5. [37] *There exists an IPS with SGMRF initial data and $t \in \mathbb{R}_+$ such that Law$(X(t))$ fails to be even a local α-MRF for any $\alpha \in \mathbb{N}$.*

There is a substantial literature on the topic of dynamic transitions of IPS from Gibbs states to non-Gibbs states which we borrow from to construct our counterexample. We consider an example introduced in [89] and further analyzed in [90]. Let G be the infinite d-regular tree for some $d > 2$ and let μ_κ be the positive-boundary Ising model for an inverse temperature-magnetic field pair (β, h) at which the Ising model experiences a phase transition. The exact structure of μ_κ is described in [41, Section 12.2]. In particular, it is a Gibbs measure with a Markovian specification, and therefore it is a local MRF. Furthermore, as an extremal countable state MRF, it is also a Markov chain on the tree [92, Corollary 2]. By Remark 6.9, this implies that it is also an SGMRF and therefore an α-SGMRF for every $\alpha \in \mathbb{N}$. Now let μ be the law of the IPS X that evolves according to infinite temperature Glauber dynamics described in [89] with initial condition $X(0)$ such that $\text{Law}(X(0)) = \mu_\kappa$. Then by [90, Theorem 3.9 and Remark 3.10], there exists an interval $[t_1, t_2) \subseteq \mathbb{R}_+$ such that for all $t \in [t_1, t_2)$, $\text{Law}(X(t))$ is non-Gibbsian. Because $\text{Law}(X(t))$ is non-Gibbs it also fails to be a local α-MRF for any $\alpha \in \mathbb{N}$.

APPENDIX D

Other Minor Results

D.1 Well-Posedness for Finite Initial Data

Under Assumption 1, well-posedness of (3.1.4) is common knowledge when the initial data is finite, but we establish it here for completeness. In this case, we also show that the trajectories also satisfy the following additional regularity property. Recall the definition of the discontinuity set $\mathrm{Disc}_t(x)$ given in (1.4.1). Recall the definition of a.s. proper trajectories from the statement of Lemma 6.3.

Proof of Proposition 4.1 and trajectories being a.s. proper: Let $(G,\xi) = (V,E,\varnothing,\overline{\kappa},\kappa,\xi)$ any a.s. finite $[\overline{\mathcal{K}},\mathcal{K}\times\mathcal{X}]$-random graph. By Lemma 3.10, it suffices to prove that (3.1.4) is strongly well-posed for (G,ξ) with a.s. proper solutions under the additional assumption that (G,ξ) is deterministic.

Let $(\mathbb{F},\mathbf{N}^G)$ be a filtration-Poisson process pair in the sense of Remark 3.6. First note that for any $(\mathbb{F},\mathbf{N}^G)$-weak solution (G,X) to (3.1.4), any distinct vertices $u,v\in V$ and any $T>0$,

$$\mathrm{Disc}_T(X_v)\cap\mathrm{Disc}_T(X_u) \subseteq \{s\in[0,T] : \mathbf{N}^G_u(\{s\}\times(0,C_{k,T}]\times\mathcal{J})\mathbf{N}^G_v(\{s\}\times(0,C_{k,T}]\times\mathcal{J}) = 1\} = \emptyset \text{ a.s.},$$

by (3.1.4) and Assumption 1 where $\{C_{k,T}\}$ is the family of constants from Assumption 1 and $k := \max\{|\mathrm{cl}_u|,|\mathrm{cl}_v|\}$. Since this holds for all T, (G,X) is a.s. proper.

Next, fix a filtration-Poisson process pair $(\mathbb{F},\mathbf{N}^G)$. Define the finite collection,

$$\mathcal{E} := \{(\tau_n,r_n,j_n,v_n)\in[0,T]\times[0,C_{K,T}]\times\mathcal{J}\times V : \mathbf{N}^G_{v_n}(\{(\tau_n,r_n,j_n)\}) = 1\},$$

$K = \max_{v\in V} |\mathrm{cl}_v|$ and $\{\tau_n\}_{n=1}^{|\mathcal{E}|}$ is increasing. Note that $\{\tau_n\}$ is, in fact, strictly increasing. define the $\mathcal{D}_T^V$-random element $X[T]$ by,

$$X_v(t) = \begin{cases} \xi_v & \text{if } 0 \le t < \tau_1, \\ X_v(\tau_n) & \text{if } \tau_n \le t < \tau_{n+1}, n < |\mathcal{E}|, \\ X_v(\tau_n) & \text{if } \tau_n \le t \le T, n = |\mathcal{E}|, \\ X_v(\tau_n) & \text{if } t = \tau_{n+1}, r_{n+1} > r^{G,v}_{j_{n+1}}(t_{n+1}, X), \\ X_v(\tau_n) + j_{n+1} & \text{if } t = \tau_{n+1}, r_{n+1} > r^{G,v}_{j_{n+1}}(t_{n+1}, X). \end{cases}$$

Then note that for $t \in [0,T]$, $X(t)$ is clearly $\mathcal{H}_t^{G,\xi,\mathbf{N}^G}$-measurable so X is a $\mathbf{N}^G$-strong solution to (3.1.4) for (G,ξ) on the interval $[0,T]$. Furthermore, it is clear that any $(\mathbb{F},\mathbf{N}^G)$-weak solution to (3.1.4) must satisfy the above display, so all $(\mathbb{F},\mathbf{N}^G)$-weak solutions equal X on the interval $[0,T]$. Because T is arbitrary, and for any $T' > T$ the corresponding solution X' satisfies $X'[T] = X[T]$ a.s., it follows that there exists a $\mathbf{N}^G$-strong solution to (3.1.4) and that solution is pathwise unique. Therefore (3.1.4) is strongly well-posed for (G,ξ). □

D.2 Characterization of Strong Well-Posedness on Random Graphs

Proof of Lemma 3.10. The key issue here is to show that conditioning on the initial data does not change the driving noise structure. Note that this is slightly non-standard as the driving noise is indexed by the vertices of the graph and thus is not completely independent of the initial data. Let $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ be a complete, filtered probability space that supports (G,ξ) and a filtration-Poisson process pair $(\mathbb{F},\mathbf{N}^G)$ such that (G,X) and (G,Y) are two $(\mathbb{F},\mathbf{N}^G)$-weak solutions of (3.1.4) for (G,ξ). To prove the lemma, it suffices to prove the following claim: for $\mathbb{P}$-a.s. $\omega \in \Omega$, setting $(H,\xi^H) := (G(\omega),\xi(\omega))$, there exists a filtration-Poisson process pair $(\mathcal{F}^H,\mathbf{N}^H)$ on some probability space $(\Omega^H,\mathcal{F}^H,\mathbb{F}^H,\mathbb{P}^H)$ and two $(\mathcal{F}^H,\mathbf{N}^H)$-weak solutions (H,X^H) and (H,Y^H) to (3.1.4) for (H,ξ^H) such that $\mathrm{Law}((G,X,Y,\mathbf{N}^G)|\mathcal{F}_0)(\omega) = \mathrm{Law}((H,X^H,Y^H,\mathbf{N}^H))$. Indeed, if the claim holds, then a.s. strong well-posedness of (3.1.4) for every realization of the random graph (G,ξ) implies that $\mathbb{P}(X = Y|\mathcal{F}_0)(\omega) = \mathbb{P}^H(X^H = Y^H) = 1$ for $\mathbb{P}$-a.s. $\omega \in \Omega$. Hence $\mathbb{P}(X = Y) = 1$, which proves strong well-posedness. This claim can be proved via direct verification of Definitions 3.3 and 3.5. We include the details for completion.

To prove the claim, let $(H, X^H, Y^H, \mathbf{N}^H)$ be a random element with law $\mathrm{Law}((G, X, Y, \mathbf{N}^G)|\mathcal{F}_0)(\omega)$ and fix a complete, filtered probability space $(\Omega^H, \mathcal{F}^H, \mathbb{F}^H, \mathbb{P}^H)$ that supports $(H, X^H, Y^H, \mathbf{N}^H)$ where $\mathbb{F}^H$ is the minimal filtration satisfying the usual conditions such that X^H, Y^H and $\mathbf{N}^H$ are all $\mathbb{F}^H$-adapted in the sense that for any $v \in V_H$, X^H_v, Y^H_v and $\mathbf{N}^H_v$ are all $\mathbb{F}^H$-adapted (point) processes. Since by assumption $\mathbf{N}^G$ is a $\mathbb{F}$-driving noise, condition 1 of Definition 3.3 immediately implies $\mathbf{N}^H$ is a collection of i.i.d. Poisson processes on $\mathbb{R}^2_+ \times \mathcal{J}$ with intensity $\mathrm{Leb}^2 \otimes \mathcal{J}$, indexed by the vertices of H, and hence that $(H, \mathbf{N}^H)$ is an $\mathbb{H}^{\mathbf{N}^H}$-driving noise.

Since (3.1.4) holds a.s., for $\mathbb{P}$-a.s. $\omega \in \Omega$, (H, X^H) and (H, Y^H) are both graphs with random càdlàg marks that, together with $(H, \mathbf{N}^H)$, $\mathbb{P}(\cdot|\mathcal{F}_0)(\omega)$-a.s. solve (3.1.4) for (H, ξ^H). Thus, (H, X^H) and (H, Y^H) satisfy conditions 1 and 3 of Definition 3.5 $\mathbb{P}$-a.s. with respect to the filtered probability space $(\Omega^H, \mathcal{F}^H, \mathbb{F}^H, \mathbb{P}^H)$ (noting that any $\mathbb{P}(\cdot|\mathcal{F}_0)(\omega)$-null event is also $\mathbb{P}^H$-null), and so it suffices to prove that $\mathbf{N}^H$ is $\mathbb{P}$-a.s. a collection of i.i.d. $\mathbb{F}^H$-Poisson processes. To do this, we note that condition 2 of Definition 3.3 implies that for any $t > 0$ and $A \in \mathcal{B}((t, \infty) \times \mathbb{R}_+ \times \mathcal{J})$, the random element $(G, \mathbf{N}^G(A))$ is conditionally independent of the $\mathcal{F}_t$-measurable random element $(G, X[t], Y[t], \mathbf{N}^G_t)$ given $\mathcal{F}_0$, where $\mathbf{N}^G_t = \mathbf{N}^G|_{[0,t] \times \mathbb{R}_+ \times \mathcal{J}}$. Thus, $\mathbb{P}$-a.s., $(H, \mathbf{N}^H(A))$ is independent of $(H, X^H[t], Y^H[t], \mathbf{N}^H_t)$. Then using a standard approximation argument exploiting the fact that Borel sigma algebras of subsets of Polish spaces are countably generated and that $\mathbb{F}^H$ is complete, it follows that $\mathbb{P}$-a.s., $\mathbf{N}^H_v(A)$ is independent of $\mathcal{F}^H_t$ for all $v \in V_H$, $t \in \mathbb{R}_+$ and $A \in \mathcal{B}((t, \infty) \times \mathbb{R}_+ \times \mathcal{J})$. Thus, $\{\mathbf{N}^H_v\}_{v \in V_H}$ is $\mathbb{P}$-a.s. a collection of $\mathbb{F}^H$-Poisson processes. Therefore, $\mathbf{N}^H$ is $\mathbb{P}$-a.s. an $\mathbb{F}^H$-driving noise in the sense of Definition 3.3, so condition 2 of Definition 3.5 is also $\mathbb{P}$-a.s. satisfied. This concludes the proof of the claim. □

D.3 A Classical Change-of-Measure Result

For convenience we rephrase here the result of [19, Theorem 15.2.7] in the specific case that the mark space is finite. Recall that the definition of local characteristics of a marked point process was given in Definition 8.14.

Theorem D.1. *[19, Theorem 15.2.7] Fix a probability space $(\Omega, \mathbb{G}, \eta^1)$ and let N_Z be a simple and locally finite point process on $\mathbb{R}_+$ with marks in $\mathcal{Z}$. Suppose that N_Z is $\mathbb{G}$-adapted and admits the $(\eta^1, \mathbb{G})$-local characteristics (λ_g, Ψ). Let $\{\theta(t)\}_{t\geq 0}$ be a non-negative $\mathbb{G}$-predictable process and let $\{h(t,z)\}_{t\geq 0, z\in\mathcal{Z}}$ be a non-negative, $\mathbb{G}$-mark predictable random function. Suppose that for all $t \geq 0$,*

$$\int_0^t \lambda_g(s)\theta(s)\, ds < \infty \ \text{a.s.}, \tag{D.3.1}$$

and

$$\sum_{z\in\mathcal{Z}} h(t,z)\Psi(t,z) = 1 \ \text{a.s.}. \tag{D.3.2}$$

If the events of N_Z are given by the sequence $\{(t_n, z_n)\}$, then define

$$L(t) := \left(\prod_{t_n\in(0,t]} \theta(t_n)h(t_n,z_n)\right) \exp\left(-\sum_{z\in\mathcal{Z}} \int_{(0,t]} \left(\theta(s)h(s,z) - 1\right) \lambda_g(s)\Psi(s,z)\, ds\right). \tag{D.3.3}$$

Then the following properties hold:

(a) $\{L(t)\}_{t\in\mathbb{R}_+}$ is a non-negative $\mathbb{G}$-local martingale. Moreover, if $\mathbb{E}\left[L(t)\right] = 1$ for all $t \geq 0$, then it is a $\mathbb{G}$-martingale.

(b) If $\mathbb{E}\left[L(T)\right] = 1$ for some $T > 0$ and if the measure η^2 is defined by

$$\frac{d\eta_T^2}{d\eta_T^1} = L(T),$$

then under η^2, N_Z admits the $(\eta^2, \mathbb{G})$-local characteristic $(\theta\lambda_g, h\Psi)$ on $[0,T]$.

D.4 ODE Results

Proof of Lemma 10.1. Let $\overline{\Delta}_n := \{b \in \mathbb{R}_+^{n-1} : \sum_{i=1}^{n-1} b_i \leq 1\}$ be the projection of the unit simplex in $\mathbb{R}^n$ to $\mathbb{R}^{n-1}$ obtained by removing the nth component of any vector in Δ_n. Let $\overline{\Delta}_n^\circ := \{b \in (0,1)^{n-1} : \sum_{i=1}^{n-1} b_i < 1\}$ be the interior of $\overline{\Delta}_n$, and let $\partial\overline{\Delta}_n := \overline{\Delta}_n \setminus \overline{\Delta}_n^\circ$ be its boundary. For any $b \in \Delta_n$, let $\overline{b} = (b_1, \ldots, b_{n-1}) \in \overline{\Delta}_n$ so that $b = (\overline{b}, 1 - \sum_{i=1}^{n-1} \overline{b}_i) = (\overline{b}, b_n)$. Define the function

$\overline{f} : \mathbb{R}_+ \times \overline{\Delta}_n^\circ \to \mathbb{R}_+^{n-1}$ by,

$$\overline{f}(\overline{b}) = \left(f_{i,n}(\overline{b}, b_n)b_n + \sum_{k \notin \{i,n\}} f_{i,k}(\overline{b}, b_n)\overline{b}_k - g_i(\overline{b}, b_n)\overline{b}_i \right)_{i=1}^{n-1}, \quad b_n = 1 - \sum_{i=1}^{n-1} \overline{b}_i.$$

It immediately follows that $\overline{f}$ is bounded, continuously differentiable and any continuously differentiable function x will satisfy (10.2.1) if and only if $\overline{x}$ satisfies the initial value problem

$$\frac{d\overline{x}}{dt} = \overline{f}(\overline{x}),$$
$$\overline{x}(0) = \overline{a},$$

where (10.2.2) ensures that $\frac{d}{dt}\left(\sum_{i=1}^n x_i\right) \equiv 0$ so that $x_n(t) = 1 - \sum_{i=1}^n \overline{x}_i(t)$ for all $t > 0$. Because $\overline{f}$ is continuously differentiable in Δ_n°, it is also locally Lipschitz. Therefore, by [6, Theorem 1.43] (which slightly extends the Picard-Lindelöf theorem), there exists an interval $0 \in (s_1, s_2) \subseteq \mathbb{R}$ and a continuously differentiable function $\overline{\varphi} : (s_1, s_2) \to \overline{\Delta}_n^\circ$ such that for any interval $0 \in I_{\overline{x}} \subseteq \mathbb{R}$ and any continuously differentiable solution $\overline{x} : I_{\overline{x}} \to \overline{\Delta}_n^\circ$ to (10.2.1), $I_{\overline{x}} \subseteq (s_1, s_2)$ and $\overline{x}(t) = \overline{\varphi}(t)$ for all $t \in I_{\overline{x}}$. As we are only interested in solutions to (10.2.1) that are well defined at non-negative times, it suffices to prove that $s_2 = \infty$, which we prove by contradiction.

Suppose that $s_2 < \infty$. Then by [6, Theorem 1.46], for each compact $K \subset [0, s_2] \times \overline{\Delta}_n^\circ$ there must exist an $\epsilon > 0$ such that $(t, \overline{\varphi}(t)) \notin K$ for each $t > s_2 - \epsilon$. In particular, let $K' \subseteq \overline{\Delta}_n^\circ$ be any compact set and let $K = [0, s_2] \times K'$. Then [6, Theorem 1.46] implies that there exists an $\epsilon > 0$ such that $\overline{\varphi}(t) \notin K'$ whenever $t > s_2 - \epsilon$. For any $\delta > 0$, let $K'_\delta = \{\overline{b} \in \overline{\Delta}_n^\circ : \min\{b_i\}_{i=1}^n \geq \delta\}$ (where recall that $b_i = \overline{b}_i$ for $i < n$ and $b_n = 1 - \sum_{i=1}^{n-1} \overline{b}_i$) and note that K'_δ is always compact. We claim that $\lim_{\delta \searrow 0} \inf\{t > 0 : \overline{\varphi}(t) \notin K'_\delta\} = \infty$. If the claim holds, then for sufficiently small δ, it follows that $\overline{\varphi}(t) \in K'_\delta$ for all $t \in [0, s_2]$. This is a contradiction, so $s_2 = \infty$.

To prove the claim, fix any $M \in \mathbb{R}_+$ such that $\max_{\{i,k\}} \max\{\|f_{i,k}\|, \|g_i\|\} \leq M$ (M exists because these functions are bounded and $n < \infty$), where $\|\cdot\|$ denotes the uniform norm. Let $\varphi : (s_1, s_2) \to \Delta_n^\circ$ be the function defined by $\varphi_i = \overline{\varphi}_i$ for $i \leq n-1$ and $\varphi_n = 1 - \sum_{i=1}^{n-1} \overline{\varphi}_i$. Then note that φ is a continuously differentiable solution to (10.2.1) and for each $t \in [0, s_2]$, $\overline{\varphi}(t) \in K'_\delta$ if and only if $\min\{\varphi_i(t)\} \geq \delta$. By boundedness of g_i, non-negativity of φ and (10.2.1), for each $i \in \{1, \ldots, n\}$ and $t \geq 0$, $\varphi_i(t) \geq y_i(t)$, where y_i solves the initial value problem,

$$\frac{dy_i}{dt} = -My_i,$$

$$y_i(0) = a_i.$$

Then $y_i(t) = a_i e^{-Mt}$, so $\lim_{\delta \searrow 0} \inf\{t > 0 : \overline{\varphi}(t) \notin K'_\delta\} = \lim_{\delta \searrow 0} \inf\{t > 0 : \min_i \varphi_i(t) \geq \delta\} \geq \lim_{\delta \searrow 0} \inf\{t > 0 : \min_i a_i e^{-Mt} < \delta\} = \infty$, which proves the claim. $\square$

Proof of Lemma 10.3. Let $\oplus$ denote the symmetric difference operator and define $C < \infty$ such that $\sup_{v,j} \|\widehat{r}_j^v\| \leq C$, where $\|\cdot\|$ is the uniform norm. Then note that,

$$\begin{aligned}
&\left\{X_U^G(t+\Delta t) = S_{v,j}(X_U^G(t))\right\} \oplus \left\{\mathbf{N}_v\left((t, t+\Delta t] \times (0, \widehat{r}_j^v(t+, X^G(t))] \times \{j\}\right) = 1\right\} \\
&\qquad \subseteq \left\{\mathbf{N}_v\left((t, t+\Delta t] \times (0, \widehat{r}_j^v(t+, X^G(t))] \times \{j\}\right) > 1\right\}.
\end{aligned}$$

By the memoryless property,

$$\begin{aligned}
&\mathbb{P}(\mathbf{N}_v\left((t, t+\Delta t] \times (0, \widehat{r}_j^v(t+, X^G(t))] \times \{j\}\right) > 1 | X_W^G(t-)) \\
&\qquad = \mathbb{P}(\mathbf{N}_v\left((t, t+\Delta t] \times (0, \widehat{r}_j^v(t+, X^G(t))] \times \{j\}\right) > 1) \\
&\qquad = O\left((\Delta t)^2\right).
\end{aligned}$$

It follows that

$$\begin{aligned}
&\mathbb{P}\left(X_U^G(t+\Delta t) = S_{v,j}(X_U(t)) \middle| X_W^G(t)\right) \\
&\qquad = \mathbb{P}\left(\mathbf{N}_v\left((t, t+\Delta t] \times (0, \widehat{r}_j^v(t+, X^G(t))] \times \{j\}\right) = 1 \middle| X_W^G(t)\right) + O\left((\Delta t)^2\right) \\
&\qquad = \widehat{r}_j^v(t+, X_{\mathrm{cl}_v}^G(t))\Delta t + O\left((\Delta t)^2\right).
\end{aligned}$$

$\square$

Bibliography

[1] M. Abért and Y. Glasner. Generic groups acting on regular trees. *Transactions of the American Mathematical Society*, 361(7):3597–3610, 2009.

[2] D. Aldous and R. Lyons. Processes on unimodular random networks. *Electronic Journal of Probability*, 12:1454–1508, 2007.

[3] David Aldous and J. Michael Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In Harry Kesten, editor, *Probability on Discrete Structures*, pages 1–72. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[4] G. Amir, R. Baldasso, and N. Beilin. Majority dynamics and the median process: connections, convergence and some new conjectures. pages 1–27, 2019.

[5] J. Aubin and H. Frankowska. *Set-Valued Analysis.* Modern Birkhäuser Classics. Springer, Boston, 1 edition, 2009.

[6] L. Barreira and C. Valls. *Ordinary Differential Equations: Qualitative Theory*, volume 137 of *Graduate studies in mathematics.* American Mathematical Society, 2012.

[7] E. Bayraktar, S. Chakraborty, and R. Wu. Graphon mean field systems. pages 1–28, 2021.

[8] E. Bayraktar and R. Wu. Mean field interaction on random graphs with dynamically changing multi-color edges. *Stochastic Processes and their Applications*, 141:197–244, 2021.

[9] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electronic Journal of Probability*, 6:1–13, 2001.

[10] S. Bhamidi, A. Budhiraja, and R. Wu. Weakly interacting particle systems on inhomogeneous random graphs. 129(6):1–31, 2019.

[11] S. Bhamidi, D. Nam, O. Nguyen, and A. Sly. Survival and extinction of epidemics on random graphs with general degree. *The Annals of Probability*, 49(1):1–39, 2021.

[12] P. Biane and R. Durrett. Ten Lectures on Particle Systems. In P. Bernard, editor, *Lectures on Probability Theory*, Ecole d'Eté de Probabilités de Saint-Flour XXIII, pages 97–201. Springer, 1993.

[13] P. Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Statistics. John Wiley & Sons, inc, 2 edition, 1999.

[14] T. Björk. *Arbitrage Theory in Continuous Time*. Oxford scholarship online. Oxford University Press, Oxford, 4 edition, 2020.

[15] P. Block et al. Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nature Human Behaviour*, 4(6):588–596, 2020.

[16] V. Bogachev. *Measure Theory*, volume 2. Springer, Berlin Heidelberg, 2007.

[17] C. Bordenave. Lecture notes on random graphs and probabilistic combinatorial optimization. `https://www.math.univ-toulouse.fr/~bordenave/coursRG.pdf`, 2016. available at: `https://www.math.univ-toulouse.fr/~bordenave/coursRG.pdf`, accessed: 2022–01-24.

[18] P. Brémaud. *Point processes and queues: martingale dynamics*. Springer-Verlag, New York, 1981.

[19] P. Brémaud. *Probability theory and stochastic processes*. Universitext. Springer, 2020.

[20] P. Brémaud and L. Massoulié. Stability of nonlinear Hawkes processes. *The Annals of Probability*, 24(3):1563–1588, 1996.

[21] A. Budhiraja, P. Dupuis, M. Fischer, and K. Ramanan. Limits of relative entropies associated with weakly interacting particle systems. *Electronic Journal of Probability*, 20, 2015.

[22] A. Budhiraja, D. Mukherjee, and R. Wu. Supermarket model on graphs. *Annals of Applied Probability*, 29(3):1740–1777, 2019.

[23] A. Carro, R. Toral, and M. San Miguel. The noisy voter model on complex networks. *Scientific Reports*, 6(April):1–14, 2016.

[24] L. Chaintron and A. Diez. Propagation of chaos: a review of models, methods and applications. I. Models and methods. 2022.

[25] L. Chaintron and A. Diez. Propagation of chaos: a review of models, methods and applications. II. Applications. 2022.

[26] K. Chung. *A course in probability theory, revised edition.* Academic Press, 2 edition, 2001.

[27] F. Coppini, H. Dietert, and G. Giacomin. A law of large numbers and large deviations for interacting diffusions on Erdos-Rényi graphs. *Stochastic Dynamics*, 2019.

[28] D. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure*, volume 2 of *Probability and its Applications*. Springer, 2 edition, 2008.

[29] S. Delattre, G. Giacomin, and E. Luçon. A note on dynamical models on random graphs and Fokker–Planck equations. *J. Stat. Phys.*, 165(4):785–798, 2016.

[30] Sylvain Delattre, Nicolas Fournier, and Marc Hoffmann. Hawkes processes on large networks. *Annals of Applied Probability*, 26(1):216–261, 2016.

[31] A. Dembo, A. Montanari, and N. Sun. Factor models on locally tree-like graphs. *The Annals of Probability*, 41(6):4162–4213, 2013.

[32] R. Durrett. *Probability: theory and examples.* 2019.

[33] S. Ethier and T. Kurtz. *Markov Processes Characterization and Convergence*. John Wiley and Sons, inc, 1986.

[34] L. Fontes, T. Mountford, D. Ungaretti, and M. Vares. Renewal contact processes: phase transition and survival. pages 1–29, 2021.

[35] A. Ganguly and A. McDonough. Rotor-routing induces the only consistent sandpile torsor structure on plane graphs. 2022.

[36] A. Ganguly and K. Ramanan. Hydrodynamic limits of non-Markovian interacting particle systems on sparse graphs. 2022.

[37] A. Ganguly and K. Ramanan. Interacting jump processes preserve semi-global markov random fields on path space. 2022. In preparation.

[38] N. Gantert and D. Schmid. The speed of the tagged particle in the exclusion process on Galton–Watson trees. *Electronic Journal of Probability*, 25:1–27, 2020.

[39] N. Gast. The power of two choices on graphs: the pair-approximation is accurate? *ACM SIGMETRICS Performance Evaluation Review*, 43(2):69–71, 2015.

[40] N. Gast. *Refinements of mean field approximation.* habilitation, 2019.

[41] Hans-Otto Georgii. *Gibbs measures and phase transitions.* De Gruyter Studies in Mathematics. Berlin/New York, 2 edition, 2011.

[42] J. Gleeson et al. Accuracy of mean-field theory for dynamics on real-world networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 85(2):16–18, 2012.

[43] S. Goldstein. Remarks on the global Markov property. *Communications in Mathematical Physics*, 74(3):223–234, 1980.

[44] B. Granovsky and N. Madras. The noisy voter model. *Stochastic Processes and their Applications*, 55:23–43, 1995.

[45] G. Grimmett and A. Stacey. Critical probabilities for site and bond percolation models. *The Annals of Probability*, 26(4):1788–1812, 1998.

[46] T. E. Harris. Nearest-neighbor Markov interaction processes on multidimensional lattices. *Advances in Mathematics*, 9:66–89, 1972.

[47] T. E. Harris. Contact interactions on a lattice. *The Annals of Probability*, 2(6):969–988, 1974.

[48] X. Huang and R. Durrett. The contact process on random graphs and Galton-Watson trees. *Alea (Rio de Janeiro)*, 17(1):159–182, 2020.

[49] R. Huo and R. Durrett. The zealot voter model. *Annals of Applied Probability*, 29(5):3128–3154, 2019.

[50] J. Jacod. *Calcul Stochastique et Problèmes de Martingales.* Lecture Notes in Mathematics. Springer-Verlag Berlin Heidelberg, 1979.

[51] S. Janson, M. Luczak, and P. Windridge. Law of large numbers for the SIR epidemic on a random graph with given degrees. *Random Structures and Algorithms*, 45(4):726–763, 2014.

[52] A. Jędrzejewski. Pair approximation for the q -voter model with independence on complex networks. *Physical Review E*, 95(1), 2017.

[53] O. Kallenberg. *Foundations of Modern Probability.* Probability Theory and Stochastic Modelling. Springer Nature Switzerland AG, 3 edition, 2021.

[54] W. Kermack and A. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society A-Mathematical, Physical and Engineering Sciences*, 115(772), 1927.

[55] C. Kipnis and C. Landim. *Scaling Limits of Interacting Particle Systems*, volume 320 of *Grundlehren der mathematischen Wissenschaften.* 1999.

[56] R J Williams K.L. Chung. *Introduction to stochastic integration.* Probability and its applications. Birkhäuser, 2 edition, 2014.

[57] V. Kolokoltsov. *Nonlinear Markov Processes and Kinetic Equations.* Cambridge Tracts in Mathematics. Cambridge University Press, 2010.

[58] T Kurtz. The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities. *Electronic Journal of Probability*, 12:951–965, 2007.

[59] D. Lacker, K. Ramanan, and R. Wu. Locally interacting diffusions as markov random fields on path space. *Stochastic Processes and their Applications*, 140:81–114, 2021.

[60] D. Lacker, K. Ramanan, and R. Wu. Marginal dynamics of interacting diffusions on unimodular Galton-Watson trees. pages 1–59, 2021.

[61] D. Lacker, K. Ramanan, and R. Wu. Local weak convergence for sparse networks of interacting processes. *The Annals of Applied Probability*, 2022. to appear.

[62] D. Lacker and J. Zhang. Stationary solutions and local equations for interacting diffusions on regular trees. pages 1–36, 2021.

[63] S. Lauritzen. *Graphical models*. Oxford University Press, New York, 1996.

[64] T. Liggett. Existence theorems for infinite particle systems. *Transactions of the American Mathematical Society*, 165(March):471, 1972.

[65] T. Liggett. *Interacting Particle Systems*. Springer-Verlag, New York, 1 edition, 1985.

[66] T. Liggett. Multiple transition points for the contact process on the binary tree. *The Annals of Probability*, 24(4):1675–1710, 1996.

[67] T Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*, volume 324 of *Runcllehren cler mathematischen Wissenschaften*. Springer, 1999.

[68] A. Mata, R. Ferreira, and S. Ferreira. Heterogeneous pair-approximation for the contact process on complex networks. *New Journal of Physics*, 16, 2014.

[69] A. Montanari, E. Mossel, and A. Sly. The weak limit of Ising models on locally tree-like graphs. *Probability Theory and Related Fields*, 152(1-2):31–51, 2012.

[70] E. Mossel and A. Sly. Exact thresholds for Ising-Gibbs samplers on general graphs. *Annals of Probability*, 41(1):294–328, 2013.

[71] D. Nam, O. Nguyen, and A. Sly. Critical value asymptotics for the contact process on random graphs. *Transactions of the American Mathematical Society*, 2022.

[72] C. Neuhauser. The role of explicit space in plant competition models. In *Perplexing problems in probability: Festschrift in honor of {H}arry {K}esten*, chapter 19, pages 355–371. 1999.

[73] K. Oelschläger. A martingale approach to the law of large numbers for weakly interacting stochastic processes. *The Annals of Probability*, 12(2):458–479, 1984.

[74] R. Oliveira and G. Reis. Interacting diffusions on random graphs with diverging average degrees: hydrodynamics and large deviations. *Journal of Statistical Physics*, 176(5):1057–1087, 2019.

[75] R. Oliveira, G. Reis, and L. Stolerman. Interacting diffusions on sparse graphs: hydrodynamics from local weak limits. *Electronic Journal of Probability*, 25(110), 2020.

[76] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200–3203, 2001.

[77] R. Pemantle. The contact process on trees. *The Annals of Probability*, 20(4):2089–2116, 1992.

[78] R. Pemantle and A. Stacey. The branching random walk and contact process on Galton-Watson and nonhomogeneous trees. *The Annals of Probability*, 29(4):1563–1590, 2001.

[79] P. Pfaffelhuber, S. Rotter, and J. Stiefel. Mean-field limits for non-linear Hawkes processes with excitation and inhibition. 2021.

[80] A. Puha. A reversible nearest particle system on the homogeneous tree. *Journal of Theoretical Probability*, 12(1):217–254, 1999.

[81] C. Putten and J. Schuppen. Invariance properties of the conditional independence relation. *The Annals of Probability*, 13(3):934–945, 1985.

[82] M. Rao. On modification theorems. *Transactions of the American Mathematical Society*, 167:443–450, May 1972.

[83] S. Redner. Reality-inspired voter models : a mini-review. *Comptes Rendus Physique*, 20(4):275–292, 2019.

[84] G. Röst, Z. Vizi, and I. Z. Kiss. Pairwise approximation for SIR-type network epidemics with non-Markovian recovery. *Proc. R. Soc. A*, 474, 2018.

[85] A.S. Sznitman. *Topics in propagation of chaos.* Springer-Verlag, 1991.

[86] T Tao. *An introduction to measure theory.* AMS, 2011.

[87] W. Truccolo. From point process observations to collective neural dynamics: nonlinear Hawkes process GLMs, low-dimensional dynamics and coarse graining. *Journal of Physiology Paris*, 110(4):336–347, 2016.

[88] R. van der Hofstad. Random Graphs and Complex Networks, 2022. available at: `https://www.win.tue.nl/~rhofstad/NotesRGCNII_colleagues_25_04_2022.pdf`, accessed: 2022–07-14.

[89] A. Van Enter et al. Possible loss and recovery of Gibbsianness during the stochastic evolution of Gibbs measures. *Communications in Mathematical Physics*, 226(1):101–130, 2002.

[90] A. Van Enter et al. Gibbs-non-Gibbs properties for evolving Ising models on trees. *Annales de l'institut Henri Poincare (B) Probability and Statistics*, 48(3):774–791, 2012.

[91] F. Vazquez and V. Eguíluz. Analytical solution of the voter model on uncorrelated networks. *New Journal of Physics*, 10, 2008.

[92] S. Zachary. Countable state space Markov random fields and Markov chains on trees. *The Annals of Probability*, 11(4):894–903, 1983.