STOCHASTIC PDES ON GRAPHS AS SCALING LIMITS OF DISCRETE INTERACTING SYSTEMS

WAI-TONG LOUIS FAN[*][†]

Stochastic partial differential equations (SPDE) on graphs were recently introduced by Cerrai and Freidlin [9]. This class of stochastic equations in infinite dimensions provides a minimal framework for the study of the effective dynamics of much more complex systems. However, how they emerge from microscopic individual-based models is still poorly understood, partly due to complications near vertex singularities. In this work, motivated by the study of the dynamics and the genealogies of expanding populations in spatially structured environments, we obtain a new class of SPDE on graphs of Wright-Fisher type which have nontrivial boundary conditions on the vertex set. We show that these SPDE arise as scaling limits of suitably defined biased voter models (BVM), which extends the scaling limits of Durrett and Fan [29]. We further obtain a convergent simulation scheme for each of these SPDE in terms of a system of Itô SDEs, which is useful when the size of the BVM is too large for stochastic simulations. These give the first rigorous connection between SPDE on graphs and more discrete models, specifically, interacting particle systems and interacting SDEs. Uniform heat kernel estimates for symmetric random walks approximating diffusions on graphs are the keys to our proofs. Some open problems are provided as further motivations of our study.

1. Introduction. Stochastic partial differential equations (SPDE) on graphs, more precisely SPDE whose spatial variables lie in a metric graph, first explicitly appear in Cerrai and Freidlin [9, 10] as asymptotic limits of SPDE on two-dimensional domains that shrink to a graph. Here a graph Γ is a continuous object consisting of all points on its edges, so the real line $\mathbb{R}$ is a trivial example which has one edge and no vertex. These equations provide a "*minimal*" framework for the study of the interplay between the solution of SPDE and the geometric properties of the underlying metric space, "*minimal*" in the sense that the metric space is essentially one-dimensional yet flexible enough to incorporate nontrivial topologies and various boundary conditions on the vertex set. Such interplay between the evolution of the quantity (e.g. density of a population or concentration of a chemical) modeled by the equation and the spatial environment of the system is of fundamental importance in scientific modeling and control. For example, an important problem in ecology is to identify mechanisms that permit the coexistence of species in different geographical environments. The role of space and stochasticity in shaping competition outcomes and biodiversity has been intensively studied in spatial evolutionary games. It has also been explored rigorously in the framework of interacting particle systems (a.k.a. stochastic cellular automata), as in Durrett [28], Lanchier and Neuhauser [66, 67] and Lanchier [65], to name just a few. See the

---

seminal articles of Durrett and Levin [30, 31] about the importance of space in modeling.

A practical motivation for our study of SPDE on graphs is to provide a theoretical foundation for previous experimental work [34, 57] and on-going modeling work on co-infection spread of defective and normal viruses. Here co-infection means simultaneous infection by two or more different types of virus particles. Instead of the traditional petri dish, designed micro-arrays of network structures are used as the container of host cells [34, 57]. Virus infections and co-infections are then systematically initiated and observed. The aim is to *predict the propagation speed and the spatial patterns for viral co-infections in spatially structured populations* of biological host cells. Insights obtained from these laboratory studies are potentially applicable to more complex real life epidemic networks, which is important in controlling epidemic spread [80].

Quantitative imaging and analysis of viral infection provides extensive spatial-temporal data for validation and refinement of models. However, a reliable mathematical framework is still missing. Deterministic models like PDE on graphs fail to capture the dynamics of viral particles or genomes, because fluctuations of propagating infection fronts are typically observed. A more reasonable macroscopic model is instead an SPDE on the graph formed by the host cell environment. The question, then, is to *deduce the "correct" SPDE on the graph*, based on the local and spatial interactions between the viral particles and the host cells. For example, where does the observed noise come from; more specifically, what is the magnitude of the noise term in the SPDE in terms of microscopic rules? Can local interactions near a vertex singularity lead to the emergence of new terms in the SPDE? In co-infection spread, what population-level signatures reveal emergence of new levels of cooperation and conflict between the defective and normal viruses? In an on-going work, the author is developing various stochastic spatial models, including individual-based models [48] such as interacting particle systems and systems of stochastic reaction-diffusion equations, to model the joint evolution of defective and normal viruses. This paper provides the theoretical foundation for [34, 57], which is still missing in the mathematics literature.

Another broader motivation for this paper is related to our long term goal to understand the *genealogies in expanding populations and the resulting patterns of genetic heterogeneity.* This is important because medical treatments in cancer or epidemics may fail due to drug resistance, if one does not have an accurate knowledge of the mutational types present. The genetic forces at work in a growing cancer tumor or in an infection spread are very similar to those in a population expanding into a new geographical area, in which most of the advantageous mutation occur near the front. See Lee and Yin [69] and Edmonds et al. [33]. Existing studies for genealogies in expanding populations mostly rely on computer simulations [60] and nonrigorous arguments [54, 55, 56, 61, 70, 78]. The first rigorous analysis for this is perhaps in Durrett and Fan [29], which provides a precise description of the lineage dynamics in terms of a coupled SPDE of Fisher-Kolmogorov-Petrovsky-Piscounov (FKPP) type. However, the spatial domain is restricted to $\mathbb{R}$. A rigorous analysis for the genealogies in $\mathbb{R}^d$ for dimensions $d \geq 2$ seems difficult. Even more, a solution theory for the stochastic FKPP is not yet available in two or higher dimensions. Therefore, besides the "minimality" mentioned in the first paragraph, SPDE on metric graphs provide a natural setting for further analysis. See Section 8.3 for some concrete open problems.

Main question and significance. Even though SPDE on graphs and their deter-

ministic counterpart arise naturally in scientific problems and discoveries, they are rather unexplored. In the mathematics literature, a subset of these equations seem to first (and so far only) explicitly have appeared in the rather recent work [9, 10]. This is partly because SPDE is still considered to be a rather new and technical modeling approach compared with deterministic models, but a more important reason is that it is not yet clear **how these equations emerge from interactions in the microscopic scale**, especially interactions near vertex singularities. This fundamental question, our focus in this work as suggested by our title, needs to be carefully investigated in order to answer more specific questions such as those raised in the previous paragraph about epidemic spread. With increasingly advanced technology, more and more experimental data describe both cell-level and population-level behavior. Thus, connecting continuum models with discrete models not only can facilitate model validation at both scales, but also provides complementary perspectives of the complex dynamics (such as tumor growth and virus spread) under study. Increasing recognition of these benefits has stimulated efforts to connect discrete and continuum models in a variety of biological and ecological contexts; see [30] for different modeling perspectives and [31, 14, 28] for some reviews.

Known results. In standard settings such as domains in $\mathbb{R}^d$, there are rigorous approximation schemes of SPDE by microscopic particle models, such as Sturm [87] and Kotelenez and Kurtz [62], where [87] is for a stochastic heat equation with a multiplicative colored noise term and [62] is for SPDE of McKean-Vlasov type. Particle representations of SPDE are obtained by Kurtz and coauthors in [64, 35, 17]. Cox, Durrett and Perkins [16] showed that the rescaled long range voter model in dimension $d \geq 2$ converges to a super-Brownian motion. Durrett, Mytnik and Perkins [32] showed convergence of multi-type contact processes to a pair of super-Brownian motions interacting through their collision local times. For genealogies of super-processes, see the snake process of Le Gall [68], the historical process of Evans and Perkins [38] and the lookdown process of Donnelly and Kurtz [27]. Our two practical motivations led us to first focus on the stochastic FKPP, the base case model for an expanding population density exhibiting noisy wavefront, which is of the form

$$(1) \qquad \partial_t u = \alpha\,\Delta u + \beta\,u(1-u) + \sqrt{\gamma\,u(1-u)}\,\dot{W},$$

where $\dot{W}$ is the space-time white noise, $\alpha > 0$ is the diffusion coefficient representing the average dispersal distance of the individuals, $\beta \geq 0$ accounts for an average increase and the last term represents fluctuations during reproduction where $\gamma \geq 0$ parametrizes the variance.

Muller and Tribe [77] gave the first rigorous convergence result that stochastic FKPP on $\mathbb{R}$ can arise as scaling limit of long range biased voter models (BVM). In [29], we generalized this by scrutinizing all possible scalings for which this type of connections between SPDE (1) and BVM are valid. These BVM are idealized individual-based models for an expanding population on the 1-dimen rescaled integer lattice $L_n^{-1}\mathbb{Z}$ whose points are called *demes*. In the n-th model in [29], there is one cell at each point of the lattice

$$(L_n^{-1}\mathbb{Z}) \times \{1, \ldots, M_n\},$$

whose cell-type is either 1 (cancer cell) or 0 (healthy cell). Each cell in deme $w \in L_n^{-1}\mathbb{Z}$ only interact with the $2M_n$ neighbors in demes $w - L_n^{-1}$ and $w + L_n^{-1}$. Type-0 cells reproduce at rate $2M_n r_n$, type-1 cells at a higher rate $2M_n(r_n + \beta_n)$ due to higher fitness. When reproduction

occurs the offspring replaces a neighbor chosen uniformly at random. In the terminology of evolutionary games, this is birth-death updating. It is shown in [29, Theorem 1] that under the scalings

$$(2) \qquad \frac{r_n M_n}{L_n^2} \to \alpha, \ M_n \beta_n \to \frac{\beta}{2}, \ \frac{L_n}{M_n} \to \frac{\gamma}{4\alpha}, \ L_n \to \infty \ \text{as} \ n \to \infty,$$

the local fraction of type 1 converges to the solution of equation (1) on $\mathbb{R}$. The scalings (2) were explained in [29, Section 1.2] and will be compared with those in this paper in Remark 3.

Connections between models of different scales, offered by these types of scaling limit theorems and also Theorems 1 and 2 in this paper, not only provide complementary insights into the underlying mechanisms of the complex dynamical system, but are also of fundamental importance for model selection and analysis. For instance, the above convergence tells us that the variance of the noise is of order L_n/M_n near the wavefront where u is bounded away from 0 and 1, which is important in predicting the propagation speed [74]. See [6] for some behaviors that are expected to hold in general for the class of the stochastic pulled fronts with weak noise.

More related work. For SDE in infinite dimensions, which is an abstract framework containing SPDE on graphs or on manifolds [89, 90], see [22] for general background and [24] a comparison between different theories. Despite recent breakthroughs including [52, 53, 49, 7], a solution theory for the stochastic FKPP in the dimensions $d \geq 2$ is still open.

One might replace the white noise by a colored noise as in [23, 87] to smooth out spatial correlations. This standard approach indeed gives well-defined SPDE in higher dimensions, but the particle approximation would be less intuitive. There is also a large body of literature about SPDE arising as the fluctuation limits of interacting particle systems, which we do not attempt to give a survey. We refer the reader to [46] for the stochastic burgers equation and [11] for reaction-diffusion equations, and the references therein.

Mathematical contributions. Our main contributions are as follows.

1. Our scaling limit theorems give **the first rigorous connections between SPDE on graphs and discrete models**. See Remark 3. Theorem 1 generalizes [29, Theorem 1] to the graph setting, paying special attention to the new spatial heterogeneity and vertex singularities. A new stochastic FKPP on graph emerges from a suitably defined BVM (Section 3). In Theorem 2, we give a convergent simulation scheme for this SPDE in terms of a system of interacting Itô SDEs. This schemeis useful when the size of the BVM is too large for stochastic simulations.
2. Besides having a different type of limit theorems compared to [9, 10], we consider more general diffusions on graphs Γ. In fact, we identify suitable conditions on Γ for the study of SPDE on graphs and point out directions for generalizations. See Remark 2. Our SPDE have an extra boundary condition at the set of vertices, and the coefficients are typically non-Lipschitz. Well-posedness of these SPDE is established via a new duality in Lemma 1.
3. An application of results like the conjunction of Theorems 1 and 2 is as follows: Given a complex system (such as cancer cell dynamics) with fine details, one (i) starts with an individual-based model which elucidate, at a more fundamental level, single-particle

interactions, spatial component and stochasticity of the system, then (ii) deduces the macroscopic evolution of the particles that emerges, as in Remark 3, which might be an SPDE/PDE describing the evolution of the particle density, and (iii) simulates the SPDE/PDE which is robust against the size of the microscopic system, using the interacting SDEs or numerical methods. We summarize this lesson for our case as:

$$\text{BVM} \;\to\; \text{SPDE} \;\to\; \text{interacting SDEs}$$

Benefits actually go both ways: an intuitive way to understand SPDE is through scaling limits of discrete approximating systems, similar to the way one interprets Brownian motion as scaling limits of random walks.

4. Besides vertex singularities, a technical challenge in the proofs is to obtain *uniform estimates of the transition kernel* of random walks on a discretized version Γ^n of Γ. For this we need to impose an assumption on Γ. The volume-doubling property and the Poincaré inequality in Assumption 1 are enough for this paper, and we point to further generalizations in Remark 1. Uniform estimates for the random walks and the diffusions obtained in Theorems 3-5 and also the local CLT are of crucial importance in analyzing regularity properties of SPDE on Γ in general.

5. The scalings discovered in Theorem 1 enable one to generalize, to the graph setting, scaling limit results for coupled SPDE such as [29, Theorem 4]. This is a key step towards the study of interacting populations of more than one species. Broadly speaking, this paper points to directions for various generalizations, such as defining SPDE on random graphs and on fractals, studying SPDE defined through Walsh diffusions instead of symmetric diffusions, extending scaling limits of contact processes in [77] to the graph settings, etc. See Sections 8.2 and 8.3 for more generalizations and open questions.

The paper is organized as follows: We give preliminaries in Sections 2 and 3, including assumptions on the graph Γ, diffusions and SPDE and the construction of BVM, before stating our main results in Section 4. Uniform heat kernel estimates for random walks on the discretized graph Γ^n and the local CLT are obtained in Section 5. Sections 6 and 7 are proofs for Theorem 1 and Theorem 2. Some extensions of our method and some open problems are offered in Section 8.2 and Section 8.3 respectively. Finally, the notions of weak solutions and mild solutions to SPDE on graphs are written down in the Appendix for completeness.

2. Diffusions and SPDE on metric graphs. A **metric graph** (Γ, d) is a topological graph $\Gamma = (V, E, \partial)$ endowed with a metric d, where V and E are disjoint countable sets and ∂ is a map. The elements of V and E are respectively the vertices and the directed edges of Γ. The combinatorial structure of the graph is described by the map $\partial : E \to V \times (V \cup \{\infty\})$ which sends every edge e to an ordered pair (e_-, e_+). For simplicity, we identify e with its image $\partial(e) = (e_-, e_+)$. We call e_- and e_+ the *initial vertex* and the *terminal vertex* of e (self-loops, i.e. $e_+ = e_-$, are allowed).

The *degree* of $v \in V$ is defined as $deg(v) := |E(v)| = |E^+(v)| + |E^-(v)|$, where $E^\pm(v) := \{e \in E : e_\pm = v\}$ consists of all edges starting $(-)$ and ending $(+)$ at v respectively, and $E(v) := E^+(v) \bigcup E^-(v)$ is their disjoint union.

There are two types of edges. If $(e_-, e_+) \in V \times V$, then e is called an *internal edge*; if $(e_-, e_+) \in V \times \{\infty\}$, then e is called an *external edge*. Each internal edge is isomorphic to

a closed and bounded interval in $\mathbb{R}$, and each external edge is isomorphic to $[0, \infty)$. We identify Γ with the union of all edges. The metric d on Γ is defined in the canonical way as the length of a shortest path between two points along the edges. Internal edges have finite lengths and external edges have infinite lengths. We equip Γ with the 1-dimensional Hausdorff measure m associated with d and write this metric measure space as (Γ, d, m).

Denote by $\mathring{e}$ the interior of the edge e and $\mathring{\Gamma} = \Gamma \setminus V = \cup_{e \in E} \mathring{e}$ be the interior of Γ. For a function $f : \Gamma \to \mathbb{R}$, we define

- $\nabla f_\pm(x)$ to be the one sided derivative of f at $x \in \mathring{e}$, along e *towards* $e_\pm$,
- $\nabla f_{e_\pm}(v)$ be the one sided derivative of f at $v = e_\mp \in V$, along e *towards* $e_\pm$

whenever they exist. A function f is said to be differentiable at $x \in \mathring{e}$ if $\nabla f_+(x) = -\nabla f_-(x)$, in which case this quantity $\nabla f(x) = \lim_{\mathring{e} \ni y \to x} \frac{f(y) - f(x)}{d(x,y)}$ is called the derivative of f at $x \in \mathring{e}$. Higher derivatives $\nabla^k f$ (for $k \geq 2$) are defined similarly.

Recall that an edge $e \in E$ is isomorphic to a closed interval in $\mathbb{R}$. The space of k-times continuously differentiable on e is

$$\mathcal{C}^k(e) := \left\{ g \in \mathcal{C}^k(\mathring{e}) : \nabla^r g \text{ is uniformly continuous on bounded subsets of } \mathring{e} \text{ for all } r \leq k \right\}.$$

If $g \in \mathcal{C}^k(e)$, then $\nabla^r g$ continuously extends to the closed set e for all $r \leq k$. We also let

- $\mathcal{C}(\Gamma)$ be the space of continuous functions on Γ,
- $\mathcal{C}^k(\Gamma) := \left\{ f \in \mathcal{C}(\Gamma) : f|_e \in \mathcal{C}^k(e) \text{ for all } e \in E \right\}$ where $f|_e$ is the restriction of f on e,
- $\mathcal{C}^\infty(\Gamma) := \cap_{k \geq 1} \mathcal{C}^k(\Gamma)$,
- $\mathcal{C}_c(\Gamma)$ be the space of continuous functions with compact support on Γ,
- $\mathcal{C}_c^k(\Gamma) := \mathcal{C}^k(\Gamma) \cap \mathcal{C}_c(\Gamma)$,
- $\mathcal{C}_c^\infty(\Gamma) := \mathcal{C}^\infty(\Gamma) \cap \mathcal{C}_c(\Gamma)$.

For an arbitrary sigma-finite measure μ on Γ (e.g., the Hausdorff measure m or the measure $\ell(x) m(dx)$ in (5)), we define the L^2-norm of f

$$\|f\|_{L^2(\mu)} := \sqrt{\int_\Gamma |f(x)|^2 \, \mu(dx)} = \sqrt{\sum_e \int_e |f(x)|^2 \, \mu(dx)}$$

and the Sobolev norm $\|f\|_{W^{1,2}(\mu)} := \|f\|_{L^2(\mu)} + \|\nabla f\|_{L^2(\mu)}$. The space $L^2(\Gamma, \mu)$ is the set of μ-measurable functions f with $\|f\|_{L^2(\mu)} < \infty$ and the Sobolev space

$$W^{1,2}(\Gamma, \mu) := \left\{ f \in L^2(\Gamma, \mu) : \nabla f \in L^2(\Gamma, \mu) \right\}.$$

Our focus is particle approximation to a class of parabolic SPDE on Γ. Our notion of solution to such an SPDE, detailed in the Appendix, involves a diffusion on Γ. This motivates us to construct such diffusions next, under assumptions on Γ and the diffusion coefficients.

Diffusions on metric graphs via Dirichlet forms. Although a nontrivial diffusion can be defined on very general Γ including fractals [4] and random graphs [1, 2], we make the following assumption on Γ, which ensures certain regularity on the transition density of symmetric diffusions. See Remark 1.

ASSUMPTION 1. *The metric graph* (Γ, d, m) *has a positive infimum for the branch lengths and satisfies the followings.*

1. *(Volume-doubling) There is a constant $C_{VD} > 0$ such that*

(3) $$m(B(x, 2r)) \leq C_{VD}\, m(B(x, r))$$

for all $x \in \Gamma$ and $r > 0$ where $B(x, r) := \{y \in \Gamma : d(x, y) < r\}$ is a ball.

2. *(Poincaré inequality) There is a constant $C_{PI} > 0$ such that*

(4) $$\int_{B(x,r)} |f(y) - \bar{f}_B|^2\, m(dy) \leq C_{PI}\, r^2 \int_{B(x,2r)} |\nabla f(y)|^2\, m(dy)$$

for all $f \in W^{1,2}(\Gamma, m)$, $x \in \Gamma$, $r > 0$, where $\bar{f}_B := m^{-1}(B) \int_B f\, dm$ is the average value of f over $B = B(x, r)$.

REMARK 1. Any graph with finitely many edges satisfies Assumption 1 with $C_{VD} = 2\sum_{v \in V} deg(v)$ and $C_{PI} = C^2 \sum_{v \in V} deg(v)$, where C is the same absolute constant in Theorem 2 in [37, Section 5.8] with $p = 2$ and $n = 1$. Many infinite graphs also satisfy Assumption 1. These include any regular infinite lattice such as $\mathbb{Z}^n$ and any infinite regular tree with constant branch length. As we shall explain in Subsection 5.3, the conjunction of (3) and (4) is *equivalent to* the existence of two-sided Gaussian bounds for the transition density of symmetric diffusions on Γ. Assumption 1 can be significantly relaxed.

We shall construct diffusions on graphs by the Dirichlet form method, under the following conditions on the diffusion coefficients and the symmetrizing measure.

ASSUMPTION 2. *Suppose we are given two functions $\alpha, \ell \in W^{1,2}(\Gamma, m)$ that are strictly elliptic, i.e. bounded above and below by positive constants.*

We now define the measure ν on Γ by

(5) $$\nu(dx) = \ell(x)\, m(dx),$$

which has full support and is locally finite, and we consider the symmetric bilinear form

$$\mathcal{E}(f, g) := \int_\Gamma \alpha(x)\, \nabla f(x) \cdot \nabla g(x)\, \nu(dx)$$

with domain $Dom(\mathcal{E}) = W^{1,2}(\Gamma, \nu) \cap \mathcal{C}^1(\Gamma)$. It can be checked that $(\mathcal{E},\, Dom(\mathcal{E}))$ is a Dirichlet form in $L^2(\Gamma, \nu)$ that possesses the local property. See for instance [13]. Furthermore, $(\mathcal{E},\, Dom(\mathcal{E}))$ is *regular* under Assumptions 1 and 2. This can be checked (for instance [1, Proposition 4.1]) by showing that the space $\mathcal{C}_c^\infty(\Gamma)$ of infinitely differentiable functions on Γ with compact support is a core of the domain. Hence by [43, Theorem 7.2.1], there is a ν-symmetric diffusion $X = \{X_t\}_{t \geq 0}$ on the graph Γ associated with this Dirichlet form. Henceforth we refer to X as the $\mathcal{E}$-**diffusion** and denote by $(\Omega, \mathcal{F}, \{\mathbb{P}_x\}_{x \in \Gamma})$ the filtered probability space on which X is defined, where $\mathbb{P}_x$ is a probability measure on $(\Omega, \mathcal{F})$ such that $\mathbb{P}_x(X_0 = x) = 1$. Whenever $\mathbb{P}$ appears in an expression, it denotes the probability measure on the space on which the random variables involved in that expression are defined.

REMARK 2. In [42, 9, 10] such a diffusion process is constructed by specifying its generator under stronger assumptions: the set of vertices V is finite, $\alpha = 1$ is a constant function and ℓ is smooth in $\mathring{\Gamma}$. We follow the notation in [9, 10] for Γ, ℓ and ν. Studying solutions to

SPDE on graphs rely crucially on our understanding about certain diffusion processes on graphs, which is a research topic by itself. See [42, 41] for theoretical foundation of diffusions on finite graphs, [71, 79, 91] for some interesting applications.

As is known [43, 13, 72], the Dirichlet form method is more robust against irregularities of both the diffusion coefficients and the underlying metric space. The price to pay, however, is that many statements about the associated process X are valid apriori only for "quasi-everywhere", that is, except for a set of capacity zero. Fortunately, most of these statements can be strengthened to be valid for "all $x \in \Gamma$", provided that we have extra knowledge about its transition density. By the usual L^2 method (see, e.g., [8]), X_t admits a transition density $p(t, x, y)$ with respect to its symmetrizing measure $\nu(dx)$. That is

$$\tag{6} \mathbb{P}_x(X_t \in dy) = p(t, x, y)\nu(dy)$$

and $p(t, x, y) = p(t, y, x)$ for all $t \geq 0$, ν-almost all $x, y \in \Gamma$. The previous 'almost all' can further be strengthened to 'all' because (3) and (4) imply Hölder continuity for $p(t, x, y)$. See Theorem 5 for the precise statement.

Gluing condition and generator. For any edge $e \in E$, it is known (for example [43, Example 1.2.3]) that every function $f \in W^{1,2}(e, m)$ coincides a.e. with an absolutely continuous function on $\mathring{e}$ having derivative defined m-a.e. and lies in $L^2(e, m)$. In particular, under Assumption 2, the one-sided limits

$$\alpha_{e\pm} := \lim_{x \to e_\pm} \alpha(x) \quad \text{and} \quad \ell_{e\pm} := \lim_{x \to e_\pm} \ell(x) \quad \text{exist for all } e \in E.$$

As in [42, Section 3], the $L^2(\Gamma, \nu)$ infinitesimal generator $\mathcal{L}$ of X can be described as a second order differential operator

$$\tag{7} \mathcal{L}f(x) := \frac{1}{2\ell(x)} \nabla\big(\ell(x)\,\alpha(x)\,\nabla f(x)\big), \quad x \in \mathring{\Gamma},$$

endowed with the gluing conditions

$$\tag{8} 0 = \big(\nabla_{out} f \cdot [\ell\alpha]\big)(v), \quad v \in V,$$

where the following *outward* derivative of f at v is used:

$$\tag{9} \big(\nabla_{out} f \cdot [g]\big)(v) := \sum_{e \in E^+(v)} \nabla f_{e-}(v)\, g_{e+} + \sum_{e \in E^-(v)} \nabla f_{e+}(v)\, g_{e-}.$$

The gluing condition (8) reduces to the Neumann boundary condition at vertices with degree 1. From (8) and integration by parts, we can check that $\mathcal{E}(f, g) = \langle -\mathcal{L}f, g\rangle_{L^2(\nu)}$ for $f \in Dom_{L^2}(\mathcal{L})$ and $g \in Dom(\mathcal{E})$, where $Dom_{L^2}(\mathcal{L})$ is the domain of the generator $\mathcal{L}$. See [13, Appendix A.4] and [72, Chapter 1] for general relations among the generator $\mathcal{L}$, the Dirichlet form and the semigroup $\{P_t\}_{t \geq 0}$ of X.

SPDE on graphs. In this work we focus on the following stochastic FKPP type equation on Γ with nontrivial boundary condition:

$$\tag{10} \begin{cases} \partial_t u = \mathcal{L}\,u + \beta\,u(1 - u) + \sqrt{\gamma\,u(1 - u)}\dot{W} & \text{on } \mathring{\Gamma} \\ \nabla_{out} u \cdot [\alpha] = -\hat{\beta}\,u(1 - u) & \text{on } V, \end{cases}$$

where $\dot{W}$ is the space-time white noise on $[0, \infty) \times \Gamma$, the operator ∇_{out} is defined in (9), the graph satisfies Assumption 1, the functions α, ℓ satisfy Assumption 2, and the functions $\beta, \gamma, \hat{\beta}$ satisfy the following assumption.

ASSUMPTION 3. *Let $\beta, \gamma : \Gamma \to [0, \infty)$ be non-negative bounded measurable functions on Γ and $\hat{\beta} : V \to [0, \infty)$ be a non-negative bounded function on V.*

Here we adopt Walsh's theory (named after Walsh's St. Flour notes [92]) and regard (10) as a shorthand for an integral equation. See Definition 2 in Appendix for this integral equation and the definition of a weak solution to (10).

Note that when $\hat{\beta}(v)$ is positive, the negative sign in the boundary condition corresponds to creation of mass (growth of u) at $v \in V$. To see this, note that $\nabla_{out}u \cdot [\alpha](v)$ is a sum of derivatives of u *away* from the vertex v, and therefore towards the interior of the edges connecting to v. These *inward* normal derivatives along the edges correspond to a growth rather than a decay of u.

While Assumptions 1 and 2 guarantee existence and basic properties of the diffusion process X on the graph, Assumption 3 will be needed for the weak uniqueness of (10). The latter will be established via duality (Lemma 1).

To have a cleaner description without losing much generality of the spatial heterogeneity across edges, we further restrict to piecewise constant functions whenever a discrete approximation is involved (namely, in Theorem 1, Theorem 2, Lemma 3 and Theorem 3).

EXAMPLE 1. *Suppose $\alpha(x) = \alpha_e$, $\beta(x) = \beta_e$ and $\gamma(x) = \gamma_e$ whenever $x \in \mathring{e}$, where $\{\alpha_e, \beta_e, \gamma_e\}$ are non-negative numbers that are uniformly bounded and $\inf_{e \in E} \alpha_e > 0$. Suppose $\ell = 1$ constant. Then SPDE (10) reduces to*

$$(11) \qquad \begin{cases} \partial_t u = \alpha_e \, \Delta u + \beta_e \, u(1-u) + \sqrt{\gamma_e \, u(1-u)} \dot{W} & \text{on } \mathring{e} \\ \nabla_{out}u \cdot [\alpha] = -\hat{\beta} \, u(1-u) & \text{on } V. \end{cases}$$

In this case $\alpha_{e+} = \alpha_{e-} := \alpha_e$ and for $v \in V$,

$$(12) \qquad \left(\nabla_{out}u \cdot [\alpha]\right)(v) = \sum_{e \in E(v)} \left(\mathbf{1}_{e_-=v} \nabla u_{e+}(v) + \mathbf{1}_{e_+=v} \nabla u_{e-}(v) \right) \alpha_e.$$

3. Rescaled biased voter model. In this section, we describe a sequence of biased voter models (BVM) indexed by $n \in \mathbb{N}$, which is a natural generalization to the one in [29] described in the introduction.

For each edge $e \in E$, we associate it with two sequences $\{L_n^e\}_{n \geq 1} \subset (0, \infty)$ and $\{M_n^e\}_{n \geq 1} \subset \mathbb{N}$, then we choose a *maximal* countable subset e^n of its interior $\mathring{e}$ such that neighboring points in e^n are of distance $1/L_n^e$. Without loss of generality, by throwing away demes that are too close to the endpoints of an edge if necessary, we suppose

$$(13) \qquad 1/L_n^e \leq d(x, v) < 2/L_n^e$$

whenever $x \in e^n$ is neighboring to vertex v. Points in the **discretized graph** $\Gamma^n := \cup_e e^n$ are called *demes*. Each deme $x \in e^n$ represents an isolated location containing a subpopulation
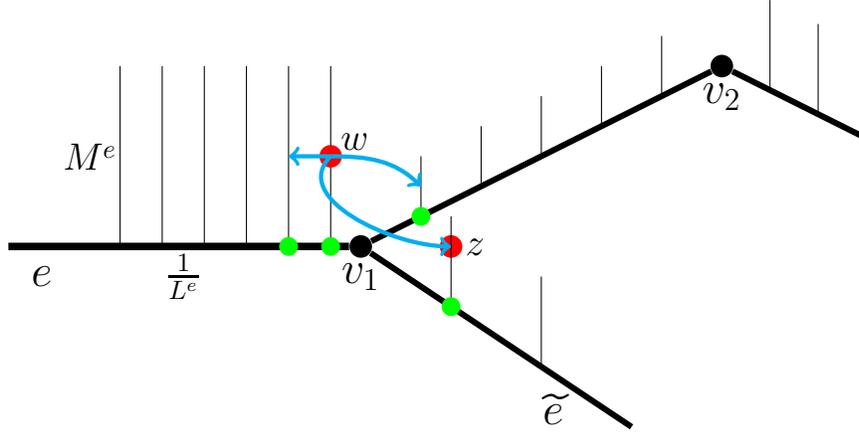
FIG 1. *Lattice* $\Lambda_n := \cup_e \Lambda_n^e$, *together with sites* $w \in e^n$ *and* $z \in \tilde{e}^n$ *such that* $w \sim z$.

of M_n^e particles. Two different demes $x, y \in \Gamma^n$ are said to be *neighbors* (denoted $x \sim y$) if either they lie on the same e and $d(x, y) = 1/L_n^e$ or if they are both adjacent to the same vertex $v \in V$.

Our n-th BVM is defined on the **discrete lattices** (**Figure 1**)

$$\Lambda_n := \cup_e \Lambda_n^e \qquad \text{where } \Lambda_n^e := \Big\{ (x, i) : \ x \in e^n, \ i \in \{1, \cdots, M_n^e\} \Big\}.$$

Points of Λ_n are called *sites*. Each site contains one individual agent/particle, which is of either type 1 or type 0. In the context of cancer dynamics, we think of an agent/particle as a biological cell, type 1 (cancer cell) and type 0 (normal cell). A site z is said to be on edge e (denoted $z \in e^n$) if $z \in \Lambda_n^e$; it is said to be in deme x if $z = (x, i)$ for some i. Two different sites $z, w \in \Lambda_n$ are said to be *neighbors* (also denoted $z \sim w$) if they are located at two neighboring demes.

Dynamics of BVM. Particles in deme x only interact with those in neighboring demes. Let $\xi_t(z) := \xi_t^n(z)$ be the type of the particle at site z at time t. Our BVM $(\xi_t)_{t \geq 0}$ can be constructed using two independent families of Poisson processes $\{P_t^{z,w} : z \sim w\}$ with rates $a^{z,w}$ and $\{\tilde{P}_t^{z,w} : z \sim w\}$ with rates $b^{z,w}$. At a jump time of $P_t^{z,w}$, the particle at z is replaced by an offspring of the one at w. At a jump time of $\tilde{P}_t^{z,w}$, the particle at z is replaced by an offspring of the one at w **only if** w **has type 1**, so there is a "bias" towards type 1. Under Assumption (1) the biased voter process $(\xi_t)_{t \geq 0} = (\xi_t^n)_{t \geq 0}$ is a Markov process with state space $\{0, 1\}^{\Lambda_n}$.

We suppress the superscript/lowerscript n's in ξ_t, $a^{z,w}$, $b^{z,w}$, L^e and M^e to simplify notation.

4. Main results. The following assumption, to be explained in Remark 4 right after the main result, is crucial to both Theorems 1 and 2. For two different directed edges $e, \tilde{e} \in E(v)$ where $v \in V$, we let $L^{e,\tilde{e}} = (d(x, v) + d(v, y))^{-1}$ where $x \in e^n$ and $y \in \tilde{e}^n$ are two neighboring demes in Γ^n. For $e = \tilde{e}$ we let $L^{e,e} = L^e$.

ASSUMPTION 4 (Symmetric conductances). *Suppose $\{C^n_{\tilde{e},e}\}_{e,\tilde{e}\in E(v),\,v\in V}$ are positive numbers satisfying symmetry $C^n_{e,\tilde{e}} = C^n_{\tilde{e},e}$ and*

$$(14) \qquad \sup_{v\in V}\ \sup_{e,\tilde{e}\in E(v)}\ \left| C^n_{\tilde{e},e} - L^{e,\tilde{e}}\,\Theta(\alpha_e,\alpha_{\tilde{e}}) \right| L^e \to 0, \qquad \text{as } n \to \infty,$$

for some symmetric continuous function Θ on $(0,\infty)^2$ such that $a \le \Theta(a,b) = \Theta(b,a) \le b$ for all $a \le b$ and that $\Theta(a,a) = a$ for all $a \in (0,\infty)$. In particular, $\frac{C^n_{e,e}}{L^e} \to \alpha_e$.

Examples include $\Theta(a,b) = \sqrt{ab}$ and power means $\left(\frac{a^p+b^p}{2}\right)^{1/p}$ with $p \in \mathbb{R} \setminus \{0\}$.

4.1. *Scaling limit of BVM.* The principal result in this paper says that the approximate densities of our BVM converge to SPDE (11) under suitable conditions, where the approximate density at deme $x \in e^n$ is defined by

$$(15) \qquad u^n_t(x) := \frac{1}{M^e} \sum_{i=1}^{M^e} \xi_t(x,i).$$

For $v \in V$, we define $u^n_t(v)$ to be the average value of $\{u^n_t(x)\}$ among demes x which are adjacent to v. We then linearly interpolate between demes (and also between vertices and demes) to define $u^n_t(x)$ for all $x \in \Gamma$. Then for all $t \ge 0$, we have $u^n_t \in \mathcal{C}_{[0,1]}(\Gamma)$, the set of continuous functions on Γ taking values in the interval $[0,1]$. Furthermore, if we equip $\mathcal{C}_{[0,1]}(\Gamma)$ with the metric

$$(16) \qquad \|\phi\| := \sum_{i=1}^{\infty} 2^{-i} \sup_{x\in K_i} |\phi(x)|$$

where $\{K_i\}$ is an increasing sequence of compact subsets of Γ with the union of the sequence being Γ. i.e., uniform convergence on compact sets, then $\mathcal{C}_{[0,1]}(\Gamma)$ is Polish and the paths $t \mapsto u^n_t$ are $\mathcal{C}_{[0,1]}(\Gamma)$-valued and càdlàg. Our main theorem is a weak convergence result for the sequence $\{u^n\}$ in the Skorohod space $D([0,\infty), \mathcal{C}_{[0,1]}(\Gamma))$.

THEOREM 1. *Let (Γ, d) be a metric graph satisfying Assumption 1. Let α, β and γ be piecewise constant functions as in Example 1, and $\hat{\beta}: V \to [0,\infty)$ be a non-negative bounded function on V. Suppose that as $n \to \infty$, the initial condition u^n_0 converges in $\mathcal{C}_{[0,1]}(\Gamma)$ to f_0, and that the following assumptions for L^e, M^e, a^{zw} and b^{zw} hold.*

(a) *$1/L^e \to 0$ uniformly for all $e \in E$ and (13) holds.*
(b) *$4\,L^e/M^e \to \gamma_e/\alpha_e$ for all $e \in E$.*
(c) *$a^{z,w} = 2\,C^n_{\tilde{e},e}L^e/M^e$ for all $z \in \tilde{e}^n$, $w \in e^n$ such that $z \sim w$, where $\{C^n_{\tilde{e},e}\}$ are positive numbers that satisfy Assumption 4.*
(d) *$b^{z,w} = B^n_e$ whenever $w \in e^n$ is not adjacent to a vertex in V and $z \sim w$, where $\{B^n_e\}$ are non-negative numbers such that $2\,B^n_e\,M^e \to \beta_e$.*
(e) *For each $v \in V$, let x^e_1 be the element on e^n adjacent to v and $x^e_2 \in e^n$ be adjacent to x^e_1. Suppose $\widehat{B}^n_{e,e} := b^{x^e_2,x^e_1}$ and $\widehat{B}^n_{e,\tilde{e}} := b^{x^e_1,x^{\tilde{e}}_1}$ for $\tilde{e} \ne e$ in $E(v)$ are non-negative numbers and*

$$(17) \qquad 4 \sum_{e\in E(v)} \sum_{\tilde{e}\in E(v)} \frac{\widehat{B}^n_{e,\tilde{e}}\,M^{\tilde{e}}}{L^e} \to \hat{\beta}(v).$$

Then the processes $(u_t^n)_{t \geq 0}$ in (15) *converge, in distribution in $D([0, \infty), \mathcal{C}_{[0,1]}(\Gamma))$, to a continuous $\mathcal{C}_{[0,1]}(\Gamma)$-valued process $(u_t)_{t \geq 0}$ which is the weak solution to the stochastic partial differential equation* (11) *with $u_0 = f_0$.*

We explain this result in the remarks below.

REMARK 3 (Identifying SPDE from microscopic rules). The significance of Theorem 1 lies in the connection it establishes between the microscopic BVM and the new macroscopic SPDE model that have fewer parameters. For example, the BVM is intractable to analyze or simulate when L^e or M^e is large, but with Theorem 1 one can take advantage of a new regularity (described by the SPDE, which is robust against the size of the particle system) that emerges.

To compare these two models we must relate microscale and macroscale parameters. The micro-parameters are $\{L^{e,\tilde{e}}\}$, $\{M^e\}$, $\{a^{z,w}\}$ and $\{b^{z,w}\}$. The macro-parameters can be found from micro-parameters as follows.

(i) $\alpha_e = 2 \lim_{n \to \infty} C_{e,e}^n / L^e$,
(ii) $\gamma_e = 4 \lim_{n \to \infty} \alpha_e L^e / M^e$,
(iii) $\beta_e = 2 \lim_{n \to \infty} B_e^n M^e$,
(iv) $\hat{\beta}(v) = 4 \lim_{n \to \infty} \sum_{e \in E(v)} \sum_{\tilde{e} \in E(v)} \widehat{B}_{e,\tilde{e}}^n M^{\tilde{e}} / L^e$.

These are generalizations of (2). Conversely, given $\alpha, \beta, \hat{\beta}$ and γ satisfying hypothesis of Theorem 1, there exist micro-parameters such that (i)-(iv) hold.

Using this connection, one can either obtain macro-parameters from microscopic (e.g. cell-level) measurements and experimental set up, or test hypothesis of microscopic interactions by using population-level measurements, or even perform model validation at both scales. Different edges can have different L_n^e or M_n^e, allowing the flexibility to model situations in which cells of different types and experimental configurations are situated on different edges [88]. This also enables us to take care of the case when a solution is simultaneously deterministic ($\gamma_e = 0$) on edge e and noisy ($\gamma_{\tilde{e}} > 0$) on another edge $\tilde{e}$.

REMARK 4 (Random walk approximations). The numbers $\{C_{e\tilde{e}}^n\}$ in Condition (c) of Theorem 1 arise naturally as the symmetric conductances of a random walk X^n which, under (14), converges in distribution to the m-symmetric diffusion X with $\ell = 1$ and $a = \alpha_e$ on $\mathring{e}$. More precisely, define the measure m_n on Γ^n by $m_n(x) := \frac{1}{L^e}$ whenever $x \in e^n$. Let $X^n := \{X_t^n\}_{t \geq 0}$ be the continuous time random walk (CTRW) on $\Gamma^n := \cup_e e^n$ associated with the Dirichlet form $\mathcal{E}^n(f, g)$ on $L^2(\Gamma^n, m_n)$, where

$$(18) \qquad \mathcal{E}^n(f, g) := \frac{1}{2} \sum_{x,y \in \Gamma^n} (f(x) - f(y))(g(x) - g(y)) C_{xy}^n \qquad \text{and}$$

$$(19) \qquad C_{xy}^n := \begin{cases} C_{e,\tilde{e}}^n & \text{if } x \in e^n, \ y \in \tilde{e}^n \text{ and } x \sim y \\ 0 & \text{if } x \text{ and } y \text{ are not adjacent in } \Gamma^n. \end{cases}$$

Observe that X^n is m_n-symmetric since $C_{xy}^n = C_{yx}^n$ for all $x, y \in \Gamma^n$. Under (14), we have $\mathcal{E}^n(f, g) \to \mathcal{E}(f, g)$ for $f, g \in \mathcal{C}_c^\infty(\Gamma)$. With extra work one can establish the weak convergence $X^n \to X$, as precisely stated in Lemma 3.

REMARK 5 (Uniform approximation). Condition (a) further implies a local central limit theorem (local CLT) and a uniform Holder continuity for X^n. The latter results, established in Theorems 4 and 5 in Section 5, will be used to obtain tightness of $\{u^n\}$ (Proposition 1). When $1/L^e \to 0$ *uniformly* for all $e \in E$ [for instance when Condition (a) is in force], we shall fix an edge $e* \in E$ and take $\varepsilon_n := 1/L^{e*}$, a representative rate at which every $1/L^e$ tends to zero. All $\{C^n_{x,y}\}_{x\sim y}$ are then of the same order $O(1/\varepsilon_n)$ as $n \to \infty$ by (14).

REMARK 6 (Generator $\mathcal{L}_n$). The transition rate of X^n from x to y is

$$\lambda(x,y) = C^n_{xy}/m_n(x) := C^n_{e\tilde{e}}L^e \quad \text{if } x \in e_n, \, y \in \tilde{e}_n, \, y \sim x.$$

Condition (14) implies that $|\lambda(x,y) - \alpha_e(L^e)^2| \to 0$ uniformly for $x, y \in e_n$ with $y \sim x$ and for all e. Hence the generator $\mathcal{L}_n$ of X^n_t can be approximated by

$$\mathcal{L}_n F(x) \approx \alpha_e \Delta_{L^e} F(x) \tag{20}$$

whenever $x \in e^n \smallsetminus \{x^e_1\}$, where Δ_{L^e} is the discrete Laplacian in (48), and

$$\mathcal{L}_n F(x^e_1) \approx \alpha_e(L^e)^2\big(F(x^e_2) - F(x^e_1)\big) + \sum_{\tilde{e}\in E(v):\tilde{e}\neq e}\big(F(x^{\tilde{e}}_1) - F(x^e_1)\big)C_{e,\tilde{e}}L^e \tag{21}$$

whenever x^e_1 is the element in e_n which is adjacent to a vertex and x^e_2 is the element in e_n which is adjacent to x^e_1. The approximations $\approx$ in (20) and (21) can be quantified by using Condition (14): the absolute difference between the left and the right is at most $o(\varepsilon_n)\|F\|_\infty$ where $o(\varepsilon_n)$ represents a term independent of F and which tends to 0 *uniformly* for all $x \in \Gamma^n$ faster than ε_n.

REMARK 7 (Local growth at $v \in V$). Results here for the simpler case $\hat{\beta} = 0$ (no extra birth on V) are already new. Condition (17) is crucially needed in (and only in) (69). It implies that, in order to have nontrivial boundary conditions $\hat{\beta} \neq 0$, the bias rates $b^{z,w}$ near vertices need to be of order at least $L^e/M^{\tilde{e}}$ which is typically higher than those in the interior of the edges. For example, $\hat{\beta} = 0$ if all $\{b^{z,w}\}$ are of the same order in n and so do all $\{M^e\}$. To try to give further interpretation, we suppose for simplicity that L^e are the same and that $\gamma_e > 0$ for all $e \in E$. Condition (b) and (17) roughly say that

$$\hat{\beta}(v) \approx 4 \sum_{e\in E(v)} \sum_{\tilde{e}\in E(v)} \widehat{B}^n_{e,\tilde{e}}\frac{\alpha_{\tilde{e}}}{\gamma_{\tilde{e}}} = 4 \sum_{\tilde{e}\in E(v)} \Big(\sum_{e\in E(v)} \widehat{B}^n_{e,\tilde{e}}\Big)\frac{\alpha_{\tilde{e}}}{\gamma_{\tilde{e}}},$$

where $\sum_{e\in E(v)} \widehat{B}^n_{e,\tilde{e}}$ can be interpreted as a local growth at v contributed by $\tilde{e}$.

REMARK 8. When $\Gamma = \mathbb{R}$, Theorem 1 reduces to [29, Theorem 1]. Under the same scalings, we can obtain the corresponding generalization of [29, Theorem 4].

4.2. *Interacting SDE as a numerical scheme.* Theorem 1 enables us to derive an SPDE which captures the macroscopic evolution of the density of particles in the BVM. See Remark 3. So when the number of particles are too large, one can simulate the more robust SPDE instead of the stochastic particle system. The next question is then: *how to simulate the SPDE?* It is known [84, 5, 64, 26, 62, 50] that SPDE can arise as the continuum scaling limit of interacting SDEs. These SDEs provide a numerical scheme for the solutions of the

SPDE and also the foundation for stochastic simulation algorithms (Gillespie algorithms [44, 45]).

In this section, we construct a system of interacting SDEs that offer a semi-discrete approximation to SPDE (11), where "semi-discrete" refers to the fact that the graph Γ is discretized into demes but populations in the demes are infinite. Our scheme utilizes the random walk X^n defined by (18)-(19).

1. **Specify step size.** Fix a sequence $\{h_n\}_{n\geq 1} \subset (0,\infty)$ which tends to 0.
2. **Discretize Γ.** Construct $\Gamma^n = \cup_e e^n$ as in Section 3, but now with $\frac{1}{L_n^e} = h_n$ for all $e \in E$. For $x \in e^n$ that is not adjacent to any vertex in V, we let I_x be the connected open interval $\{y \in \Gamma : d(y,x) < \frac{1}{2L^e}\}$.
3. **Independent Brownian motions.** For all $x \in e^n$ that is not adjacent to any vertex in V, we let $B_x = (B_x(t))_{t\geq 0}$ be a standard Brownian motion and that these Brownian motions $\{B_x\}_x$ are independent.
4. **Interacting SDE $\{U_x := U_x^n\}_{x\in\Gamma^n}$.** Consider the system of SDEs

$$(22) \qquad dU_x(t) = \left[\mathcal{L}_n U_x + \beta_e\, U_x(1-U_x)\right] dt + \sqrt{\gamma_e\, L^e\, U_x(1-U_x)}\, dB_x(t)$$

whenever $x \in e_n$ is not adjacent to any vertex in V, and

$$(23) \qquad dU_x(t) = \left[\mathcal{L}_n U_x + L^e\, \frac{\hat{\beta}(v)}{deg(v)}\, U_x(1-U_x)\right] dt$$

whenever $x \in e_n$ is adjacent to $v \in V$, where $\mathcal{L}_n$ is the generator of the CTRW X^n defined in Remark 4.

Our second result says that the interacting SDEs (22)-(23) and the SPDE (11) are close at least in distribution. Write $U_x^n = \left(U_x^n(t)\right)_{t\geq 0}$ and let $\{U_x^n\}_{x\in\Gamma^n}$ be a weak solution of (22)-(23) for the Itô SDE. We further interpolate over space to define U_x^n for all $x \in \Gamma$, as was done for (15). Let $\mathcal{C}([0,\infty), \mathcal{C}_{[0,1]}(\Gamma))$ be the space of continuous functions from $[0,\infty)$ to $\mathcal{C}_{[0,1]}(\Gamma)$, equipped with the topology of local uniform convergence.

THEOREM 2. *Let (Γ, d) be a metric graph satisfying Assumption 1. Let α, β and γ be piecewise constant functions as in Example 1, and $\hat{\beta} : V \to [0,\infty)$ be a non-negative bounded function on V. Suppose Condition (a) of Theorem 1 and Assumption 4 hold. Let $u(t,x)$ be the weak solution of SPDE (11) with initial condition $u_0 \in \mathcal{C}_{[0,1]}(\Gamma)$, and $\{U_x^n\}_{x\in\Gamma^n}$ be a weak solution to (22)-(23) with initial condition $\{U_x^n(0) = u_0(x)\}_{x\in\Gamma^n}$. Let $\{U_x^n\}_{x\in\Gamma}$ be the linear interpolation described above. Then as $n \to \infty$, the processes U^n converge to u in distribution in $\mathcal{C}([0,\infty), \mathcal{C}_{[0,1]}(\Gamma))$.*

The proof of Theorem 2 will be given in Section 7.

REMARK 9 (Coupling). The above construction gives a coupling of the interacting SDEs (22)-(23) and the SPDE (11) if we choose the Brownian motions to be

$$(24) \qquad\qquad B_x(t) := \sqrt{L^e}\, \dot{W}([0,t] \times I_x),$$

where $\dot{W}$ is the same white noise in (11), and if we can establish strong existence of solution for (22)-(23). We do not know whether $\sup_{x\in\Gamma^n} \mathbb{E}|U_x^n(t) - u(t,x)|^p$ converge to zero for some $p > 0$, under Standard L^p estimate as in Section 3 of Mueller [73] does not seem to work, due to the non-Lipschitz square root terms for U^n and u.

4.3. *Duality for stochastic FKPP with inhomogeneous coefficients.* Before turning to the proofs of Theorem 1 and Theorem 2, we settle the well-posedness of SPDE (11). While existence of weak solution follows from tightness (Proposition 1) and (72) for free, weak uniqueness requires a separate argument. We will establish a duality relation that implies weak uniqueness.

Duality between the standard stochastic FKPP and a branching coalescing Brownian motion is a known result due to Shiga [84] and [3]. Here we generalize this result to stochastic FKPP on a metric graph Γ, with inhomogeneous coefficients and nontrivial boundary conditions.

LEMMA 1 (Duality). *Suppose that Assumptions 1, 2 and 3 hold. Let u be a weak solution of the SPDE (10) with initial condition $u_0 \in \mathcal{C}_{[0,1]}(\Gamma)$, and $\{x_i(t) : 1 \le i \le n(t)\}$ be the positions at time t of a system of particles performing branching coalescing $\mathcal{E}$-diffusions on Γ, in which*

- *branching for a particle X_t occurs at rate $\beta(X_t)dt + \hat{\beta}(X_t)dL_t^V$, where L_t^V is the local time on the set of vertices V of X; i.e. the particle splits into two when the additive functional*

$$(25) \qquad \int_0^t \beta(X_s)ds + \hat{\beta}(X_s)dL_s^V$$

  exceeds an independent mean one exponential random variable.

- *two particles X_t, Y_t coalesce at rate $\frac{\gamma(X_t)}{\ell(X_t)} dL_t^{(X,Y)}$, where $L_t^{(X,Y)}$ is the local time of (X_t, Y_t) on the diagonal $\{(x,x) : x \in \Gamma\}$; i.e. the two particles become one when the additive functional*

$$(26) \qquad \int_0^t \frac{\gamma(X_s)}{\ell(X_s)} dL_s^{(X,Y)}$$

  exceeds an independent mean one exponential random variable.

Then we have the duality formula

$$(27) \qquad \mathbb{E} \prod_{i=1}^{n(0)} \big(1 - u_t(x_i(0))\big) = \mathbb{E} \prod_{i=1}^{n(t)} \big(1 - u_0(x_i(t))\big) \quad \text{for all } t \ge 0.$$

The local times L^V and $L^{(X,Y)}$ are defined as the positive continuous additive functionals (PCAF) corresponding to, respectively, the Revuz measures (Chapter 4 of [13]) $\sum_{v \in V} \delta_v$ on Γ and m_* on $\Gamma \times \Gamma$, where m_* is supported on the diagonal defined by $m(A) = m_*\{(x,x) : x \in A\}$. With this new dual process, a proof of (27) follows by a modification of Section 8.1 of [29] and heat kernel estimates of X (Theorem 5); more details of the proof are given in the Appendix.

Lemma 1 is useful to obtaining distributional properties of (10). In particular it yields the following weak uniqueness.

LEMMA 2 (Weak uniqueness). *Suppose Assumptions 1, 2 and 3 hold. Then there exists a unique weak solution to SPDE (10). In particular, (11) has a unique weak solution.*

Proofs. The rest of the paper is devoted to the proofs of Theorems 1 and 2. We start with some uniform heat kernel estimates of independent interest.

5. Uniform heat kernel estimates. In this Section, we establish some uniform estimates for the transition densities for both the random walks X^n (defined by (18)-(19)) and the $\mathcal{E}$-diffusion X. The key point here is that the constants involved do not depend on n, which in particular implies the local CLT. Besides having independent interest, these results are essential to the proof of tightness in Theorem 1 and of Theorem 2.

5.1. *Invariance principle.* As pointed out in Remark 4, under (14) we have the following generalization of Donsker's invariance principle. Recall that ν defined in (5) is equal to m when $\ell = 1$.

LEMMA 3. *(Invariance principle) Let (Γ, d) be a metric graph satisfying Assumption 1 and ℓ, α be given in Example 1. Suppose Condition (a) of Theorem 1 and Assumption 4 hold. Suppose, as $n \to \infty$, the law of X_0^n converges weakly to μ. Then for all $T \in (0, \infty)$ the random walks X^n converge in distribution in the Skorohod space $D([0,T], \Gamma)$ to X, the m-symmetric $\mathcal{E}$-diffusion with initial distribution μ.*

Lemma 3 follows from the more general result [2, Theorem 1]. The latter gives invariance principles for random walks on random trees based on Dirichlet form methods.

5.2. *Discrete heat kernel.* Let $p^n(t, x, y)$ be the transition density of the random walk X^n defined in (18)-(19), with respect to its measure m_n. That is,

(28)
$$p^n(t, x, y) := \frac{\mathbb{P}(X_t^n = y \mid X_0^n = x)}{m_n(y)}.$$

Then $p^n(t, x, y) = p^n(t, y, x)$ for all $t \geq 0$, $x, y \in \Gamma^n$. Recall from Remark 5 that ε_n is a representative rate at which every $1/L^e$ tends to zero.

THEOREM 3. *Suppose Assumptions 1, 2 and 4 and Condition (a) of Theorem 1 hold. Then the transition densities $p^n(t, x, y)$ enjoy the following uniform estimates: For any $T \in (0, \infty)$, there exist positive constants $\{C_k\}_{k=1}^{7}$ and σ such that for all $n \in \mathbb{N}$ and $x, y \in \Gamma^n$, the followings hold.*

1. *("Gaussian" upper bounds)*

(29)
$$p^n(t, x, y) \leq \frac{C_1}{\varepsilon_n \vee t^{1/2}} \exp\left(-C_2 \frac{|x-y|^2}{t}\right) \quad \text{for } t \in [\varepsilon_n, T] \quad and$$

(30)
$$p^n(t, x, y) \leq \frac{C_3}{\varepsilon_n \vee t^{1/2}} \exp\left(-C_4 \frac{|x-y|}{t^{1/2}}\right) \quad \text{for } t \in (0, \epsilon_n];$$

2. *(Gaussian lower bound)*

$$(31) \qquad p^n(t,x,y) \geq \frac{C_5}{\varepsilon_n \vee t^{1/2}} \exp\left(-C_6 \frac{|x-y|^2}{t}\right) \quad for \ t \in (0,T];$$

3. *(Hölder continuity)*

$$(32) \qquad |p^n(t,x,y) - p^n(t',x',y')| \leq C_7 \frac{(|t-t'|^{1/2} + |x-x'| + |y-y'|)^\sigma}{(t \wedge t')^{(1+\sigma)/2}}$$

for all $(t,x,y), (t',x',y') \in (0,T] \times \Gamma^n \times \Gamma^n$.

REMARK 10. It is known (for instance [25, 12]) that the standard Gaussian upper bound fails to hold for small time for continuous time random walks, but we can use the weaker estimate (30). This small time caveat is not present for the diffusion. See Theorem 5.

PROOF. The proof is an application of the famous De Giorgi-Moser-Nash theory to the metric graph setting. It is known, for manifold [82] and for discrete graphs [25], that the two sided Gaussian estimates for reversible Markov chains is equivalent to the parabolic Harnack inequality (PHI), that these estimates are characterized by geometric conditions, namely volume doubling (VD) plus Poincaré inequality (PI), and that (PHI) implies Holder inequality. See [63] for a review of results in this area.

By [25, Theorem 1.7], all these estimates hold for p_n for fixed n, but the key point here is that the *constants involved do not depend on n*. Recall that Assumptions 1, 2 and 4 and Condition (a) of Theorem 1 are enforced throughout. These clearly give

$$(33) \qquad \inf_{n \geq 1} \inf_{(x,y):x \sim y} \frac{C_{x,y}^n}{C_x^n} > 0 \quad \text{where } C_x^n := \sum_{z:z \sim x} C_{x,z}^n,$$

which implies that the geometric condition $\Delta(\alpha) > 0$ in [25, Theorem 1.7] holds uniformly in n.

For the rest of the proof we argue that a uniform (in n) PHI holds and that PHI implies all the desired estimates for p^n. For simplicity, we assume $L_n^e = L_n$ for all $e \in E$. Then $1/L_n$ is the common rate at which $1/L^e$ tends to zero. We also assume the functions α and ℓ are positive constants. The proof for Theorem 3 remains valid without this simplification: the constants will be different but still independent of n.

Now $m_n(x) = 1/L_n$ for all $x \in \Gamma^n$ and $n \geq 1$. Furthermore, $C_{x,y}^n \asymp L_n$ in the sense that for some constants $\kappa, \tilde{\kappa} \in (0,\infty)$ independent of $x,y \in \Gamma^n$ and $n \geq 1$, we have $\kappa L_n \leq C_{xy}^n \leq \tilde{\kappa} L_n$. From this, we have

$$Vol_n(A) \asymp L_n^2 m_n(A),$$

where $Vol_n(A) = \sum_{x \in A} C_x^n$ is the "volume" in the graph setting in [25]. Denote by $B(x,r) := \{y \in \Gamma : d(x,y) < r\}$ a ball of the original metric graph Γ. Note that our ball $B(x,r) \cap \Gamma^n$ is approximately a rescaled version of the usual and more discrete notion as that in [25]. Precisely, for $x \in \Gamma^n$, we let

$$B_n(x,r) := \{y \in \Gamma^n : \#(x,y) < r\}$$

18

and $\#(x,y)$ is the number of edges in Γ^n in a shortest path connecting x to y. Then due to (13), for $r > 3/L_n$ and n large enough such that $L_n > 2$, we have

$$(34) \qquad B_n\big(x, r(L_n - 2)\big) \subset B(x,r) \cap \Gamma^n \subset B_n\big(x, r(L_n + 2)\big).$$

Using (5.2), (34) and our standing assumptions, we can verify the followings two geometric conditions while keeping track of the constants: volume-doubling property and the Poincaré inequality. That is, there are constants κ_1, $\kappa_2 > 0$ such that *for all* $n \geq 1$, we have respectively

$$(\text{VD}[\kappa_1]) \qquad m_n(B(x,2r) \cap \Gamma^n) \leq \kappa_1 \, m_n(B(x,r) \cap \Gamma^n)$$

for all $x \in \Gamma^n$ and $r > 0$, and

$$(\text{PI}[\kappa_2]) \qquad \sum_{y \in B(x,r) \cap \Gamma^n} |f(y) - \bar{f}_B|^2 \, m_n(y) \leq \kappa_2 \, r^2 \, \mathcal{E}_n(f \, 1_{B(x,2r)})$$

for all functions $f : \Gamma^n \to \mathbb{R}$, points $x \in \Gamma^n$ and $r > 0$, where $\bar{f}_B := m_n^{-1}(B) \sum_{z \in B} f(z) \, m_n(z)$ is the average value of f over $B = B(x,r) \cap \Gamma^n$.

Equipped with $(\text{VD}[\kappa_1])$ and $(\text{PI}[\kappa_2])$ as well as (33), the Moser iteration argument as detailed in [25, Section 2] yields the parabolic Harnack inequality for some constant $\kappa_{\mathcal{H}} > 0$ that is independent of n. That is, for all $x_0 \in \Gamma$, $s \in \mathbb{R}$, $r > 0$ and non-negative solution U_n of the parabolic equation $\partial_t U_n = \mathcal{L}_n U_n$ on the space-time rectangle $Q := [s, \, s+r^2] \times B_n(x_0, rL_n)$, we have

$$(\text{PHI}[\kappa_{\mathcal{H}}]) \qquad \sup_{Q_-} U_n \leq \kappa_{\mathcal{H}} \inf_{Q_+} U_n,$$

where $Q_- := \big[s + \frac{r^2}{4}, \, s + \frac{r^2}{2}\big] \times B_n(x_0, \frac{r}{2}L_n)$ and $Q_+ := \big[s + \frac{3r^2}{4}, \, s + r^2\big] \times B_n(x_0, \frac{r}{2}L_n)$.

The proof of $(\text{PHI}[\kappa_{\mathcal{H}}])$ follows from [25, Section 2] which is an adaptation of [82]. Roughly, $(\text{VD}[\kappa_1])$, $(\text{PI}[\kappa_2])$ and (33) imply a weighted Poincaré inequality [25, Proposition 2.2] and a Sobolev inequality [25, Proposition 2.4]. The Sobolev inequality and $(\text{VD}[\kappa_1])$ yield mean value inequalities for positive sub- and super-solutions [25, Lemma 2.7]. The weighted Poincaré inequality, on other hand, yields information on the size of $\log v$ in [25, Lemma 2.8] for any positive super-solution $v > 0$. This information and [25, Lemma 2.9] allows us to stick together mean value inequalities on v and $1/v$, giving the parabolic Harnack inequality for positive super-solutions $v > 0$. This and the mean value inequality for sub-solutions finally give the desired $(\text{PHI}[\kappa_{\mathcal{H}}])$ for non-negative solutions. All the constants involved in these inequality are independent of $n \geq 1$. Since the argument is standard, we refer the reader to [25, Section 2] for the details of the proof.

Following standard arguments such as those in [82, Section 3.3] or [86], one can show that $(\text{PHI}[\kappa_{\mathcal{H}}])$ implies the desired inequalities (29)-(32). Indeed, the Gaussian upper bounds (29)-(30) follow from [25, Propositions 3.1 and 3.4], since the constants in the latter results depend only on $\kappa_{\mathcal{H}}$. [25, Propositions 3.1] also gives the on-diagonal lower bound that is uniform in n. This uniform (in n) on-diagonal bound then gives the desired Gaussian lower bound (31) by a standard chaining argument [86, page 329]. Finally, the Hölder inequality (32) is implied by the Gaussian bounds (29)-(31) by a standard oscillation argument [86, Theorem II.1.8-Corollary II.1.9]. $\qquad \square$

The uniform Hölder continuity (32) together with the invariance principle imply the following local central limit theorem.

THEOREM 4. *(Local CLT) Suppose all assumptions in Theorem 3 hold. Then*

$$\lim_{n \to \infty} \sup_{t \in I} \sup_{x,y \in K \cap \Gamma^n} \left| p^n(t,x,y) - p(t,x,y) \right| = 0$$

for any compact interval $I \subset (0,\infty)$ and compact subset $K \subset \Gamma$, where p is the transition density (6) of an $\mathcal{E}$-diffusion with ℓ and α given in Example 1.

PROOF. The proof follows from a compactness argument, as detailed in [12, Theorem 2.12]. For each $n \geq 1$ and $t > 0$, we extend $p^n(t,\cdot,\cdot)$ to a continuous non-negative function on $\Gamma \times \Gamma$ by interpolation; see how it was done for (15). We then consider the family $\{t^{1/2}p^n\}_{n \geq 1}$ of continuous functions on $(0,\infty) \times \Gamma \times \Gamma$. The Gaussian upper bound and the Hölder continuity in Theorem 3 give us uniform pointwise bound and equi-continuity respectively. Hence, by the Arzela-Ascoli Theorem, any sub-sequence has a further sub-sequence that converges to a continuous function $q : (0,\infty) \times \Gamma \times \Gamma \longrightarrow [0,\infty)$ locally uniformly. Lemma 3 and the continuity of the transition density p then characterize the limit.   □

Related results and ideas can be found in Croydon and Hambly [21] who investigated general conditions under which the local CLT for random walks on graphs is implied by weak convergence.

5.3. *Heat kernel for diffusions on Γ.* Hölder continuity of $p(t,x,y)$ then follows directly from the local CLT and (32). The two-sided Gaussian bounds for $p(t,x,y)$ do *not* directly follow from Theorem 3, but we can establish them in the same way. In fact we will establish these estimates for general $\mathcal{E}$-diffusion rather than only those in Example 1: The volume-doubling property we need for the diffusion is exactly stated in Assumption 1. Recall that $\nu(dx) = \ell(x)m(dx)$. The PI in Assumption 1 implies that

$$\int_{B(x,r)} |f(y) - \bar{f}_B|^2 \, \nu(dy) \leq (\sup_{\Gamma} \ell) \int_{B(x,r)} |f(y) - \bar{f}_B|^2 \, m(dy)$$

(35)
$$\leq \frac{C_{PI} \, (\sup_{\Gamma} \ell)}{\inf_{\Gamma} \alpha} \, r^2 \, \mathcal{E}\big(f \, 1_{B(x,2r)}\big)$$

for all $f \in W^{1,2}(\Gamma,m)$, $x \in \Gamma$, $r > 0$. With the VD (3) and PI (35), it is well-known (see for instance [47, 51] and the references therein) that the transition density $p(t,x,y)$ satisfies the parabolic Harnack inequality, which is equivalent to two sided Gaussian estimates and implies the Hölder continuity.

We summarize these important properties about transition densities $p(t,x,y)$ of diffusions on graphs in the following theorem. These properties applies to the diffusions considered in [42, 41, 9, 10].

THEOREM 5. *Suppose Assumptions 1 and 2 hold. Then the transition density of the $\mathcal{E}$-diffusion on Γ, defined in (6), enjoys the following properties: For any $T \in (0,\infty)$, there exist positive constants $\{C_k\}_{k=1}^5$ and σ such that we have*

1. *(Two-sided Gaussian bounds)*

$$(36) \qquad \frac{C_1}{t^{1/2}} \exp\left(-C_2 \frac{|x-y|^2}{t}\right) \leq p(t,x,y) \leq \frac{C_3}{t^{1/2}} \exp\left(-C_4 \frac{|x-y|^2}{t}\right)$$

for all $x, y \in \Gamma$ and $t \in (0,T]$; and

2. *(Hölder continuity)*

$$(37) \qquad |p(t,x,y) - p(t',x',y')| \leq C_5 \frac{(|t-t'|^{1/2} + |x-x'| + |y-y'|)^\sigma}{(t \wedge t')^{(1+\sigma)/2}}$$

for all $(t,x,y), (t',x',y') \in (0,T] \times \Gamma \times \Gamma$.

Theorem 5 implies many useful properties of the diffusion X, including exit time estimates and strong continuity of the semigroup $\{P_t\}$ on $C_b(\Gamma)$, the space of bounded continuous functions with local uniform norm.

6. Proof of Theorem 1. Equipped with Theorems 3-5, we can follow the outline of the proofs in [77, 29] to finish the proof of Theorem 1. We shall emphasize new terms and new difficulties that did not appear in [77, 29] in our calculations. The dynamics of $(\xi_t)_{t \geq 0}$ is concisely described by the equation

$$\xi_t(z) = \xi_0(z) + \sum_{w \sim z} \int_0^t (\xi_{s-}(w) - \xi_{s-}(z)) \, dP_s^{z,w}$$

$$(38) \qquad\qquad\qquad + \sum_{w \sim z} \int_0^t \xi_{s-}(w)(1 - \xi_{s-}(z)) \, d\tilde{P}_s^{z,w}.$$

In the space-time graphical representation (see, for instance [29, Fig. 1]), we draw an arrow $z \leftarrow w$ when there is a jump for the Poisson processes.

6.1. *Approximate martingale problem.* For functions $f, g : \Gamma^n \to \mathbb{R}$, we write

$$(39) \qquad \langle f, g \rangle_e := \frac{1}{L^e} \sum_{x \in e^n} f(x) g(x) \quad \text{and} \quad \langle f, g \rangle := \langle f, g \rangle_{m_n} := \sum_e \langle f, g \rangle_e.$$

We also identify f with a function on Λ_n by setting $f(z) := f(x)$ when $z = (x, i)$.

Let $\phi : [0, \infty) \times \Gamma \to \mathbb{R}$ be a continuous function that is continuously differentiable in t, twice continuously differentiable and has compact support in $x \in \Gamma$ and satisfies the boundary condition
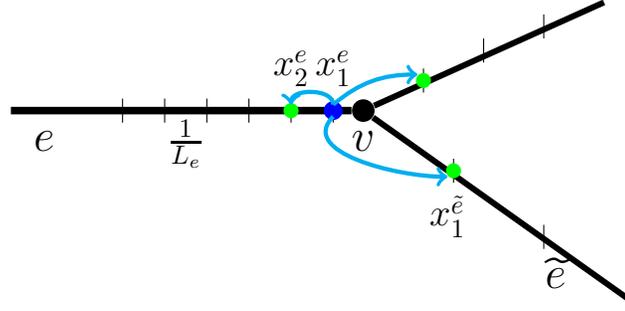
$$(40) \qquad\qquad (\nabla_{out} \phi_t \cdot [\alpha])(v) = 0 \quad \text{for } t \geq 0,$$

where we write $\phi_t(x) = \phi(t, x)$ interchangeably.

Applying integration by parts to $\xi_t(z)\phi_t(z)$, using (38), and summing over x, we obtain for all $t > 0$ and edge e, we have

$$\langle u_t^n, \phi_t \rangle_e - \langle u_0^n, \phi_0 \rangle_e - \int_0^t \langle u_s^n, \partial_s \phi_s \rangle_e \, ds$$

$$(41) \qquad = (M^e L^e)^{-1} \sum_{z \in \Lambda_n^e} \sum_{w \sim z} \int_0^t (\xi_{s-}(w) - \xi_{s-}(z)) \phi_s(z) \, dP_s^{z,w}$$

$$(42) \qquad + (M^e L^e)^{-1} \sum_{z \in \Lambda_n^e} \sum_{w \sim z} \int_0^t \xi_{s-}(w)(1 - \xi_{s-}(z)) \phi_s(z) \, d\tilde{P}_s^{z,w}.$$

FIG 2. *Illustration of demes x_k^e on the discretized graph Γ^n*

Note that w in (41) and (42) can be on a different edge $\tilde{e} \in E(v) \smallsetminus e$ if z is adjacent to a vertex v.

Outline. To describe the local behavior near a vertex and to simplify the presentation, we first consider the case when Γ consists of $deg(v)$ positive half real lines starting from a single common vertex v. For each $e \in E(v)$ we enumerate the set e^n as $(x_1^e, x_2^e, x_3^e, \cdots)$ along the direction of e away from v, so x_1^e is closest to v. We also identify $x_k^e + (L^e)^{-1}$ with x_{k+1}^e. See **Figure 2**. Observe that when both z and w are on e but *not* in x_1^e, then $a^{z,w} = a^{w,z}$. Hence we split the double sum in (41) as

$$(43) \qquad \sum_{z \in \Lambda_n^e} \sum_{w \sim z} \mathbf{1}_{\{w,z \notin x_1^e\}} + \Big(\sum_{z \in x_1^e} \sum_{w \sim z} + \sum_{z \in x_2^e} \sum_{w \in x_1^e} \Big) \coloneqq (41)(i) + (41)(ii),$$

where $(41)(ii)$ is the term $(\sum_{z \in x_1} \sum_{w \sim z} + \sum_{z \in x_2} \sum_{w \in x_1})$. The sum in $(41)(i)$ is symmetric and we shall makes use the gluing condition (40) to treat this term; while treatment of the (boundary) term $(41)(ii)$ required the choices of parameters in Conditions (a)-(c) of Theorem 1. The next 3 subsections are computations for terms $(41)(i)$, $(41)(ii)$ and (42) respectively, which lead to the martingale problem for u^n.

As we shall see, since ϕ has compact support and the branch lengths of Γ is bounded below by a positive constant (Assumption 1), all sums involving ϕ are finite sums and the proofs in Subsections 6.1-6.3 work equally well for any graph satisfying Assumption 1.

6.1.1. *Term $(41)(i)$: White noise, Laplacian and Glueing condition.* For $(41)(i)$, some ideas in [29, Section 5] can be reused which we now briefly describe. Let $\xi_t^c(z) \coloneqq 1 - \xi_t(z)$ and rewrite the integrand of (41) as

$$(\xi_{s-}(w) - \xi_{s-}(z))\, \phi_s(z)$$
$$= [\xi_{s-}(w)\, \xi_{s-}^c(z) - \xi_{s-}(z)\, \xi_{s-}^c(w)]\, \phi_s(z)$$
$$(44) \qquad = \xi_{s-}(w)\, \xi_{s-}^c(z)\phi_s(w) - \xi_{s-}(z)\, \xi_{s-}^c(w)\phi_s(z)$$
$$(45) \qquad + \xi_{s-}(w)\, \xi_{s-}^c(z)\, \big(\phi_s(z) - \phi_s(w)\big).$$

As we will see, the white noise will come from (44) while the Laplacian term and the gluing term come from (45). To simplify notation we denote $A_{\tilde{e},e}^n \coloneqq 2\, C_{\tilde{e},e}^n L^e / M^e$, so $a^{z,w} = A_{\tilde{e},e}^n$ in Condition (c) in the statement of Theorem 1.

White noises. We first work with (44). Interchanging the roles of z and w in the first double sum and writing $Q_s^{z,w} = P_s^{w,z} - P_s^{z,w}$, this part of (41)(i) becomes a martingale

$$(46) \qquad Z_t^e(\phi) := (M^e L^e)^{-1} \sum_{z \in \Lambda_n^e} \sum_{w \sim z} \mathbf{1}_{\{w, z \notin x_1\}} \int_0^t \xi_{s-}(z) \xi_{s-}^c(w) \phi_s(z) \, dQ_s^{z,w}.$$

Since $a^{z,w} = a^{w,z}$ and the quadratic variance process $\langle Q^{z,w} \rangle_t = 2A_{e,e}^n t$, we have quadratic variation $\langle Z^e(\phi) \rangle_t$ converges to

$$\gamma_e \int_0^t \int_e u_s(x)(1 - u_s(x)) \phi_s(x)^2 \, m(dx) \, ds$$

since $A_{e,e}^n / L^e \to \gamma_e/4$ by Conditions (a)-(c) and by the smoothness assumption of ϕ, assuming $\mathcal{C}$-tightness of $\{u^n\}$. Details for this part is the same as those in [29, Section 5]. *Here, and in what follows in this Subsection, the claimed convergences follow once we have proved $\mathcal{C}$-tightness. See Subsection 6.3 for the proof, which does not use any of the convergences claimed in this Subsection. See Remark 11.*

Laplacian and gluing condition. Next we work with (45). We write $L = L^e$ and $M = M^e$ for simplicity when there is no confusion, and denote the discrete gradient and the discrete Laplacian respectively by

$$(47) \qquad \nabla_L f(x) := L \left(f(x + L^{-1}) - f(x) \right) \qquad\qquad \text{if } x \in e^n,$$

$$(48) \qquad \Delta_L f(x) := L^2 \left(f(x + L^{-1}) + f(x - L^{-1}) - 2f(x) \right) \quad \text{if } x \in e^n \setminus \{x_1^e\}.$$

We break (45) into the average terms and a fluctuation term

$$(49) \qquad (ML)^{-1} \sum_{z \in \Lambda_n^e} \sum_{w \sim z} \mathbf{1}_{\{w, z \notin x_1\}} \int_0^t \xi_{s-}(w) \xi_{s-}^c(z) [\phi_s(z) - \phi_s(w)] A_{e,e}^n \, ds$$

$$(50) \qquad + (ML)^{-1} \sum_{z \in \Lambda_n^e} \sum_{w \sim z} \mathbf{1}_{\{w, z \notin x_1\}} \int_0^t \xi_{s-}(w) \xi_{s-}^c(z) [\phi_s(w) - \phi_s(z)] (dP_s^{z,w} - A_{e,e}^n \, ds).$$

We can replace ξ_{s-}^c by 1 in (49) without changing its value, because by symmetry (valid since we are summing over the *same* set for z and w),

$$\sum_z \sum_{w \sim z} \xi_{s-}(w) \xi_{s-}(z) [\phi_s(z) - \phi_s(w)] = 0 \qquad \text{for all } s > 0.$$

Doing the double sum over w and then over $z \sim w$, and recalling that we identify $x_k^e + (L^e)^{-1}$ with x_{k+1}^e, we see that the integrand of (49) is

$$(M^e L^e)^{-1} \sum_{z \in \Lambda_n^e} \sum_{w \sim z} \mathbf{1}_{\{w, z \notin x_1^e\}} \xi_{s-}(w) [\phi_s(z) - \phi_s(w)] A_{e,e}^n$$

$$= \frac{A_{e,e}^n M}{L^3} \sum_{x \geq x_3^e} u_{s-}^n(x) \Delta_L \phi_s(x) + \frac{A_{e,e}^n M}{L} u_{s-}^n(x_2^e) [\phi_s(x_3^e) - \phi(x_2^e)]$$

$$= \frac{A_{e,e}^n M}{L^3} \sum_{x \geq x_2^e} u_{s-}^n(x) \Delta_L \phi_s(x) - \frac{A_{e,e}^n M}{L} u_{s-}^n(x_2^e) [\phi_s(x_1^e) - \phi(x_2^e)]$$

$$(51) \qquad = \frac{A_{e,e}^n M}{L^2} \left\{ \left(\frac{1}{L} \sum_{x \geq x_2^e} u_{s-}^n(x) \Delta_L \phi_s(x) \right) + u_{s-}^n(x_2^e) \nabla_L \phi_s(x_1^e) \right\}.$$

By our assumption on ϕ, $\Delta_L\phi_s$ converges to $\Delta\phi_s$ uniformly on compact subsets of e and $\lim_{n\to\infty}\frac{A^n_{e,e}M}{L^2}=\alpha_e$, so (49) converges to

$$(52)\qquad \alpha_e\int_0^t\Big(\int_e u_s(x)\Delta\phi_s(x)\,dx\Big)+u_s(v)\,\nabla_{e+}\phi_s(v)\,ds.$$

Using the *gluing assumption* (40), we see that the last term will vanish upon summation over $e\in E(v)$. The other term, (50), is a martingale $E_t^{(1)}(\phi)$ with

$$\langle E^{(1)}(\phi)\rangle_t \le \frac{A^n_{e,e}}{(ML)^2}\sum_z\sum_{w\sim z}\int_0^t(\phi_s(z)-\phi_s(w))^2\,ds$$

$$=\frac{2A^n_{e,e}}{L^2}\int_0^t\langle 1,|\nabla_L\phi_s|^2\rangle_e\,ds\to 0$$

since $A^n_{e,e}/L^e\to\gamma/4$ and $L^e\to\infty$, by (14).

6.1.2. *Term* (41)(*ii*): *Matching boundary condition.* We now deal with (41)(*ii*) which are the remaining terms $\sum_{z\in x_1^e}\sum_{w\sim z}+\sum_{z\in x_2^e}\sum_{w\in x_1^e}$ in (43). Our goal is to show that this term converges to zero under the choice for C^n_{xy}'s specified in condition (c) of Theorem 1. For this we further take a summation over all edges in $E(v)$ to obtain

$$\sum_{e\in E(v)}(M^eL^e)^{-1}\sum_{z\in x_1^e}\sum_{w\sim z}\int_0^t(\xi_{s-}(w)-\xi_{s-}(z))\phi_s(z)\,dP_s^{z,w}$$

$$+\sum_{e\in E(v)}(M^eL^e)^{-1}\sum_{z\in x_2^e}\sum_{w\in x_1^e}\int_0^t(\xi_{s-}(w)-\xi_{s-}(z))\phi_s(z)\,dP_s^{z,w}$$

$$(53)\qquad =\sum_{e\in E(v)}(M^eL^e)^{-1}\sum_{\tilde e\in E(v)\setminus e}\sum_{z\in x_1^e}\sum_{w\in x_1^{\tilde e}}\int_0^t(\xi_{s-}(w)-\xi_{s-}(z))\phi_s(z)\,dP_s^{z,w}$$

$$(54)\qquad +\sum_{e\in E(v)}(M^eL^e)^{-1}\Big(\sum_{z\in x_2^e}\sum_{w\in x_1^e}+\sum_{w\in x_2^e}\sum_{z\in x_1^e}\Big)\int_0^t\big(\xi_{s-}(w)-\xi_{s-}(z)\big)\phi_s(z)\,dP_s^{z,w}.$$

We break (53) into an average term and a fluctuation term

$$(55)\qquad \sum_{e\in E(v)}(M^eL^e)^{-1}\sum_{\tilde e\in E(v)\setminus e}\sum_{z\in x_1^e}\sum_{w\in x_1^{\tilde e}}\int_0^t(\xi_{s-}(w)-\xi_{s-}(z))\phi_s(z)\,A^n_{e,\tilde e}\,ds$$

$$(56)\qquad +\sum_{e\in E(v)}(M^eL^e)^{-1}\sum_{\tilde e\in E(v)\setminus e}\sum_{z\in x_1^e}\sum_{w\in x_1^{\tilde e}}\int_0^t(\xi_{s-}(w)-\xi_{s-}(z))\phi_s(z)\big(dP_s^{z,w}-A^n_{e,\tilde e}\,ds\big).$$

Grouping terms in *unordered pairs* of distinct elements in $E(v)$, the average term (55) is

$$\sum_{e\in E(v)}\sum_{\tilde e\in E(v)\setminus e}\frac{A^n_{e,\tilde e}}{M^eL^e}\sum_{z\in x_1^e}\sum_{w\in x_1^{\tilde e}}\int_0^t(\xi_{s-}(w)-\xi_{s-}(z))\phi_s(z)\,ds$$

$$=\sum_{e\in E(v)}\sum_{\tilde e\in E(v)\setminus e}2C^n_{e,\tilde e}\int_0^t\big(u^n_{s-}(x_1^{\tilde e})-u^n_{s-}(x_1^e)\big)\phi_s(x_1^e)\,ds$$

$$(57)\qquad =\int_0^t\sum_{(e,\tilde e)\text{unordered}}\big(2C^n_{e,\tilde e}-2C^n_{\tilde e,e}\big)\phi_s(x_1^e)\big(u^n_{s-}(x_1^{\tilde e})-u^n_{s-}(x_1^e)\big)\,ds+err_1(t)$$

$$(58)\qquad =0+err_1(t),$$

where

$$(59) \qquad err_1(t) = \int_0^t \sum_{(e,\tilde{e})\text{unordered}} 2C_{\tilde{e},e}^n \left(\phi_s(x_1^e) - \phi_s(x_1^{\tilde{e}})\right)\left(u_{s-}^n(x_1^{\tilde{e}}) - u_{s-}^n(x_1^e)\right)ds.$$

The equality (58) says that the first sum in (57) is zero. To see this, we group the terms of this sum in pairs in such a way that $\sum_{(e,\tilde{e})\text{unordered}}$ is over all unordered pairs of distinct elements in $E(v)$. The last equality (58) then follows by symmetry of the conductances.

The error term $err_1(t)$ is $o(1)$, i.e. tends to 0 as $n \to \infty$, since $\sup_{s\in[0,T]}\sup_{n\geq 1} L^e|\phi_s(x_1^e) - \phi_s(v)| < C_T$ and (by Remark 5)

$$(60) \qquad \sup_{n\geq 1} \sum_{e\in E(v)} \sum_{\tilde{e}\in E(v)} C_{\tilde{e},e}^n \left(\frac{1}{L^e} + \frac{1}{L^{\tilde{e}}}\right) < \infty.$$

The fluctuation term (56) has quadratic variation

$$\sum_{e\in E(v)} \sum_{\tilde{e}\in E(v)\setminus e} \frac{A_{e,\tilde{e}}^n}{(M^e L^e)^2} \sum_{z\in x_1^e} \sum_{w\in x_1^{\tilde{e}}} \int_0^t \left(\xi_{s-}(w) - \xi_{s-}(z)\right)^2 \phi_s^2(z)\, ds$$

$$\leq \sum_{e\in E(v)} \sum_{\tilde{e}\in E(v)\setminus e} \frac{2C_{e,\tilde{e}}^n}{L^e} \frac{M^{\tilde{e}}}{M^e} \int_0^t \left(\frac{u_{s-}^n(x_1^{\tilde{e}})}{M^e} + \frac{u_{s-}^n(x_1^e)}{M^{\tilde{e}}}\right)\phi_s^2(x_1^e)\, ds$$

$$\leq \sum_{e\in E(v)} \sum_{\tilde{e}\in E(v)\setminus e} \frac{2C_{e,\tilde{e}}^n}{L^e} \frac{M^{\tilde{e}}}{M^e}\left(\frac{1}{M^e} + \frac{1}{M^{\tilde{e}}}\right) \int_0^t \phi_s^2(x_1^e)\, ds$$

which tends to 0 as $n \to \infty$ since

$$(61) \qquad \sum_{(e,\tilde{e})} \frac{C_{e,\tilde{e}}^n M^{\tilde{e}}}{L^e M^e}\left(\frac{1}{M^e} + \frac{1}{M^{\tilde{e}}}\right) \leq C \sum_{(e,\tilde{e})} \frac{M^{\tilde{e}}}{M^e}\left(\frac{1}{M^e} + \frac{1}{M^{\tilde{e}}}\right) \to 0,$$

by (14).

Similarly, we break (54) into an average term and a fluctuation term

$$\sum_{e\in E(v)} (M^e L^e)^{-1} \sum_{z\in x_2^e} \sum_{w\in x_1^e} \int_0^t \Big[(\xi_{s-}(w) - \xi_{s-}(z))\phi_s(z)\, a^{zw}$$

$$(62) \qquad\qquad\qquad + (\xi_{s-}(z) - \xi_{s-}(w))\phi_s(w)\, a^{wz}\Big]ds$$

$$+ \sum_{e\in E(v)} (M^e L^e)^{-1} \sum_{z\in x_2^e} \sum_{w\in x_1^e} \int_0^t (\xi_{s-}(w) - \xi_{s-}(z))\phi_s(z)\left(dP_s^{z,w} - a^{zw}\, ds\right)$$

$$(63) \qquad\qquad\qquad + (\xi_{s-}(z) - \xi_{s-}(w))\phi_s(w)\left(dP_s^{w,z} - a^{wz}\, ds\right).$$

The average term is equal to

$$\sum_{e\in E(v)} (M^e L^e)^{-1} \frac{2C_{ee}^n L^e}{M^e} \sum_{z\in x_2^e} \sum_{w\in x_1^e} \int_0^t (\xi_{s-}(w) - \xi_{s-}(z))\phi_s(z) + (\xi_{s-}(z) - \xi_{s-}(w))\phi_s(w)\, ds$$

$$= \sum_{e\in E(v)} 2C_{ee}^n \int_0^t \left(u_{s-}^n(x_2^e) - u_{s-}^n(x_1^e)\right)\left(\phi_s(x_1^e) - \phi_s(x_2^e)\right)ds$$

$$(64)$$
$$= err_2(t)$$

where $err_2(t)$ is the $o(1)$ error term

$$(65) \qquad err_2(t) = -\sum_{e \in E(v)} \frac{2C_{ee}^n}{L^e} \int_0^t \left(u_{s-}^n(x_2^e) - u_{s-}^n(x_1^e)\right) \nabla_L \phi_s(x_1^e) \, ds.$$

The variance of the fluctuation term is

$$= \sum_{e \in E(v)} \frac{A_{e,e}^n}{(M^e L^e)^2} \sum_{z \in x_2^e} \sum_{w \in x_1^e} \int_0^t \left(\xi_{s-}(w) - \xi_{s-}(z)\right)^2 \left[\phi_s^2(x_2^e) + \phi_s^2(x_1^e)\right] ds$$

$$\leq \sum_{e \in E(v)} \frac{2 A_{e,e}^n}{(L^e)^2} \int_0^t \left[\phi_s^2(x_2^e) + \phi_s^2(x_1^e)\right] ds$$

$$= \sum_{e \in E(v)} \frac{4 C_{ee}^n}{L^e M^e} \int_0^t \left[\phi_s^2(x_2^e) + \phi_s^2(x_1^e)\right] ds \to 0$$

since $1/M^e \to 0$ and C_{ee}^n is of order L^e, as $n \to \infty$, by Remark 5.

6.1.3. Term (42)**. Drift term.** We break (42) into an average term and a fluctuation term

$$(66) \qquad (M^e L^e)^{-1} \sum_{z \in \Lambda_n^e} \sum_{w \sim z} \int_0^t \xi_{s-}(w)(1 - \xi_{s-}(z))\phi_s(z) \, b^{z,w} \, ds$$

$$(67) \qquad + (M^e L^e)^{-1} \sum_{z \in \Lambda_n^e} \sum_{w \sim z} \int_0^t \xi_{s-}(w)(1 - \xi_{s-}(z))\phi_s(z) \, (d\widetilde{P}_s^{z,w} - b^{z,w} \, ds).$$

Recalling the definition of the density u^n in (15), we check that (66) becomes

$$\frac{M^e}{L^e} \sum_{x \in e^n} \int_0^t [u_{s-}^n(x - L^{-1})](1 - u_{s-}^n(x))\phi_s(x) \, B_e^n \, ds$$

$$+ \frac{M^e}{L^e} \sum_{x \in e^n \setminus \{x_1^e, x_2^e\}} \int_0^t [u_{s-}^n(x + L^{-1})](1 - u_{s-}^n(x))\phi_s(x) \, B_e^n \, ds$$

$$+ \frac{M^e}{L^e} \int_0^t u_{s-}^n(x_1^e)(1 - u_{s-}^n(x_2^e))\phi_s(x_2^e) \, \widehat{B}_{e,e}^n \, ds$$

$$(68) \qquad + \sum_{\tilde{e} \in E(v) \setminus e} \frac{M^{\tilde{e}}}{L^e} \int_0^t u_{s-}^n(x_1^{\tilde{e}})(1 - u_{s-}^n(x_1^e))\phi_s(x_1^e) \, \widehat{B}_{e,\tilde{e}}^n \, ds$$

where $\widehat{B}_{e,e}^n = b^{x_2^e, x_1^e}$ and $\widehat{B}_{e,\tilde{e}}^n = b^{x_1^e, x_1^{\tilde{e}}}$. The sum of the first two terms converges to

$$\frac{\beta_e}{2} \int_0^t \int_e 2u_s(x)(1 - u_s(x))\phi_s(x) \, dx \, ds$$

as $n \to \infty$ by Condition (d). After a further summation over e, we see from (17) that the sum of the last two terms tends to

$$(69) \qquad \frac{\hat{\beta}(v)}{4} \int_0^t 2u_s(v)(1 - u_s(v))\phi_s(v) \, ds.$$

The second term (67) is a martingale $E_t^{(2)}(\phi)$ with

$$\langle E^{(2)}(\phi)\rangle_t \le \frac{1}{(M^e L^e)^2} \sum_{z \in \Lambda_n^e} \sum_{w \sim z} \int_0^t \phi_s^2(z)\, b^{z,w}\, ds$$

$$\le \frac{(M^e)^2 B_e^n + M^e M^{\tilde{e}} \widehat{B}_{e,\tilde{e}}^n (|E(v)| - 1)}{(M^e L^e)^2} \int_0^t \langle 1, \phi_s^2 \rangle_e\, ds$$

$$= \Big(\frac{B_e^n}{(L^e)^2} + \frac{M^{\tilde{e}} \widehat{B}_{e,\tilde{e}}^n}{L^e} \frac{(|E(v)| - 1)}{M^e L^e}\Big) \int_0^t \langle 1, \phi_s^2 \rangle_e\, ds \to 0.$$

since $\frac{B_e^n}{(L^e)^2} \to 0$, $\frac{M^{\tilde{e}} \widehat{B}_{e,\tilde{e}}^n}{L^e}$ is bounded and $\frac{(|E(v)|-1)}{M^e L^e} \to 0$.

Limiting martingale problem. Combining our calculations, we can characterize any sub-sequential limit u of u^n. Suppose u is the distributional limit of a sub-sequence of u^n in $D([0,\infty), \mathcal{C}_{[0,1]}(\Gamma))$. Then there is a further sub-sequence that converges almost surely to u which, by our calculation in this section 6.1, satisfies the following: for any $\phi \in C_c^{1,2}([0,\infty) \times \Gamma)$ which satisfies the gluing condition (8) for all t,

$$\int_\Gamma u_t(x)\, \phi_t(x)\, m(dx) - u_0(x)\, \phi_0(x)\, m(dx) - \int_0^t \int_\Gamma u_s(x)\, \partial_s \phi_s(x)\, m(dx)\, ds$$

$$- \int_0^t \int_\Gamma \alpha(x) u_s(x) \Delta \phi_s(x) - \beta(x)\, u_s(x)(1 - u_s(x))\, \phi_s(x)\, m(dx)\, ds$$

(70)
$$- \frac{1}{2} \int_0^t \sum_{v \in V} \hat{\beta}(v)\, u_s(v)(1 - u_s(v))\, \phi_s(v)\, ds$$

is a continuous martingale with quadratic variation

(71)
$$\int_0^t \gamma(x) \int_\Gamma u_s(x)(1 - u_s(x)) \phi_s^2(x)\, m(dx)\, ds,$$

which is the martingale problem formulation of (11); see p. 536-537 in [77]. From this, one can construct on a probability space (see, for instance, [81, Section V20]) a white noise $\dot{W}$ on the Polish space $\Gamma \times [0,\infty)$ such that (103) holds for all $\phi \in C_c(\Gamma) \cap C^2(\hat{\Gamma})$. Hence u solves the (11) weakly. We have shown that, under convergence in distribution in $D([0,\infty), \mathcal{C}_{[0,1]}(\Gamma))$,

(72)   any sub-sequential limit of $\{u^n\}$ solves the SPDE (11) weakly.

REMARK 11.   All calculation in this section *before* taking $n \to \infty$ hold for a more general class of test functions ϕ. Namely $\phi : [0,\infty) \times \Gamma^n \to \mathbb{R}$ is merely defined on Γ^n for the spatial variable, but it is continuously differentiable in t and such that all sums that appeared in the above calculations are well-defined (e.g. when ϕ is bounded and has compact support in Γ). In particular, the gluing condition (40) is not needed in the pre-limit calculations and it is legitimate to apply these calculations to the test function to be defined in (74).

In the next two subsections, we establish tightness of $\{u^n\}$. Weak uniqueness of (11) (Lemma 2) together with (72) then completes the proof of Theorem 1.

6.2. *Green's function representation.* Following [29, Section 4], our proof of tightness begins with the Green's function representation of u^n. This will be obtained in (79)-(80) in this subsection. New terms that do not appear in [29] will be pointed out in our derivations.

Denote $(P_t^n)_{t\geq 0}$ to be the semigroup of X^n, defined by

$$(73) \qquad P_t^n f(x) := \mathbb{E}[f(X_t^n)|X_0^n = x] = \sum_{y\in\Gamma^n} f(y)\, p^n(t,x,y)\, m_n(y)$$

for bounded measurable functions f, where we recall from (28) that $p^n(t,x,y) := \frac{\mathbb{P}(X_t^n=y|X_0^n=x)}{m_n(y)}$.

Observe that $\langle f, g\rangle$ defined before can be written as $\sum_{x\in\Gamma^n} f(x)g(x)m_n(x)$ and that for any $y \in \Gamma^n$,

$$u_t^n(y) = \langle u_t^n, \sum_e \mathbf{1}_{y\in e^n}\, L^e\, \mathbf{1}_y\rangle = \langle u_t^n, \phi_t\rangle,$$

where $\mathbf{1}_y$ is the indicator function and

$$(74) \qquad \phi_s(x) := \phi_s^{t,y}(x) := \begin{cases} p^n(t-s,x,y) & \text{for } s \in [0,t], \\ 0 & \text{otherwise.} \end{cases}$$

Applying the approximate martingale problem (41)-(42) with test function $\phi_s := \phi_s^{t,y}$ in (74) (see Remark 11 for why we can do this) and using the facts that

- $\partial_s \phi_s + \alpha_e \Delta_{L^e} \phi_s = o(1)$, where $o(1)$ is a term which tends to 0 as $n \to \infty$, uniformly for $x \in e^n \smallsetminus \{x_1^e\}$, $e \in E$ and $s \in (0,\infty)$ (see Remark 6); and
- $\langle u_0^n, \phi_0^{t,y}\rangle = P_t^n u_0^n(y)$ for all $y \in \Gamma^n$,

we obtain

$$(75) \qquad u_t^n(y) = P_t^n u_0^n(y) + \sum_e \frac{1}{L^e} \int_0^t u_s^n(x_1^e)\partial_s\phi_s(x_1^e)\,ds$$

$$(76) \qquad + \sum_e \left(Y_t^e(\phi) + Z_t^e(\phi) + E_t^{(1,e)}(\phi) + E_t^{(2,e)}(\phi) \right)$$

$$(77) \qquad + \sum_e \left(T_t^e(\phi) + U_t^e(\phi) + V_t^e(\phi) + E_t^{(3,e)}(\phi) + E_t^{(4,e)}(\phi) \right)$$

for $t \geq 0$ and $y \in \Gamma^n$. Here the terms $Z_t^e(\phi)$, $E_t^{(1,e)}(\phi)$, $E_t^{(2,e)}(\phi)$, $E_t^{(3,e)}(\phi)$ and $E_t^{(4,e)}(\phi)$ are defined in (46), (50), (67), (56) and (63) respectively; $Y_t^e(\phi)$ is defined in (66); the new term

$$(78) \qquad T_t^e(\phi) := \int_0^t \alpha_e\, u_{s-}^n(x_2^e)\, \nabla_L \phi_s(x_1^e)\,ds$$

is obtained from (51); note that the $\alpha_e \Delta_L \phi_s(x)$ term is killed due to our choice of ϕ in (74). New terms $\sum_e U_t^e(\phi)$ and $\sum_e V_t^e(\phi)$ are defined in (55) and (62) respectively.

New technical challenge in proving tightness: The four terms (76) are analogous to terms in (36) of [29], but all five terms in (77) and the $\partial_s \phi_s$ term in (75) are new: they come from boundary terms at vertices of Γ. Treating these new terms requires the uniform estimates for the transition density $p^n(t,x,y)$ of random walks on graph, as well as the careful choice of $C_{x,y}^n$ in Condition (14).

Cancellation and simplification: An important observation is that, by our choice of $C_{x,y}^n$, equations (75)-(77) simplify to

$$(79) \qquad u_t^n(y) = P_t^n u_0^n(y) + \sum_e \left(Y_t^e(\phi) + Z_t^e(\phi) + E_t^{(1,e)}(\phi) + E_t^{(2,e)}(\phi) \right)$$

$$(80) \qquad\qquad + \sum_e \left(E_t^{(3,e)}(\phi) + E_t^{(4,e)}(\phi) \right) + o(1).$$

To see this, note that $U_t^e(\phi) = err_1(t)$ and $V_t^e(\phi) = err_2(t)$ defined in (59) and (65) respectively. That is,

$$U_t^e(\phi) = \int_0^t \sum_{(e,\tilde{e})\text{unordered}} 2C_{\tilde{e},e}^n \left(\phi_s(x_1^e) - \phi_s(x_1^{\tilde{e}}) \right) \left(u_{s-}^n(x_1^{\tilde{e}}) - u_{s-}^n(x_1^e) \right) ds$$

$$V_t^e(\phi) = - \int_0^t \alpha_e \left(u_{s-}^n(x_2^e) - u_{s-}^n(x_1^e) \right) \nabla_L \phi_s(x_1^e) \, ds.$$

On other hand, by (21)

$$\mathcal{L}_n F(x_1^e) \approx \alpha_e (L^e)^2 \left(F(x_2^e) - F(x_1^e) \right) + \sum_{\tilde{e}\in E(v):\,\tilde{e}\neq e} \left(F(x_1^{\tilde{e}}) - F(x_1^e) \right) C_{e,\tilde{e}}^n L^e.$$

where the approximation $\approx$ is quantified in the last sentence in Remark 6. So by our choice of ϕ and Komogorov's equation,

$$(81) \qquad \partial_s \phi_s(x_1^e) \approx -\alpha_e (L^e)^2 \left(\phi_s(x_2^e) - \phi_s(x_1^e) \right) - \sum_{\tilde{e}\in E(v):\,\tilde{e}\neq e} C_{e,\tilde{e}}^n L^e \left(\phi_s(x_1^{\tilde{e}}) - \phi_s(x_1^e) \right).$$

From these it is easy (for example V cancels with the first terms of (81), T cancels with part of U) to check, by using symmetry $C_{e,\tilde{e}}^n = C_{\tilde{e},e}^n$, that we have cancellations

$$(82) \qquad 0 \approx \sum_e \frac{1}{L^e} \int_0^t u_s^n(x_1^e) \partial_s \phi_s(x_1^e) \, ds + T_t^e(\phi) + U_t^e(\phi) + V_t^e(\phi),$$

giving the desired (79) and (80).

6.3. *Tightness of approximate densities.* Our goal of this section is to prove the following $\mathcal{C}$-tightness result.

PROPOSITION 1. *Suppose the assumptions in Theorem 1 hold. Then the sequence $\{u^n\}_{n\geq 1}$ is tight in $D([0,T], \mathcal{C}_{[0,1]}(\Gamma))$ for every $T > 0$. Moreover, any subsequential limit has a continuous version.*

Tightness in $D([0,T], \mathcal{M}(\Gamma))$, where $\mathcal{M}(\Gamma)$ is the space of positive Radon measures on Γ, is much easier to prove than tightness in $D([0,T], \mathcal{C}_{[0,1]}(\Gamma))$ since the former can be reduced to one-dimensional tightnesses (by integrating a fixed function against the measures). However, it is not easy to identify subsequential limits of measures as those whose density solves SPDE (11) weakly.

PROOF OF PROPOSITION 1. The desired $\mathcal{C}$-tightness follows once we can show that (i) the "weak" compact containment condition (condition (a) of Theorem 7.2 in [36, Chapter 3]) holds and (ii) for any $\epsilon > 0$, one has

$$(83) \qquad \lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}\left(\sup_{\substack{t_1 - t_2 < \delta \\ 0 \le t_2 \le t_1 \le T}} \left\| u_{t_1}^n - u_{t_2}^n \right\| > \epsilon \right) = 0.$$

Here and in what follows the norm is the one defined in (16). It is enough to show that (83) holds with u^n replaced by any term in the decomposition given in (79)-(80).

First term in (79). Upon linearly interpolating $P_t^n u_0^n(x)$ in space, triangle inequality and the contraction property of the semigroup $\{P_t^n\}_{t \ge 0}$ gives

$$(84) \qquad \sup_{\substack{t_1 - t_2 < \delta \\ 0 \le t_2 \le t_1 \le T}} \left\| P_{t_1}^n u_0^n - P_{t_2}^n u_0^n \right\| \le 2 \left\| u_0^n - f_0 \right\| + \sup_{\substack{t_1 - t_2 < \delta \\ 0 \le t_2 \le t_1 \le T}} \left\| P_{t_1}^n f_0 - P_{t_2}^n f_0 \right\|,$$

where f_0 is the initial condition for u functions in Theorem 1. So to show that (83) holds with u_t^n replaced by $P_t^n u_0^n$, it suffices to show that (83) holds with u_t^n replaced by $P_t^n f_0$. The latter can be checked by using the uniform Hölder continuity (32).

Second term in (79). For simplicity, we write

$$\widehat{u}_t^n(y) := \sum_e \left(Y_t^e(\phi) + Z_t^e(\phi) + E_t^{(1,e)}(\phi) + E_t^{(2,e)}(\phi) \right).$$

The next moment estimate for space and time increments is similar to [77, Lemma 6] and [29, Lemma 4].

LEMMA 4. *For any $p \ge 2$ and $T \ge 0$ and compact subset K of Γ, there exists a constant $C_{T,p,K} \in (0, \infty)$ such that*

$$(85) \qquad \mathbb{E}|\widehat{u}_{t_1}^n(y_1) - \widehat{u}_{t_2}^n(y_2)|^p \le C_{T,p,K} \left(|t_1 - t_2|^{p/4} + d(y_1, y_2)^{p/2} + \left(\inf_e M^e \right)^{-p} \right)$$

for all $0 \le t_2 \le t_1 \le T$, $y_1, y_2 \in \Gamma^n \cap K$ and $n \ge 1$.

PROOF OF LEMMA 4. The proof of this result requires a number of computations involving estimations of the heat kernel $p^n(t, x, y)$. By (68), $Y_t^e(\phi)$ is equal to

$$(86) \qquad \frac{M^e B_e^n}{L^e} \sum_{x \in e^n} \int_0^t [u_{s-}^n(x - L^{-1})](1 - u_{s-}^n(x)) \phi_s(x) \, ds$$

$$(87) \qquad + \frac{M^e B_e^n}{L^e} \sum_{x \in e^n \setminus \{x_1^e, x_2^e\}} \int_0^t [u_{s-}^n(x + L^{-1})](1 - u_{s-}^n(x)) \phi_s(x) \, ds$$

$$(88) \qquad + \frac{M^e B_{e,e}}{L^e} \int_0^t u_{s-}^n(x_1^e)(1 - u_{s-}^n(x_2^e)) \phi_s(x_2^e) \, ds$$

$$(89) \qquad + \sum_{\tilde{e} \in E(v) \setminus e} \frac{M^{\tilde{e}} B_{e,\tilde{e}}}{L^e} \int_0^t u_{s-}^n(x_1^{\tilde{e}})(1 - u_{s-}^n(x_1^e)) \phi_s(x_1^e) \, ds$$

To estimate $Y_{t_1}^e(\phi^{t_1,y_1}) - Y_{t_2}^e(\phi^{t_2,y_2})$, the key is to observe that all four terms (86)-(89) are sum of integrals of the form

$$I_t(y) := \int_0^t v_{s-}^n(x)\,\phi_s^{t,y}(x)\,ds$$

where $|v_{s-}^n(x)| \le 1$ (for example, $v_{s-}^n(x) = [u_{s-}^n(x - L^{-1})](1 - u_{s-}^n(x))$ for (86)), and that

$$|I_{t_1}(y_1) - I_{t_2}(y_2)| = \Big| \int_{t_2}^{t_1} v_{s-}^n(x)p_{t_1-s}^n(x,y_1)ds + \int_0^{t_2} v_{s-}^n(x)[p_{t_1-s}^n(x,y_1) - p_{t_2-s}^n(x,y_2)]ds \Big|$$

$$\le \int_{t_2}^{t_1} p_{t_1-s}^n(x,y_1)\,ds + \int_0^{t_2} \big| p_{t_1-s}^n(x,y_1) - p_{t_2-s}^n(x,y_2) \big|\,ds.$$

Hence

$$|Y_{t_1}^e(\phi^{t_1,y_1}) - Y_{t_2}^e(\phi^{t_2,y_2})|$$

$$\le \frac{2M^e B_e^n}{L^e} \Big(\sum_{x \in e^n} \int_{t_2}^{t_1} p_{t_1-s}^n(x,y_1)\,ds + \int_0^{t_2} \big| p_{t_1-s}^n(x,y_1) - p_{t_2-s}^n(x,y_2) \big|\,ds \Big)$$

$$+ \frac{M^e B_{e,e}}{L^e} \Big(\int_{t_2}^{t_1} p_{t_1-s}^n(x_2^e,y_1)\,ds + \int_0^{t_2} \big| p_{t_1-s}^n(x_2^e,y_1) - p_{t_2-s}^n(x_2^e,y_2) \big|\,ds \Big)$$

$$+ \sum_{\tilde{e} \in E(v) \setminus e} \frac{M^{\tilde{e}} B_{e,\tilde{e}}}{L^e} \Big(\int_{t_2}^{t_1} p_{t_1-s}^n(x_1^e,y_1)\,ds + \int_0^{t_2} \big| p_{t_1-s}^n(x_1^e,y_1) - p_{t_2-s}^n(x_1^e,y_2) \big|\,ds \Big).$$

The new error terms $E^{(3,e)} + E^{(4,e)}$

$$\sum_e \Big(E_t^{(3,e)}(\phi) + E_t^{(4,e)}(\phi) \Big)$$

$$= \sum_{e \in E(v)} (M^e L^e)^{-1} \sum_{\tilde{e} \in E(v) \setminus e} \sum_{z \in x_1^e} \sum_{w \in x_1^{\tilde{e}}} \int_0^t (\xi_{s-}(w) - \xi_{s-}(z))\phi_s(z)\big(dP_s^{z,w} - A_{e,\tilde{e}}^n ds \big)$$

$$+ \sum_{e \in E(v)} (M^e L^e)^{-1} \sum_{z \in x_2^e} \sum_{w \in x_1^e} \int_0^t (\xi_{s-}(w) - \xi_{s-}(z))\phi_s(z)\big(dP_s^{z,w} - A_{e,e}^n ds \big)$$

$$+ (\xi_{s-}(z) - \xi_{s-}(w))\phi_s(w)\big(dP_s^{w,z} - A_{e,e}^n ds \big)$$

can be treated in the same way, so do terms Z^e, $E^{(1,e)}$ and $E^{(2,e)}$. The proof of Lemma 4 can be completed as in [29, Section 6], using the uniform estimates in Theorem 3. □

Observe that the last term in (85) $\lim_{n\to\infty}(\inf_e M^e)^{-p} = 0$, by Conditions (a)-(b) and the assumption that $\inf_e \alpha_e > 0$. It can then be shown, as in [29, Section 5] that (85) implies (83) holds for $\widehat{u}^n$. This idea is described in the paragraph before Lemma 7 in [77] and page 648 of [59]: we approximate the càdlàg process $\widehat{u}^n$ by a *continuous* process $\tilde{u}$ and invoke a tightness criterion inspired by Kolmogorov's continuity theorem. Finally, the "weak" compact containment condition (condition (a) of Theorem 7.2 in [36, Chapter 3]) follows from the fact $0 \le u^n \le 1$, (84) and Lemma 4 with $t_1 = t_2$.

The proof of Proposition 1 is complete. □

PROOF OF THEOREM 1. The proof of Theorem 1 is complete by Proposition 1, (72) and Lemma 2. □

7. Proof of Theorem 2. The proof follows the footsteps of that of Theorem 1, and many estimates are simpler and very similar. So we provide only the key steps here. As before we begin with the approximate martingale problems for U^n. Using this, we identify any sub-sequential limit and obtain a Green's function representation for U^n. The proof is complete by establishing tightness, based on the Green's function representation.

Before proving convergence, we first argue that U^n, after the described interpolation, is an element in $\mathcal{C}([0,\infty), \mathcal{C}_{[0,1]}(\Gamma))$. Indeed, if the initial condition $U^n(0)$ is bounded between 0 and 1, then there exists a weak solution $\{U^n_x\}_{x \in \Gamma^n}$ to (22)-(23) with initial condition $U^n(0)$ and takes values in [0,1] for all $t \geq 0$. This follows from an approximation argument as for the SPDE (10): one first approximates $\sqrt{x(1-x)}$ by Lipschitz functions $\{a_k(x)\}_k$ taking values in $[0,1]$ such that $a_k(0) = a_k(1) = 0$ and constructs solutions $\{U^{n,k}\}_k$ using standard theory, then applies the comparison principle and passing to a sub-sequential limit $k \to \infty$. See, for instance, [75, Section 2.1]. Furthermore, weak uniqueness of (22)-(23) follows from duality argument as in [26]. Therefore $U^n(t) \in \mathcal{C}_{[0,1]}(\Gamma)$ for all $t \geq 0$. Finally, since $t \mapsto U^n_x(t)$ is continuous for each $x \in \Gamma^n$, and a compact subset of Γ contains only finitely many $x \in \Gamma^n$, we have $U^n \in \mathcal{C}([0,\infty), \mathcal{C}_{[0,1]}(\Gamma))$ based on the norm of $\mathcal{C}_{[0,1]}(\Gamma)$ in (16).

7.1. *Approximate martingale problem.* This section is similar to Section 6.1. Recall the inner products (39) and let $\phi : [0,\infty) \times \Gamma \to \mathbb{R}$ be a continuous function that is continuously differentiable in t, twice continuously differentiable and has compact support in $x \in \Gamma$ and satisfies the gluing condition (40).

As before we focus on a single vertex $v \in V$, and for each $e \in E(v)$ we enumerate the set e^n as $(x^e_1, x^e_2, x^e_3, \cdots)$ along the direction of e away from v; see **Figure 2**. From our construction (22)-(23), for $t \geq 0$ and $e \in E(v)$,

$$\langle U^n(t), \phi_t \rangle_e - \langle U^N(0), \phi_0 \rangle_e - \int_0^t \langle U^n(s), \partial_s \phi_s \rangle_e \, ds$$

(90)

$$= \frac{1}{L^e} \sum_{x \in e^n \setminus \{x^e_1\}} \left\{ \int_0^t \phi_s(x) \Big[\mathcal{L}_n U_x + \beta_e U_x (1 - U_x) \Big] ds + \int_0^t \phi_s(x) \sqrt{\gamma_e L^e U_x (1 - U_x)} dB_x(s) \right\}$$

(91)

$$+ \frac{1}{L^e} \int_0^t \phi^e_1 \left[\mathcal{L}_n U^e_1 + L^e \frac{\hat{\beta}(v)}{deg(v)} U^e_1 (1 - U^e_1) \right] ds,$$

where we used the abbreviations $U^e_k := U_{x^e_k}$ and $\phi^e_k := \phi_t(x^e_k)$.

Laplacian and gluing condition. We first consider the terms involving the operator $\mathcal{L}_n$ in (90)-(91). The careful construction of the generator $\mathcal{L}_n$ of the CTRW X^n in Remark 4, including condition (14) for the conductances $\{C_{e,\tilde{e}}\}$, as well as the gluing condition for ϕ, are crucial for the lemma below.

LEMMA 5. *Suppose $\{U^n\}$ converge almost surely in $C([0,\infty), C_{[0,1](\Gamma)})$ along a sub-sequence to a limit U^∞. Then along that sub-sequence, almost surely,*

$$\sum_{e \in E(v)} \frac{1}{L^e} \sum_{x \in e^n} \phi_s(x) \mathcal{L}_n U^n_x(s) \to \sum_{e \in E(v)} \alpha_e \int_e U^\infty(s,x) \Delta\phi(s,x) \, m(dx).$$

uniformly for all $t \in [0,\infty)$ and $v \in V$.

Proof. By Condition (14),

$$(92) \qquad \sup_{s \geq 0} \frac{1}{L^e} \sum_{x \in e^n \setminus \{x_1^e\}} |\mathcal{L}_n U_x^n(s) - \alpha_e \Delta_L U_x^n(s)| \to 0$$

uniformly for all edges and vertices, where Δ_L is the discrete Laplacian defined in (48); see Remark 6. Applying summation by parts twice and omitting the superscript e for simplicity,

$$\sum_{k=2}^{N} \phi_k (\nabla U_k - \nabla U_{k-1}) = \phi_N \nabla U_N - \phi_2 \nabla U_1 - \nabla \phi_{N-1} U_N + \nabla \phi_2 U_2 + \sum_{j=3}^{N-1} U_j (\nabla \phi_j - \nabla \phi_{j-1}),$$

where $\nabla U_k = U_{k+1} - U_k$. Since ϕ_s has compact support on Γ,

$$(93) \qquad \sum_{k \geq 2} \phi_k^e \Delta_L U_k^e = -(L^e)^2 \left[\phi_2^e (U_2^e - U_1^e) + (\phi_3^e - \phi_2^e) U_2^e \right] + \sum_{j \geq 3} U_j^e \Delta_L \phi_j^e$$

From the above two displays and also (21), we see that by adding up the terms in (90)-(91) that involve the operator $\mathcal{L}_n$,

$$\frac{1}{L^e} \sum_{x \in e^n} \phi_s(x) \mathcal{L}_n U_x$$

$$= \frac{\alpha_e}{L^e} \left\{ -(L^e)^2 \left[\phi_2^e (U_2^e - U_1^e) + (\phi_3^e - \phi_2^e) U_2^e \right] + \sum_{j \geq 3} U_j^e \Delta_L \phi_j^e \right\}$$

$$+ \frac{1}{L^e} \phi_1^e \left\{ \alpha_e (L^e)^2 (U_2^e - U_1^e) + \sum_{\tilde{e} \in E(v): \tilde{e} \neq e} (U_1^{\tilde{e}} - U_1^e) C_{e,\tilde{e}} L^e \right\} + \kappa_n \|\phi_s\|_\infty$$

$$(94) \qquad = \alpha_e L^e \left\{ -(\phi_3^e - \phi_2^e) U_2^e - (\phi_2^e - \phi_1^e)(U_2^e - U_1^e) \right\}$$

$$(95) \qquad + \phi_1^e \sum_{\tilde{e} \in E(v): \tilde{e} \neq e} (U_1^{\tilde{e}} - U_1^e) C_{e,\tilde{e}} + \frac{\alpha_e}{L^e} \sum_{j \geq 3} U_j^e \Delta_L \phi_j^e + \kappa_n \|\phi_s\|_\infty,$$

where, by (92), κ_n is a constant that tends to 0 uniformly for all $e \in E(v)$, $v \in V$ and $s \geq 0$.

We now sum over $e \in E(v)$. Suppose $\{U^n\}$ converge almost surely along a sub-sequence. Then by the *gluing condition* (40), the term (94) (after summing over $e \in E(v)$) will tend to zero along that sub-sequence. The first term of (95) also tends to zero (after summing over $e \in E(v)$), by the same argument for (58)-(59) based on the symmetry of the conductances $\{C_{e,\tilde{e}}\}$ and the uniform continuity of ϕ. The second term of (95), converges to $\alpha_e \int_e U^\infty(s, x) \Delta \phi(s, x) m(dx)$. The proof is complete. □

Limiting martingale problem. Convergences of the other terms in (90)- (91) follow directly from the continuity of ϕ. For example, for the term involving $\hat{\beta}(v)$,

$$\sum_{e \in E(v)} \frac{1}{L^e} \int_0^t \phi_1^e L^e \frac{\hat{\beta}(v)}{deg(v)} U_1^e (1 - U_1^e) \, ds \to \hat{\beta}(v) \int_0^t \phi_s(v) U_v^\infty(s) (1 - U_v^\infty(s)) \, ds,$$

which explains the term $deg(v)$ in the denominator. Combining our calculations as in (70)-(71), we have shown that, under convergence in distribution in $\mathcal{C}([0, \infty), \mathcal{C}_{[0,1]}(\Gamma))$,

$$(96) \qquad \text{any sub-sequential limit of } \{U^n\} \text{ solves the SPDE (11) weakly.}$$

7.2. *Green's function representation.* This section is similar to Section 6.2. We shall obtain a Green's function representation of U^n that is analogous to (79)-(80). Such representation will be useful for proving tightness.

Applying the approximate martingale problem (90)-(91) with the same test function $\phi_s := \phi_s^{t,y}$ in (74), namely

$$\phi_s(x) := \phi_s^{t,y}(x) := p^n(t-s, x, y) \quad \text{for } s \in [0, t]$$

and zero otherwise, we obtain that $\partial_s \phi_s + \mathcal{L}_n \phi_s = 0$ and hence for $t \geq 0$ and $y \in \Gamma^n$,

$$(97) \quad U_y^n(t) - \left(P_t^n u_0\right)(y) + \int_0^t \langle U^n(s), \mathcal{L}_n \phi_s \rangle - \langle \mathcal{L}_n U^n(s), \phi_s \rangle \, ds$$

$$(98) \quad = \sum_e \frac{1}{L^e} \sum_{x \in e^n \setminus \{x_1^e\}} \left\{ \int_0^t \phi_s(x) \beta_e U_x(1 - U_x) \, ds + \int_0^t \phi_s(x) \sqrt{\gamma_e L^e U_x(1 - U_x)} \, dB_x(s) \right\}$$

$$(99) \quad + \sum_e \frac{1}{L^e} \int_0^t \phi_1^e L^e \frac{\hat{\beta}(v)}{deg(v)} U_1^e(1 - U_1^e) \, ds.$$

7.3. *Tightness.* This section is similar to Section 6.3. Our goal is to prove the following tightness result.

PROPOSITION 2. *Suppose the assumptions in Theorem 2 hold. Then the sequence $\{U^n\}_{n \geq 1}$ is tight in $C([0, T], \mathcal{C}_{[0,1]}(\Gamma))$ for every $T > 0$.*

PROOF. The desired tightness follows once we can show that (i) the "weak" compact containment condition (condition (a) of Theorem 7.2 in [36, Chapter 3]) holds and (ii) for any $\epsilon > 0$, one has

$$(100) \quad \lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P} \left(\sup_{\substack{t_1 - t_2 < \delta \\ 0 \leq t_2 \leq t_1 \leq T}} \left\| U^n(t_1) - U^n(t_2) \right\| > \epsilon \right) = 0.$$

To show (100), it is enough to show that (100) holds with U^n replaced by any other terms in the decomposition given in (97)-(99).

Note that the summation-by-parts calculation (93) still holds if ϕ and U are interchanged. From this, the fact that $0 \leq U^n \leq 1$ and the Gaussian upper bound (29)-(30) for ϕ_s, we obtain that the integral in (97)

$$(101) \quad \left| \int_0^t \langle U^n(s), \mathcal{L}_n \phi_s \rangle - \langle \mathcal{L}_n U^n(s), \phi_s \rangle \, ds \right| \leq \kappa_n \sqrt{t}$$

for $t \geq 0$, where κ_n is a constant that tends to 0 uniformly for all $e \in E(v)$, $v \in V$ and $t \geq 0$.

Furthermore, as in Lemma 4, if we let $\widehat{U}_y^n(t) = (98) + (99)$, then we have the following moment estimate: For any $p \geq 2$ and $T \geq 0$ and compact subset K of Γ, there exists a constant $C_{T,p,K} \in (0, \infty)$ such that

$$(102) \quad \mathbb{E} \left| \widehat{U}_{y_1}^n(t_1) - \widehat{U}_{y_2}^n(t_2) \right|^p \leq C_{T,p,K} \left(|t_1 - t_2|^{p/4} + d(y_1, y_2)^{p/2} \right)$$

for all $0 \leq t_2 \leq t_1 \leq T$, $y_1, y_2 \in \Gamma^n \cap K$ and $n \geq 1$.

The "weak" compact containment condition (condition (a) of Theorem 7.2 in [36, Chapter 3]) follows from the fact $0 \leq U^n \leq 1$ and (102) with $t_1 = t_2$. With (101), (102) and the uniform heat kernel estimates in Theorem 3, the proof can be completed as in Section 6.2.

The proof of Proposition 2 is complete.   □

PROOF OF THEOREM 2. The proof of Theorem 2 is complete by Proposition 2, (72) and Lemma 2. □

8. Appendix. In this appendix, we specify the notions of weak solution and mild solution to SPDE on graphs, give a proof of Lemma 1, point out some generalizations of our results and some open problems.

8.1. *Solutions to SPDE on graphs.* For completeness we give the precise definition for the notions of weak solutions and of mild solutions for SPDE (10):

$$
\begin{cases}
\partial_t u = \mathcal{L}\, u + \beta\, u(1-u) + \sqrt{\gamma\, u(1-u)}\dot{W} & \text{on } \overset{\circ}{\Gamma} \\
\nabla_{out} u \cdot [\alpha] = -\hat{\beta}\, u(1-u) & \text{on } V.
\end{cases}
$$

These definitions are analogous to the usual ones (see [92]) but have extra boundary conditions.

DEFINITION 1. *Let $\dot{W}$ be the space-time white noise on Γ endowed with the product of Lebesque measures $\mu := m(dx) \otimes dt$. That is, $\dot{W} := \{\dot{W}(A)\}_{A \in \mathfrak{B}(\Gamma \times [0,\infty))}$ are centered Gaussian random variables with covariance $\mathbb{E}[\dot{W}(A)\dot{W}(B)] = \mu(A \cap B)$. The white noise process $(W_t)_{t \geq 0}$ is defined by $W_t(U) := \dot{W}(U \times [0,t])$ where $U \in \mathfrak{B}(\Gamma)$. Denote by $\mathcal{F}_t$ to the sigma-algebra generated by $\{W_s(U) : 0 \leq s \leq t, U \in \mathfrak{B}(\Gamma)\}$ and call $(\mathcal{F}_t)_{t \geq 0}$ the filtration generated by $\dot{W}$.*

It can be checked as in [92] that $\{W_t(U), \mathcal{F}_t\}_{t \geq 0, U \in \mathfrak{B}(\Gamma)}$ is an orthogonal (hence worthy) martingale measure, so that we have a well-defined notion of stochastic integral with respect to W for a class of integrands which contains the collection of all predictable functions f such that

$$
\mathbb{E} \int_{(0,T]} \Big(\int_\Gamma |f(x,t)|\, m(dx) \Big)^2 dt < \infty
$$

for all $T > 0$. By the Gaussian upper bound in (36), the stochastic integrals that appear throughout this paper, including (104) below, are well-defined.

The notions of weak solution and mild solution are given below. Let $\mathcal{B}_{[0,1]}(\Gamma)$ be the space of Borel measurable functions on Γ taking values in the interval $[0,1]$. Note that we do not need an absolute sign inside the square root $\sqrt{\gamma u(1-u)}$.

DEFINITION 2. *A process $u = (u_t)_{t \geq 0}$ taking values in $\mathcal{B}_{[0,1]}(\Gamma)$ is a **weak solution** to SPDE (10) with initial condition u_0 if there is a space-time white noise $\dot{W}$ on $\Gamma \times [0,\infty)$ such that (i) u is adapted to the filtration generated by $\dot{W}$ and (ii) for any $\phi \in \mathcal{C}_c(\Gamma) \cap \mathcal{C}^2(\overset{\circ}{\Gamma})$*

satisfying the gluing condition (8), *we have*

$$\int_\Gamma u_t(x)\phi(x)\,\nu(dx) = \int_\Gamma u_0(x)\phi(x)\,\nu(dx) + \frac{1}{2}\int_0^t \int_\Gamma \nabla\Big(\ell\,\alpha\,\nabla\phi\Big)(x)\,u_s(x)\,m(dx)\,ds$$

$$+ \int_0^t \int_\Gamma \beta(x)\,u_s(x)(1-u_s(x))\,\phi(x)\,\nu(dx)\,ds$$

$$+ \int_{[0,t]\times\Gamma} \phi(x)\sqrt{\gamma(x)\,u_s(x)\big(1-u_s(x)\big)}\,\ell(x)\,dW(s,x)$$

(103) $$+ \frac{1}{2}\int_0^t \sum_{v\in V} \phi(v)\hat\beta(v)\,u_s(v)(1-u_s(v))\,\ell(v)\,ds$$

for all $t \geq 0$, *almost surely, where* $\nu(dx) = \ell(x)m(dx)$.

DEFINITION 3. *A process* $u = (u_t)_{t\geq 0}$ *taking values in* $\mathcal{B}_{[0,1]}(\Gamma)$ *is a* **mild solution** *to SPDE* (10) *with initial condition* u_0 *if there is a space-time white noise* $\dot W$ *on* $\Gamma \times [0,\infty)$ *such that (i)* u *is adapted to the filtration generated by* $\dot W$ *and (ii)* u *solves the integral equation*

$$u_t(x) = P_t u_0(x) + \int_0^t P_{t-s}\big(\beta\,u_s(1-u_s)\big)(x)\,ds$$

$$+ \int_{[0,t]\times\Gamma} p(t-s,x,y)\,\ell(y)\sqrt{\gamma(y)\,u_s(y)\big(1-u_s(y)\big)}\,dW(s,y)$$

(104) $$+ \frac{1}{2}\int_0^t \sum_{v\in V} p(t-s,x,v)\,\ell(v)\,\hat\beta(v)\,u_s(v)(1-u_s(v))\,ds,$$

where $p(t,x,y)$ *is defined in* (6) *and* $P_t f(x) = \mathbb{E}_x f(X_t) = \int_\Gamma f(y)p(t,x,y)\nu(dy)$.

A weak solution may fail to be a mild solution in general (see, e.g., [15]). However, we have the following.

LEMMA 6. *Suppose Assumptions* 1, 2 *and* 3 *hold and* $u_0 \in \mathcal{C}_{[0,1]}(\Gamma)$. *An adapted process* u *with* $u_t \in \mathcal{B}_{[0,1]}(\Gamma)$ *for all* $t \geq 0$ *is a weak solution of* (10) *if and only if it is a mild solution.*

PROOF. A proof follows from that of Shiga's result [85, Theorem 2.1], thanks to our heat kernel estimates in Theorem 5 and the boundedness of u. We provide a sketch of proof.

Suppose u is a weak solution of (10). Then (103) holds for the test function $\phi(x) = p(t,y,x)$ for $x \in \Gamma$, where $(t,y) \in (0,\infty) \times \Gamma$ is fixed. This is because (i) $\phi \in Dom_{L^2}(\mathcal{L})$ and hence ϕ satisfies the gluing condition, and (ii) $Dom_{L^2}(\mathcal{L}) \subset Dom(\mathcal{E}) = W^{1,2}(\Gamma,\nu) \cap \mathcal{C}^1(\Gamma)$ and so ϕ is the limit, under the Sobolev norm $\|\cdot\|_{W^{1,2}(\nu)}$, of a sequence in $\mathcal{C}_c(\Gamma) \cap \mathcal{C}^2(\mathring\Gamma)$ that satisfy the gluing condition. Furthermore, the time-derivatives of $p(t,x,y)$ also possess Gaussian bounds and Hölder continuity; see [51, Theorem 2.32]. This allows us to extend (103) to time-dependent test functions $\phi : [0,T] \times \Gamma \to \mathbb{R}$, where $\phi(t,x) = p(T-t,y,x)$ and T,y are fixed. This gives

$$P_{T-t}u_t(x) = P_T u_0(x) + \int_0^t P_{T-s}\big(\beta\,u_s(1-u_s)\big)(x)\,ds$$

$$+ \int_{[0,t]\times\Gamma} p(T-s,x,y)\,\ell(y)\sqrt{\gamma(y)\,u_s(y)\big(1-u_s(y)\big)}\,dW(s,y)$$

(105) $$+ \frac{1}{2}\int_0^t \sum_{v\in V} p(T-s,x,v)\,\ell(v)\,\hat\beta(v)\,u_s(v)(1-u_s(v))\,ds$$

and hence (104) by letting $t \to T$. Therefore u is also a mild solution.

The converse is straightforward. Suppose u is a mild solution. We insert (104) into $\int_0^t \int_\Gamma \nabla\big(\ell\,\alpha\,\nabla\phi\big)(x)\,u_s(x)\,m(dx)\,ds$ and apply a stochastic Fubini theorem to verify (103). $\qquad\square$

For existence and uniqueness of solutions, results in [9] cannot be directly applied since we have a non-Lipschitz coefficient $\sqrt{\gamma\,u(1-u)}$. Nonetheless, existence of a mild solution taking values in $\mathcal{B}_{[0,1]}(\Gamma)$ follows from an approximation and tightness argument. See, for instance, [75, Section 2.1]. Furthermore, weak uniqueness for (10) holds by Lemma 2.

PROOF OF LEMMA 1 (DUALITY). Recall that Assumptions 1, 2 and 3 in force. Recall from (5) and (6) that $\nu(dy) = \ell(y)m(dy)$ and $\mathbb{P}_x(X_t \in dy) = p(t,x,y)\,\nu(dy)$.

To simplify notation, we write $p^\epsilon(x,y) = p(\epsilon,x,y)$ for the transition density of the symmetric diffusion X and define $z := 1 - u$ and $\bar{z}_t(x) := \int_\Gamma z_t(y)p^\epsilon(x,y)\nu(dy)$.

Using the weak formulation (103) with test function $\phi^{\epsilon,x}(y) = p^\epsilon(x,y)$ (we can do so as explained in the proof of Lemma 6), we obtain

$$
\begin{aligned}
\bar{z}_t(x) - \bar{z}_0(x) = {} & \frac{1}{2}\int_0^t \int_\Gamma \nabla_y\big(\ell(y)\,\alpha(y)\,\nabla_y p^\epsilon(x,y)\big)\,z_s(y)\,m(dy)\,ds \\
& - \int_0^t \int_\Gamma \beta(y)\,z_s(y)(1 - z_s(y))\,p^\epsilon(x,y)\,\ell(y)\,m(dy)\,ds \\
& + \int_0^t \int_\Gamma p^\epsilon(x,y)\sqrt{\gamma(y)\,z_s(y)\big(1 - z_s(y)\big)}\,\ell(y)\,dW(s,y) \\
& + \frac{1}{2}\int_0^t \sum_{v \in V} p^\epsilon(x,v)\,\hat{\beta}(v)z_s(v)\big(1 - z_s(v)\big)\,\ell(v)\,ds.
\end{aligned}
$$

Fix $x_1, \ldots x_n$ and let $\mathcal{L}_{\bar{z}}$ be the generator of the process $(\bar{z}_t(x_1), \ldots, \bar{z}_t(x_n))_{t \geq 0}$. Using the Itô's formula [41, Lemma 2.3] (each process $(\bar{z}_t(x))_{t \geq 0}$ is a semi-martingale so this is legitimate), we see that as in Section 8.1 of [29],

$$
\begin{aligned}
\operatorname{drift}\left(\mathcal{L}_{\bar{z}}\prod_i \bar{z}_t(x_i)\right) = {} & \frac{1}{2}\sum_{i=1}^n \prod_{j \neq i}\bar{z}_t(x_j)\int_\Gamma \nabla_y\big(\alpha(y)\,\nabla_y p^\epsilon(x_i,y)\big)\,z_t(y)\,m(dy) \\
& + \sum_{i=1}^n \prod_{j \neq i}\bar{z}_t(x_j)\int_\Gamma \beta(y)[z_t^2(y) - z_t(y)]p^\epsilon(x_i,y)\,\ell(y)\,m(dy) \\
& + \sum_{i=1}^{n-1}\sum_{j=i+1}^n \prod_{k \neq i,j}\bar{z}_t(x_k)\int_\Gamma \gamma(y)[z_t(y)(1 - z_t(y))]p^\epsilon(x_i,y)p^\epsilon(x_j,y)\ell(y)m(dy) \\
& + \frac{1}{2}\sum_{i=1}^n \sum_{v \in V}\prod_{j \neq i}\bar{z}_t(x_j)\,\hat{\beta}(v)\big(z_t^2(v) - z_t(v)\big)p^\epsilon(x_i,v)\,\ell(v).
\end{aligned}
$$
(106)

The dual process is a system of branching coalescing particles performing $\mathcal{E}$-diffusions on Γ. During their lifetime all particles independently give birth (here means the particle splits into two) at rate $\beta(x)$ at the interior $x \in \mathring{\Gamma}$ and give birth at rate $\hat{\beta}(v)L_t^{i,v}$ at vertex $v \in V$. In addition, for $i < j$, particle j is killed by particle i at rate $\frac{\gamma(x_i)}{\ell(x_i)}L_t^{i,j}$ where $L_t^{i,j}$ denotes the

local time of the process $x_j - x_i$ at 0. Writing $\mathcal{L}_x$ for the generator of this particle system,

$$
\begin{aligned}
\text{drift}\left(\mathcal{L}_x \prod_i \bar{z}_t(x_i)\right) &= \frac{1}{2} \sum_{i=1}^n \prod_{j \neq i} \bar{z}_t(x_j) \cdot \nabla\left(\alpha \nabla \bar{z}_t\right)(x_i) \\
&+ \sum_{i=1}^n \prod_{j \neq i} \bar{z}_t(x_j) \cdot \beta(x_i)[\bar{z}_t^2(x_i) - \bar{z}_t(x_i)] \\
&+ \sum_{i=1}^{n-1} \sum_{j=i+1}^n \prod_{k \neq i,j} \bar{z}_t(x_k) \cdot \left[\frac{\gamma(x_i)}{\ell(x_i)} \bar{z}_t(x_i)(1 - \bar{z}_t(x_j))\right] \delta_{\{x_j = x_i\}} \\
(107) \qquad &+ \frac{1}{2} \sum_{i=1}^n \sum_{v \in V} \prod_{j \neq i} \bar{z}_t(x_j) \cdot \hat{\beta}(v)[\bar{z}_t^2(v) - \bar{z}_t(v)]\delta_{\{v = x_i\}}
\end{aligned}
$$

in which we used the formal notation $dL_t^{i,j} = \delta_{\{x_j(t) = x_i(t)\}}\, dt$ and $dL_t^{i,v} = \delta_{\{x_i(t) = v\}}\, dt$.

We continue to follow [29, Section 8.1] to assert that

$$
(108) \qquad \mathbb{E}F(\bar{z}_t, x(0)) - \mathbb{E}F(\bar{z}_0, x(t)) = \mathbb{E} \int_0^t G(\bar{z}_{t-s}, x(s)) - H(\bar{z}_{t-s}, x(s))\, ds.
$$

with $G(\bar{z}, x) = \mathcal{L}_{\bar{z}} F(\bar{z}, x)$ and $H(\bar{z}, x) = \mathcal{L}_x F(\bar{z}, x)$, where

$$
F(\bar{z}, x) = \prod_{i=1}^n \bar{z}(x_i) \quad \text{if } x = (x_1, \cdots, x_n).
$$

To prove the desired duality formula (27), i.e.

$$
(109) \qquad \mathbb{E} \prod_{i=1}^{n(0)} z_t(x_i(0)) = \mathbb{E} \prod_{i=1}^{n(t)} z_0(x_i(t)), \quad t \geq 0,
$$

it remains to argue that the RHS of (108) tends to zero as $\epsilon \to 0$.

The first term in (106) agrees with that of (107) on the RHS of (108). The second and the third terms of (106) and (107) can be teated in the same way as in [29, Section 8.1], thanks to Theorem 5 and Assumption 3. The reason we need to divide $\gamma(x_i)$ by $\ell(x_i)$ in the third term of (107) is as follows: there are two p^ϵ factors but only one ℓ in the third term of (106). This explains why the coalescence rate should be $\frac{\gamma(x_i)}{\ell(x_i)} L_t^{i,j}$ rather than $\gamma(x_i) L_t^{i,j}$ as in [29, Section 8.1].

Finally, we consider the new contribution to the RHS of (108) from the last term of (106) and (107), coming from the branchings in the vertex set V.

The contribution to (108) from the last term of (107) is

$$
(110) \qquad \mathbb{E} \int_0^t \sum_{i=1}^{n(s)} \sum_{v \in V} \prod_{j \neq i} \bar{z}_{t-s}(x_j(s)) \cdot \hat{\beta}(v)[\bar{z}_{t-s}^2(v) - \bar{z}_{t-s}(v)]\, dL_s^{i,v}
$$

which converges as $\epsilon \to 0$ (first by dominated convergence for fixed $v \in V$ and then by monotone convergence for $\sum_{v \in V}$) to

$$
(111) \qquad \mathbb{E} \int_0^t \sum_{i=1}^{n(s)} \sum_{v \in V} Y_{s,t}^{i,v}\, dL_s^{i,v},
$$

38

where

$$Y_{s,t}^{i,v} = \prod_{j \neq i} z_{t-s}(x_j(s)) \cdot \hat{\beta}(v)[z_{t-s}^2(v) - z_{t-s}(v)]$$

is non-negative and bounded from above uniformly for $1 \leq i \leq n(s)$, $s \in [0, t]$ and $v \in V$, according to Assumption 3. The contribution to (108) from the last term of (106) is

$$(112) \qquad \frac{1}{2}\mathbb{E} \int_0^t \sum_{i=1}^{n(s)} \sum_{v \in V} Y_{s,t}^{i,v} \, p^\epsilon(x_i(s), v) \, \ell(v) \, ds$$

which also converges to (111) as $\epsilon \to 0$, first by dominated convergence for fixed $v \in V$ and then by monotone convergence for $\sum_{v \in V}$ (details as in pages 1724–1725 of [3]). The proof of (109) is complete. $\qquad \square$

8.2. *Generalizations.* Assumptions 1 and 2 can both be significantly relaxed. The Dirichlet form method enables one to construct more general processes, such as singular diffusions on graphs as in [58, 83], and diffusions on $\mathbb{R}$-trees [18, 19, 20, 1] and fractals [4]. These two assumptions should be compared with conditions on the boundary of a domain and on the coefficients respectively in the construction of reflected diffusions [13]. The function ℓ contributes to a drift for the diffusion which we ignored for simplicity, i.e., we assumed $\ell = 1$. To incorporate such as drift in discrete approximations, see [39, Section 6].

Furthermore, with the techniques developed in this paper, one can immediately generalize the coupled SPDE in [29, Theorem 4] to the graph setting. This enables us to investigate the role of space in shaping coexistence and competition outcomes of interacting species. In ongoing work, we plan to apply such generalizations to study the interactions of virus and sub-virus particles during co-infection in a population of susceptible cells.

8.3. *Open problems.* The study of SPDE on graphs leads to an explosion of interesting open questions. We mention only three of the important ones below.

PROBLEM 1 (Speed of FKPP on Γ). Wavefront propagation and asymptotic speed for deterministic FKPP on a class of regular trees has recently been obtained using large deviation techniques in [40]. Can we define and obtain the *asymptotic speed* of wavefront propagation for the weak solution of the stochastic PDE (11) in terms of properties of Γ ? For instance, in terms of k (and α, β, γ) when Γ is the infinite k-regular tree with unit branch length? The duality formula in Lemma 1 can be very useful for this study (see [6, 26, 74] for the case $\Gamma = \mathbb{R}$).

PROBLEM 2 (Contact process). Techniques developed here enables one to analyse other SPDE on graphs such as scaling limits of branching random walks and those of *contact processes*

$$(113) \qquad \partial_t u = \alpha \, \Delta u + \beta \, u - \delta \, u^2 + \sqrt{\gamma \, u} \, \dot{W} \quad \text{on } \overset{\circ}{\Gamma}.$$

A fundamental and challenging question in epidemiology is to estimate (and show existence of) the *threshold infection rate* $\beta_c \in (0, \infty)$ in terms of geometric properties of Γ, such that when $\beta > \beta_c$ the infection sustain with positive probability, and when $0 \leq \beta < \beta_c$ the infection dies out with probability one (assuming β is a constant function and $\hat{\beta} = 0$ on V for simplicity). See Muller and Tribe [76, Theorem 1] for such a threshold when $\Gamma = \mathbb{R}$.

PROBLEM 3 (Other SPDE and mixture models). In this paper, we obtain the term $\hat{\beta}u(1-u)$ in (11) because we specify biased voter rule not just on edges but also at the vertex set. Some other polynomial terms can arise if we specify other microscopic rules (such as contact process) near the vertices. This is left open for exploration. For example, the interacting particle system can be a mixture of the biased voter model and the contact process. See Lanchier and Neuhauser [66, 67] who introduced one such mixture to investigate how the interactions in spatially explicit host-symbiont systems are shaping plant community structure.

REFERENCES

[1] S. Athreya, M. Eckhoff, and A. Winter. Brownian motion on $\mathbb{R}$-trees. *Trans. Amer. Math. Soc.*, 365(6):3115–3150, 2013.

[2] S. Athreya, W. Löhr, and A. Winter. Invariance principle for variable speed random walks on trees. *Ann. Probab.*, 45(2):625–667, 2017.

[3] S. Athreya and R. Tribe. Uniqueness for a class of one-dimensional stochastic PDEs using moment duality. *Ann. Probab.*, 28(4):1711–1734, 2000.

[4] M. T. Barlow and E. A. Perkins. Brownian motion on the Sierpiński gasket. *Probab. Theory Related Fields*, 79(4):543–623, 1988.

[5] D. Blount. Fourier analysis applied to SPDEs. *Stochastic Process. Appl.*, 62(2):223–242, 1996.

[6] E. Brunet, B. Derrida, A. Mueller, and S. Munier. Phenomenological theory giving the full statistics of the position of fluctuating pulled fronts. *Phys. E*, 73(5):056126, 2006.

[7] F. Caravenna, R. Sun, and N. Zygouras. Universality in marginally relevant disordered systems. *arXiv preprint arXiv:1510.06287*, 2015.

[8] E. A. Carlen, S. Kusuoka, and D. W. Stroock. Upper bounds for symmetric markov transition functions. In *Annales de l'IHP Probabilités et statistiques*, volume 23, pages 245–287, 1987.

[9] S. Cerrai and M. Freidlin. SPDEs on narrow domains and on graphs: an asymptotic approach. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(2):865–899, 2017.

[10] S. Cerrai and M. Freidlin. Fast flow asymptotics for stochastic incompressible viscous fluids in $\mathbb{R}^2$ and spdes on graphs. *Probability Theory and Related Fields*, 173(1-2):491–535, 2019.

[11] Z.-Q. Chen and W.-T. L. Fan. Fluctuation limit for interacting diffusions with partial annihilations through membranes. *J. Stat. Phys.*, 164(4):890–936, 2016.

[12] Z.-Q. Chen and W.-T. L. Fan. Hydrodynamic limits and propagation of chaos for interacting random walks in domains. *Ann. Appl. Probab.*, 27(3):1299–1371, 2017.

[13] Z.-Q. Chen and M. Fukushima. *Symmetric Markov processes, time change, and boundary theory*, volume 35 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2012.

[14] E. A. Codling, M. J. Plank, and S. Benhamou. Random walk models in biology. *Journal of the Royal Society Interface*, 5(25):813–834, 2008.

[15] D. Conus and D. Khoshnevisan. Weak nonmild solutions to some SPDEs. *Illinois J. Math.*, 54(4):1329–1341 (2012), 2010.

[16] J. T. Cox, R. Durrett, and E. A. Perkins. Rescaled voter models converge to super-Brownian motion. *Ann. Probab.*, 28(1):185–234, 2000.

[17] D. Crisan, C. Janjigian, and T. G. Kurtz. Particle representations for stochastic partial differential equations with boundary conditions. *arXiv preprint arXiv:1607.08909*, 2016.

[18] D. Croydon. Convergence of simple random walks on random discrete trees to Brownian motion on the continuum random tree. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(6):987–1019, 2008.

[19] D. A. Croydon. Volume growth and heat kernel estimates for the continuum random tree. *Probab. Theory Related Fields*, 140(1-2):207–238, 2008.

[20] D. A. Croydon. Scaling limits for simple random walks on random ordered graph trees. *Adv. in Appl. Probab.*, 42(2):528–558, 2010.

[21] D. A. Croydon and B. M. Hambly. Local limit theorems for sequences of simple random walks on graphs. *Potential Anal.*, 29(4):351–389, 2008.

[22] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014.

[23] K. D. Dalang, R.C. and F. Rassoul-Agha. *A minicourse on stochastic partial differential equations*. Springer-Verlag, 2009.

[24] R. C. Dalang and L. s. Quer-Sardanyons. Stochastic integrals for spde's: a comparison. *Expo. Math.*, 29(1):67–109, 2011.

[25] T. Delmotte. Parabolic Harnack inequality and estimates of Markov chains on graphs. *Rev. Mat. Iberoamericana*, 15(1):181–232, 1999.

[26] C. R. Doering, C. Mueller, and P. Smereka. Interacting particles, the stochastic Fisher-Kolmogorov-Petrovsky-Piscounov equation, and duality. *Phys. A*, 325(1-2):243–259, 2003. Stochastic systems: from randomness to complexity (Erice, 2002).

[27] P. Donnelly and T. G. Kurtz. Genealogical processes for Fleming-Viot models with selection and recombination. *Ann. Appl. Probab.*, 9(4):1091–1148, 1999.

[28] R. Durrett. Special invited paper: coexistence in stochastic spatial models. *Ann. Appl. Probab.*, 19(2):477–496, 2009.

[29] R. Durrett and W.-T. L. Fan. Genealogies in expanding populations. *Ann. Appl. Probab.*, 26(6):3456–3490, 2016.

[30] R. Durrett and S. Levin. The importance of being discrete (and spatial). *Theor. Pop. Biol.*, 46(3):363–394, 1994.

[31] R. Durrett and S. Levin. Stochastic spatial models: a user's guide to ecological applications. *Philos. Trans. R. Soc. London. Ser. B*, 343(1305):329–350, 1994.

[32] R. Durrett, L. Mytnik, and E. Perkins. Competing super-Brownian motions as limits of interacting particle systems. *Electron. J. Probab.*, 10:no. 35, 1147–1220, 2005.

[33] C. A. Edmonds, A. S. Lillie, and L. L. Cavalli-Sforza. Mutations arising in the wave front of an expanding population. *Proc. Natl. Acad. Sci*, 101(4):975–979, 2004.

[34] E. E. Endler, K. A. Duca, P. F. Nealey, G. M. Whitesides, and J. Yin. Propagation of viruses on micropatterned host cells. *Biotechnology and bioengineering*, 81(6):719–725, 2003.

[35] A. M. Etheridge and T. G. Kurtz. Genealogical constructions of population models. *arXiv preprint arXiv:1402.6724*, 2014.

[36] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[37] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[38] S. N. Evans and E. A. Perkins. Collision local times, historical stochastic calculus, and competing superprocesses. *Electron. J. Probab.*, 3:No. 5, 120, 1998.

[39] W.-T. L. Fan. Discrete approximations to local times for reflected diffusions. *Electron. Commun. Probab.*, 21:Paper No. 16, 12, 2016.

[40] W.-T. L. Fan, W. Hu, and G. Terlov. Wave propagation for reaction-diffusion equations on infinite random trees. *arXiv preprint arXiv:1907.12962*, 2019.

[41] M. Freidlin and S.-J. Sheu. Diffusion processes on graphs: stochastic differential equations, large deviation principle. *Probab. Theory Related Fields*, 116(2):181–220, 2000.

[42] M. I. Freidlin and A. D. Wentzell. Diffusion processes on graphs and the averaging principle. *Ann. Probab.*, 21(4):2215–2245, 1993.

[43] M. Fukushima, Y. Oshima, and M. Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, extended edition, 2011.

[44] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4):403–434, 1976.

[45] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977.

[46] P. c. Gonçalves, M. Jara, and S. Sethuraman. A stochastic Burgers equation from a class of microscopic interactions. *Ann. Probab.*, 43(1):286–338, 2015.

[47] A. Grigor'yan and A. Telcs. Two-sided estimates of heat kernels on metric measure spaces. *Ann. Probab.*, 40(3):1212–1284, 2012.

[48] V. Grimm and S. F. Railsback. *Individual-based modeling and ecology.* Princeton Series in Theoretical and Computational Biology. Princeton University Press, Princeton, NJ, 2005.

[49] M. Gubinelli, P. Imkeller, and N. Perkowski. Paracontrolled distributions and singular PDEs. *Forum Math. Pi*, 3:e6, 75, 2015.

[50] I. Gyöngy. On stochastic finite difference schemes. *Stoch. Partial Differ. Equ. Anal. Comput.*, 2(4):539–583, 2014.

[51] P. Gyrya and L. Saloff-Coste. *Neumann and Dirichlet heat kernels in inner uniform domains.* Société mathématique de France, 2011.

[52] M. Hairer. Solving the KPZ equation. *Ann. of Math. (2)*, 178(2):559–664, 2013.

[53] M. Hairer. A theory of regularity structures. *Invent. Math.*, 198(2):269–504, 2014.

[54] O. Hallatschek, P. Hersen, S. Ramanathan, and D. R. Nelson. Genetic drift at expanding frontiers promotes gene segregation. *Proc. Natl. Acad. Sci*, 104(50):19926–19930, 2007.

[55] O. Hallatschek and D. R. Nelson. Gene surfing in expanding populations. *Theor. Pop. Biol.*, 73(1):158–170, 2008.

[56] O. Hallatschek and D. R. Nelson. Life at the front of an expanding population. *Evolution*, 64(1):193–206, 2010.

[57] B. Inankur. *Development of Tools to Study how Virus-Host Cell Interactions Influence Infection Spread.* PhD dissertation. The University of Wisconsin-Madison, 2015.

[58] U. Kant, T. Klauss, J. Voigt, and M. Weber. Dirichlet forms for singular one-dimensional operators and on graphs. *J. Evol. Equ.*, 9(4):637–659, 2009.

[59] S. Kliem. Convergence of rescaled competing species processes to a class of SPDEs. *Electron. J. Probab.*, 16:no. 22, 618–657, 2011.

[60] S. Klopfstein, M. Currat, and L. Excoffier. The fate of mutations surfing on the wave of a range expansion. *Molecular biology and evolution*, 23(3):482–490, 2005.

[61] K. S. Korolev, M. J. Müller, N. Karahan, A. W. Murray, O. Hallatschek, and D. R. Nelson. Selective sweeps in growing microbial colonies. *Physical biology*, 9(2):026008, 2012.

[62] P. M. Kotelenez and T. G. Kurtz. Macroscopic limits for stochastic partial differential equations of McKean-Vlasov type. *Probab. Theory Related Fields*, 146(1-2):189–222, 2010.

[63] T. Kumagai. Heat kernel estimates and parabolic Harnack inequalities on graphs and resistance forms. *Publ. Res. Inst. Math. Sci.*, 40(3):793–818, 2004.

[64] T. G. Kurtz and J. Xiong. Particle representations for a class of nonlinear SPDEs. *Stochastic Process. Appl.*, 83(1):103–126, 1999.

[65] N. Lanchier. Stochastic spatial model of producer-consumer systems on the lattice. *Adv. in Appl. Probab.*, 45(4):1157–1181, 2013.

[66] N. Lanchier and C. Neuhauser. Stochastic spatial models of host-pathogen and host-mutualist interactions. I. *Ann. Appl. Probab.*, 16(1):448–474, 2006.

[67] N. Lanchier and C. Neuhauser. Stochastic spatial models of host-pathogen and host-mutualist interactions II. *Stoch. Models*, 26(3):399–430, 2010.

[68] J.-F. Le Gall. A class of path-valued Markov processes and its applications to superprocesses. *Probab. Theory Related Fields*, 95(1):25–46, 1993.

[69] Y. Lee and J. Yin. Detection of evolving viruses. *Nature biotechnology*, 14(4):491–493, 1996.

[70] R. Lehe, O. Hallatschek, and L. Peliti. The rate of beneficial mutations surfing on the wave of a range expansion. *PLoS computational biology*, 8(3):e1002447, 2012.

[71] A. Lejay. Simulating a diffusion on a graph. application to reservoir engineering. *Monte Carlo Methods and Applications mcma*, 9(3):241–255, 2003.

[72] Z. M. Ma and M. Röckner. *Introduction to the theory of (nonsymmetric) Dirichlet forms.* Universitext. Springer-Verlag, Berlin, 1992.

[73] C. Mueller. On the support of solutions to the heat equation with noise. *Stochastics Stochastics Rep.*, 37(4):225–245, 1991.

[74] C. Mueller, L. Mytnik, and J. Quastel. Effect of noise on front propagation in reaction-diffusion equations of KPP type. *Invent. Math.*, 184(2):405–453, 2011.

[75] C. Mueller, L. Mytnik, and L. Ryzhik. The speed of a random front for stochastic reaction-diffusion equations with strong noise. *arXiv preprint arXiv:1903.03645*, 2019.

[76] C. Mueller and R. Tribe. A phase transition for a stochastic PDE related to the contact process. *Probab. Theory Related Fields*, 100(2):131–156, 1994.

[77] C. Müller and R. Tribe. Stochastic p.d.e.'s arising from the long range contact and long range voter processes. *Probab. Theory Related Fields*, 102(4):519–545, 1995.

[78] J. Nullmeier and O. Hallatschek. The coalescent in boundary-limited range expansions. *Evolution*, 67(5):1307–1320, 2013.

[79] J. M. Ramirez. Multi-skewed Brownian motion and diffusion in layered media. *Proc. Amer. Math. Soc.*, 139(10):3739–3752, 2011.

[80] S. Riley, K. Eames, V. Isham, D. Mollison, and P. Trapman. Five challenges for spatial epidemic models. *Epidemics*, 10:68–71, 2015.

[81] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition.

[82] L. Saloff-Coste. Parabolic Harnack inequality for divergence-form second-order differential operators. *Potential Anal.*, 4(4):429–467, 1995. Potential theory and degenerate partial differential operators (Parma).

[83] C. Seifert and J. Voigt. Dirichlet forms for singular diffusion on graphs. *Oper. Matrices*, 5(4):723–734, 2011.

[84] T. Shiga. Stepping stone models in population genetics and population dynamics. In *Stochastic processes in physics and engineering (Bielefeld, 1986)*, volume 42 of *Math. Appl.*, pages 345–355. Reidel, Dordrecht, 1988.

[85] T. Shiga. Two contrasting properties of solutions for one-dimensional stochastic partial differential equations. *Canad. J. Math.*, 46(2):415–437, 1994.

[86] D. W. Stroock. Diffusion semigroups corresponding to uniformly elliptic divergence form operators. In *Séminaire de probabilités XXII*, pages 316–347. Springer, 1988.

[87] A. Sturm. On convergence of population processes in random environments to the stochastic heat equation with colored noise. *Electron. J. Probab.*, 8:no. 6, 39, 2003.

[88] X. Su, A. B. Theberge, C. T. January, and D. J. Beebe. Effect of microculture on cell metabolism and biochemistry: do cells get stressed in microchannels? *Analytical chemistry*, 85(3):1562–1570, 2013.

[89] S. Tindel and F. Viens. On space-time regularity for the stochastic heat equation on Lie groups. *J. Funct. Anal.*, 169(2):559–603, 1999.

[90] S. Tindel and F. Viens. Almost sure exponential behaviour for a parabolic SPDE on a manifold. *Stochastic Process. Appl.*, 100:53–74, 2002.

[91] M. Wallace, R. Feres, and G. Yablonsky. Reaction-diffusion on metric graphs: from 3D to 1D. *Comput. Math. Appl.*, 73(9):2035–2052, 2017.

[92] J. B. Walsh. An introduction to stochastic partial differential equations. In *École d'été de probabilités de Saint-Flour, XIV—1984*, volume 1180 of *Lecture Notes in Math.*, pages 265–439. Springer, Berlin, 1986.