LOCALLY INTERACTING DIFFUSIONS AS MARKOV RANDOM FIELDS ON PATH SPACE

DANIEL LACKER, KAVITA RAMANAN, AND RUOYU WU

ABSTRACT. We consider a countable system of interacting (possibly non-Markovian) stochastic differential equations driven by independent Brownian motions and indexed by the vertices of a locally finite graph $G = (V, E)$. The drift of the process at each vertex is influenced by the states of that vertex and its neighbors, and the diffusion coefficient depends on the state of only that vertex. Such processes arise in a variety of applications including statistical physics, neuroscience, engineering and math finance. Under general conditions on the coefficients, we show that if the initial conditions form a second-order Markov random field on d-dimensional Euclidean space, then at any positive time, the collection of histories of the processes at different vertices forms a second-order Markov random field on path space. We also establish a bijection between (second-order) Gibbs measures on $(\mathbb{R}^d)^V$ (with finite second moments) and a set of (second-order) Gibbs measures on path space, corresponding respectively to the initial law and the law of the solution to the stochastic differential equation. As a corollary, we establish a Gibbs uniqueness property that shows that for infinite graphs the joint distribution of the paths is completely determined by the initial condition and the specifications, namely the family of conditional distributions on finite vertex sets given the configuration on the complement. Along the way, we establish approximation and projection results for Markov random fields on locally finite graphs that may be of independent interest.

CONTENTS

1. INTRODUCTION

1.1. Discussion of results. Given a finite or locally finite infinite graph G with vertex set V and edge set E, and a positive integer d, consider interacting diffusions that satisfy the following

stochastic differential equation (SDE) system:

$$dX_v(t) = b_v(t, X_v(t), X_{N_v(G)}(t)) \, dt + \sigma_v(t, X_v(t)) \, dW_v(t), \quad v \in V, \quad t \geq 0,$$

where the initial condition $X(0) = (X_v(0))_{v \in V}$ is distributed according to some given probability measure on $(\mathbb{R}^d)^V$. Here, $N_v(G) \subset V$ denotes the neighborhood of v in the graph G, $(b_v, \sigma_v)_{v \in V}$ are given drift and diffusion coefficients, and $(W_v)_{v \in V}$ are independent standard d-dimensional Brownian motions. Diffusions indexed by vertices of a large graph arise in a variety of contexts, including statistical physics [5, 24], as well as recent extensions of models stemming from neuroscience [19, 20] and systemic risk [22]. Under suitable conditions on the coefficients that guarantee the existence of a unique weak solution to the SDE, for any $t > 0$, we study the random field on the space $\mathcal{C}_t^V$ generated by the collection of trajectories $(X_v[t] := (X_v(s))_{s \leq t})_{v \in V}$, where $\mathcal{C}_t$ (resp. $\mathcal{C}$) denotes the space of $\mathbb{R}^d$-valued continuous functions on $[0, t]$ (resp. $[0, \infty)$).

Our first set of results (Theorems 2.4 and 2.7) show that, under modest conditions on the drift and diffusion coefficients that guarantee a unique weak solution to the SDE system on any locally finite graph, if $(X_v(0))_{v \in V}$ is a second-order Markov random field on $(\mathbb{R}^d)^V$ (as specified in Definition 1.1) then for each $t > 0$, $(X_v[t])_{v \in V}$ is a second-order Markov random field on $\mathcal{C}_t^V$. In fact, we establish this result for a more general class of SDEs, defined at the beginning of Section 2, with possibly non-Markovian dynamics (and potentially infinite memory).

Our next set of results relate to an interpretation of the law of the SDE as a Gibbs measure on path space (see Section 2.3 for precise definitions). Specifically, Theorem 2.9 establishes a bijection between (second-order) Gibbs measures on $(\mathbb{R}^d)^V$ (with finite second moments) and a set of (second-order) Gibbs measures on $\mathcal{C}^V$, corresponding respectively to the initial law and the law of the solution to the SDE. As a consequence, we deduce a Gibbs uniqueness property, which shows that the law of the SDE system is completely determined by its initial condition and its specifications, namely the family of conditional distributions on finite sets given the configuration on the complement. In particular, together these show (see Corollary 2.10) that when the initial distribution is the unique second-order Gibbs measure associated with some specifications on $(\mathbb{R}^d)^V$, then for each $t > 0$, the law of the SDE system is the unique second-order Gibbs measure associated with corresponding specifications on path space $\mathcal{C}_t^V$.

A key motivation for our study stems from recent results in [17] that show how a second-order Markov property is useful for obtaining an autonomous description of the marginal (local) dynamics of a particle and its neighborhood when the underlying graph G is a tree. For this purpose, a stronger *global* Markov property is derived in [17, Proposition 3.15] in the setting of an infinite regular tree G (or, more generally, a unimodular Galton-Watson tree) and homogeneous coefficients, $(b_v, \sigma_v) = (b, \sigma)$ for all $v \in V$. Notably, the characterization of the local dynamics in [17] relies on the precise order of the Markov random field (equivalently, range of interaction of the Gibbs state), and not merely the Gibbs property. Such an autonomous description, together with the local convergence result [16, Theorem 3.7] provides a complete and tractable law of large number result for interacting diffusions on certain growing sparse networks.

In addition, such characterizations of SDEs in terms of Markov random fields are also of broader interest, although past work has mainly focused on the case of interacting diffusions on $\mathbb{Z}^m$, and often with an additional gradient structure imposed on the drift. The earliest work in this direction appears to be that of Deuschel [10], who considered smooth, uniformly bounded drifts of finite range that are locally of gradient type. The approach in [10] is based on estimates of Dobrushin's contraction coefficient, and crucially relies on the uniform boundedness and gradient structure of the drift (see [10, Remark (3.5)i)]). This result was later generalized to the case of unbounded drifts, though still of locally gradient form, by Cattiaux, Roelly and Zessin [3]. They used Malliavin calculus, a variational characterization and an integration-by-parts formula (see also [21, 26]), all of which exploit the gradient structure of the drift. In [21], an alternative

cluster expansion method was also used when the gradient system can be viewed as a small perturbation of a free field. Space-time cluster expansion techniques have also been applied to non-gradient systems (although to address the slightly different question of the Gibbsian nature of marginals; see also Remark 3.6), but these only apply for sufficiently small time horizons [24] or sufficiently small memory and interaction strengths, measured in terms of the norm of the drift functional [25]. An important work that goes beyond the gradient setting is that of Dereudre and Roelly [8], which considers a system of interacting one-dimensional SDEs on $\mathbb{Z}^m$ with an adapted (possibly history-dependent) drift that is homogeneous (i.e., the same at all vertices of $\mathbb{Z}^m$) and satisfies a linear growth condition, and with a constant diffusion coefficient equal to the identity matrix. In order to circumvent the restrictive bounds on the the uniform norm of the drift b or the time duration that arise in cluster expansion approaches, the work [8] employs another technique, referred to as the so-called entropy method. However, the latter work imposes the restriction that the initial condition is stationary (i.e., shift-invariant on $\mathbb{Z}^m$), and has finite specific entropy and second moment, and the drift has sublinear growth, and show that the trajectories of the SDE system form a Gibbs or Markov random field on $\mathcal{C}_t^{\mathbb{Z}^m}$, though in a somewhat different sense that involves conditioning on all initial positions $(X_v(0))_{v \in V}$, not just those outside a given finite collection of vertices.

As mentioned above, we consider the much more general setting of (possibly non-Markovian and time-inhomogeneous) interacting diffusions on arbitrary locally finite graphs $G = (V, E)$, with vertex-dependent, adapted, finite-range drift coefficients, and diffusion coefficients that may depend both on the vertex and *state* (although we do impose continuity of the drift coefficients since we want uniqueness of weak solutions). Moreover, we do not assume that the initial conditions are shift-invariant (or, more generally, invariant under automorphisms of the graph). Thus, we have to develop new techniques to prove our results. First, to establish the Markov random field property for SDEs on finite graphs (Theorem 2.4), we apply a version of the Hammersley-Clifford characterization of Markov random fields on finite graphs. The result for the infinite graph (Theorem 2.7) is then obtained by a delicate approximation by finite-dimensional systems in a way that preserves the Markov random field property, along with careful relative entropy estimates. Along the way, we establish approximation results for Markov random fields on locally finite graphs (see Section 5) that may be of independent interest. Thus, our methods are quite different from those of the prior works described above. They do not require any gradient structure of the drift, and are not restricted to small perturbations of a free field. Like [8], we allow for unbounded drifts and non-Markovian dynamics involving path-dependent coefficients. However, in constrast to [8] we consider arbitrary locally finite graphs and more general initial conditions. Indeed, although the work of [8] also uses finite-volume approximations, the pervasive stationarity assumptions they impose allow them to rely on specific entropy, which cannot be used in our setting, and they also use a growth bound on $\mathbb{Z}^m$ (e.g., in deriving equation (30) in the proof of Proposition 3.4 therein), which does not hold for arbitary locally finite graphs. In addition, while the main goal in [8] is to construct shift-invariant solutions of their SDE, with the Gibbsian (or Markov random field) description of their process coming as a by-product, our objective is to identify the extent to which the second-order Markov random field property holds, for which no shift-invariance is required. Lastly, the correspondence between Gibbs measures on initial configurations on $(\mathbb{R}^d)^V$ and Gibbs measures on the whole path space $\mathcal{C}^V$ that we establish in Theorem 2.9 and Corollary 2.10 does not appear in [8], although in the gradient seetting, the idea of such a correspondence appears in in [3] (see Theorems 3.7 and 4.9 therein).

Finally, we also provide examples (see Section 3.3) that demonstrate that the Markov random field property we establish cannot in general be significantly strengthened. Precisely, even on a finite graph with gradient drift, in general the collection of histories $(X_v[t])_{v \in V}$ do not form

a first-order Markov random field, nor do the time-t marginals $(X_v(t))_{v \in V}$ exhibit any non-trivial conditional independence structure. This highlights the natural problem of identifying special classes of systems for which simpler Markov random field properties are preserved, a problem which we do not address but which has attracted considerable attention in certain contexts. Specifically, in the context of diffusions, the papers [6, 7, 15, 24, 25, 32, 33] have studied the phenomenon of Gibbs-non-Gibbs transitions and the propagation (or lack thereof) of the Gibbs property at the level of the time-t marginals, specifically whether the initial law of $X(0)$ being a Gibbs state on $(\mathbb{R}^d)^V$ implies that the marginal law of $X(t)$ is also a Gibbs state on $(\mathbb{R}^d)^V$. See Remark 3.6 for a more detailed description of these works.

The next section introduces some common notation and basic definitions used throughout the paper. The main results of the paper are stated in Section 2, with their proofs relegated to Sections 3–6.

1.2. **Notation and basic definitions.** For any vectors $a, b \in \mathbb{R}^d$, we use $a \cdot b$ or $\langle a, b \rangle$ to denote the inner product. In this paper, unless explicitly stated otherwise, a *graph* $G = (V, E)$ always has a finite or countably infinite vertex set, is simple (no self-edges or multi-edges), and is locally finite (i.e., the degree of each vertex is finite). We abuse notation by writing $v \in G$ to mean $v \in V$. For a graph $G = (V, E)$ and a vertex $v \in V$, we write $N_v(G) = \{u \in V : (u, v) \in E\}$ for the set of neighbors of v in G, noting that this set is empty if v is an isolated vertex. A *rooted graph* $G = (V, E, \emptyset)$ is a graph equipped with a distinguished vertex $\emptyset \in V$, called the *root*. For two vertices $u, v \in V$, let $d(u, v)$ denote the graph distance, i.e., the length of the shortest path from u to v (with $d(u, u) := 0$). Also, let $\mathrm{diam}(A)$ denote the diameter of a set $A \subset V$; precisely, $\mathrm{diam}(A) = \sup\{d(u, v) : u, v \in A\}$. For a subset $A \subset V$, we define the first and second boundaries

$$
\begin{aligned}
\partial_G A &= \{u \in V \backslash A : (u, v) \in E \text{ for some } v \in A\}, \\
\partial_G^2 A &= \partial_G A \cup \partial_G(A \cup \partial_G A).
\end{aligned}
\tag{1.1}
$$

We will often omit the subscript, writing simply $\partial^2 A$ in place of $\partial_G^2 A$, when the underlying graph G is clear. A *clique* in a graph G is a complete subgraph of G, i.e., a set $A \subset V$ such that $(u, v) \in E$ for every $u, v \in A$. Equivalently, a clique is a set $A \subset V$ of diameter at most 1. Define $\mathrm{cl}_1(G)$ to be the set of all cliques of the graph G. Similarly, we will say that any subset $A \subset V$ with diameter at most 2 is a *2-clique* of the graph G and let $\mathrm{cl}_2(G)$ denote the set of 2-cliques of G. Moreover, given a graph $G = (V, E)$, $H = (V_H, E_H)$ is said to be an induced subgraph of G if $V_H \subset V$ and $E_H = E \cap \{(u, v) : u, v \in V_H\}$.

For a set $\mathcal{X}$ and a graph $G = (V, E)$, we may write either $\mathcal{X}^V$ or $\mathcal{X}^G$ for the configuration space $\{(x_v)_{v \in V} : x_v \in \mathcal{X} \text{ for every } v \in V\}$. We make use of a standard notation for configurations on subsets of vertices: For $x = (x_v)_{v \in V} \in \mathcal{X}^V$ and $A \subset V$, we write x_A for the element $x_A = (x_v)_{v \in A}$ of $\mathcal{X}^A$. When $\mathcal{X}$ is a Polish space, we write $\mathcal{B}(\mathcal{X})$ for the Borel σ-field, and write $\mathcal{P}(\mathcal{X})$ for the set of Borel probability measures on $\mathcal{X}$, endowed always with the topology of weak convergence. Given any measurable space $\mathcal{X}$, $A \subset V$, and measure $\nu \in \mathcal{P}(\mathcal{X}^V)$, $\nu[A]$ represents the restriction of ν to the set $\mathcal{X}^A$, that is, the image measure under the restriction map $\mathcal{X}^V \ni (x_v)_{v \in V} \mapsto (x_v)_{v \in A} \in \mathcal{X}^A$.

Fixing $d \in \mathbb{N}$, we let $\mathcal{C} = C(\mathbb{R}_+; \mathbb{R}^d)$ denote the path space of $\mathbb{R}^d$-valued continuous functions on $\mathbb{R}_+ = [0, \infty)$, endowed with the topology of uniform convergence on compacts. For $t > 0$, let $\mathcal{C}_t = C([0, t]; \mathbb{R}^d)$ denote its restriction to the time interval $[0, t]$, endowed with the uniform topology. For $x \in \mathcal{C}$ and $t > 0$ we define $\|x\|_{*,t} := \sup_{s \in [0, t]} |x(s)|$, and let $x[t] = (x(s))_{s \leq t} \in \mathcal{C}_t$ denote the restriction of the path x to the time interval $[0, t]$. We assume that $\mathcal{C}$ and $\mathcal{C}_t$ are endowed with their respective Borel σ-algebras. Also, for any countable set A and probability measure Q on $\mathcal{C}^A$, we write Q_t for the image under Q of the map $\mathcal{C}^A \ni (x_v)_{v \in A} \mapsto (x_v[t])_{v \in A} \in \mathcal{C}_t^A$. The σ-algebra on a product space will always be the product σ-algebra, unless explicitly

stated otherwise. Given $J, m \in \mathbb{N}$, a measurable function $f : [0, \infty) \times \mathcal{C}^J \mapsto \mathbb{R}^m$, is said to be *progressively measurable* if for each $t \geq 0$, $f(t, (x_u)_{u=1,\ldots,J}) = f(t, (y_u)_{u=1,\ldots,J})$ whenever $x_u[t] = y_u[t]$ for all $u = 1, \ldots, J$.

We end this section by recalling the notion of a (first-order or second-order) Markov random field, which plays a central role in the paper.

Definition 1.1. Given a measurable space $\mathcal{X}$, and a (possibly infinite) locally finite graph $G = (V, E)$, let $(Y_v)_{v \in V}$ be a random element of $\mathcal{X}^V$ with some distribution $\nu \in \mathcal{P}(\mathcal{X}^V)$. We say that $(Y_v)_{v \in V}$, or equivalently its law ν, is a *first-order Markov random field* (abbreviated as 1MRF) on $\mathcal{X}^V$ if Y_A is conditionally independent of $Y_{(A \cup \partial A)^c}$ given $Y_{\partial A}$, for every finite set $A \subset V$. On the other hand, we say that $(Y_v)_{v \in V}$, or equivalently its law ν, is a *second-order Markov random field* (abbreviated as 2MRF) on $\mathcal{X}^V$ if Y_A is conditionally independent of $Y_{(A \cup \partial^2 A)^c}$ given $Y_{\partial^2 A}$, for every finite set $A \subset V$. When the space $\mathcal{X}^V$ is clear from the context, we will simply say that $(Y_v)_{v \in V}$, or equivalently its law ν, is a 1MRF or 2MRF.

Remark 1.2. In Definition 1.1, it is important to stress that the sets A are required to be finite even when the graph G is infinite. Allowing infinite sets A results in the stronger *global Markov property*, which we do not study in this paper.

2. MAIN RESULTS

Given a locally finite graph $G = (V, E)$ with a finite or countably infinite vertex set, we are interested in a system of (possibly non-Markovian) interacting stochastic processes, indexed by the vertices of the graph, that satisfy a (functional) SDE of the form

$$dX_v(t) = b_v(t, X_v, X_{N_v(G)}) \, dt + \sigma_v(t, X_v) \, dW_v(t), \quad v \in V, \tag{2.1}$$

where $(W_v)_{v \in V}$ are independent Brownian motions, and the initial law $\mu_0 \in \mathcal{P}((\mathbb{R}^d)^V)$, of $(X_v(0))_{v \in V}$ and the coefficients $(b_v, \sigma_v)_{v \in V}$ satisfy the conditions stated in Assumption **A** or Assumption **B** below, depending on whether the graph is finite or infinite. As mentioned in the introduction, our main results concern the characterization of the law of the solution to the SDE (2.1) as a 2MRF on the path space $\mathcal{C}^V$ (see Definition 1.1).

2.1. **The finite graph case.** We first consider the case when G is finite, and the conditions stated in Assumption **A** below are satisfied. Recall, from Section 1.2, the definition of 2-cliques, the notation for trajectories, $x[t] = (x(s))_{s \leq t} \in \mathcal{C}_t$ for $x \in \mathcal{C}$, and the notion of a progressively measurable functional.

Assumption A. We say that (G, b, σ, μ_0) satisfy Assumption **A** if $G = (V, E)$ is a finite graph and if $b = (b_v)_{v \in V}$, $\sigma = (\sigma_v)_{v \in V}$, and $\mu_0 \in \mathcal{P}((\mathbb{R}^d)^V)$ satisfy the following:

(**A**.1) There exist $\lambda_v \in \mathcal{P}(\mathbb{R}^d), v \in V$, such that the probability measure μ_0 is absolutely continuous with respect to the product measure $\mu_0^* = \prod_{v \in V} \lambda_v \in \mathcal{P}((\mathbb{R}^d)^V)$ and the density satisfies

$$\frac{d\mu_0}{d\mu_0^*}(x) = \prod_{K \in \mathrm{cl}_2(G)} f_K(x_K), \qquad x \in (\mathbb{R}^d)^V, \tag{2.2}$$

for some measurable functions $f_K : (\mathbb{R}^d)^K \to \mathbb{R}_+$, $K \in \mathrm{cl}_2(G)$, where $\mathrm{cl}_2(G)$ is the set of 2-cliques of G. In addition, μ_0 has a finite second moment.

(**A**.2) For each $v \in V$, the drift $b_v : \mathbb{R}_+ \times \mathcal{C} \times \mathcal{C}^{N_v(G)} \mapsto \mathbb{R}^d$ is progressively measurable. Moreover, for each $T \in (0, \infty)$ there exists $C_T < \infty$ such that

$$|b_v(t, x, y_{N_v(G)})| \leq C_T \left(1 + \|x\|_{*,t} + \sum_{u \in N_v} \|y_u\|_{*,t} \right),$$

for all $v \in V$, $t \in [0, T]$, $x \in \mathcal{C}$, and $y_{N_v(G)} = (y_u)_{u \in N_v(G)} \in \mathcal{C}^{N_v(G)}$.

(**A**.3) The diffusion matrices $\sigma_v : \mathbb{R}_+ \times \mathcal{C} \to \mathbb{R}^{d \times d}$, $v \in V$, satisfy the following:

(**A**.3a) For each $v \in V$, σ_v is bounded, progressively measurable and invertible, with bounded inverse.

(**A**.3b) For each $v \in V$, the following driftless SDE system admits a unique in law weak solution starting from any initial position $x \in \mathbb{R}^d$:

$$dX_v(t) = \sigma_v(t, X_v)\,dW_v(t), \quad X_v(0) = x. \tag{2.3}$$

Remark 2.1. A necessary condition for Assumption (**A**.1) is that μ_0 is a 2MRF and is absolutely continuous with respect to the product measure μ_0^*; this follows from a form of the Hammersley-Clifford theorem stated in Proposition 3.2 below. If the density $d\mu_0/d\mu_0^*$ is strictly positive, then it factorizes as in (2.2) if and only if μ_0 is a 2MRF.

Remark 2.2. If $\sigma_v(t, x) = \widetilde{\sigma}_v(t, x(t))$ depends only on the current state, not the history, and satisfies the additional continuity condition $\lim_{y \to x} \sup_{0 \leq s \leq T} |\widetilde{\sigma}_v(s, y) - \widetilde{\sigma}_v(s, x)| = 0$ for all $v \in V$, then Assumption (**A**.3b) holds as a consequence of Assumption (**A**.3a) and [29, Chapter 7].

The following proposition shows that, as a simple consequence of Girsanov's theorem, Assumption **A** guarantees weak existence and uniqueness in law of the SDE system (2.1). Its proof is given in Section 3.2, along the way to proving Theorem 2.4 below.

Proposition 2.3. *When (G, b, σ, μ_0) satisfy Assumption **A**, the SDE system (2.1) has a weak solution that is unique in law.*

We now state our main result for the SDE system on finite graphs.

Theorem 2.4. *Suppose $(G = (V, E), b, \sigma, \mu_0)$ satisfy Assumption **A**, and let $(X_v)_{v \in V}$ be the unique (in law) solution of the SDE system (2.1) with initial law μ_0. Then, for each $t > 0$, $(X_v[t])_{v \in V}$ is a 2MRF on $\mathcal{C}_t^V$. Moreover, $(X_v)_{v \in V}$ is a 2MRF on $\mathcal{C}^V$.*

The proof of Theorem 2.4, given in Section 3.2, relies on a certain factorization property (stated in Proposition 3.2) of the density of the law of the SDE on finite graphs with respect to a reference measure. Notice that in Assumption **A**, and throughout the paper, we assume there is no interaction in the diffusion coefficients (i.e., no dependence of σ_v on $X_{N_v(G)}$), a restriction made also in the prior works [3, 10, 21, 26]; the general case seems out of reach of our approach, because the reference measure in the factorization property must crucially be a *product measure*. This factorization property is also used in Sections 3.3.1 and 3.3.2 to show that, even when the initial states $(X_v(0))_{v \in V}$ are i.i.d., for $t > 0$, in general $(X_v[t])_{v \in V}$ fails to be a 1MRF, and the time-t marginals $(X_v(t))_{v \in V}$ can fail to be a Markov random field of either first or second order. In fact, the counterexamples show that this does not hold even on a finite graph when σ is the identity covariance matrix, and the drift is of gradient type. This shows that, in a sense, Theorem 2.4 cannot be strengthened.

2.2. **The infinite graph case.** We now consider the SDE system (2.1) in the case when G is an infinite, though still locally finite, graph. The well-posedness of the SDE system is no longer obvious and in particular does not follow from Girsanov's theorem as it did when the graph was finite. Indeed, on an infinite graph, when $b_v \equiv 1$ and $\sigma_v \equiv I_d$, $v \in V$, for instance, it is straightforward to argue, using the law of large numbers, that the law of a weak solution of (2.1) up to some time $t > 0$ and the law of the corresponding drift-free equation are mutually singular. This necessitates the following additional assumptions compared to Assumption **A**. Recall from Section 1.2 that given a measure ν on $\mathcal{X}^V$ for some Polish space $\mathcal{X}$, and $A \subset V$,

$\nu[A]$ denotes the restriction of ν to A. Also, we use the notation $\pi_1 \sim \pi_2$ to denote that the measures π_1 and π_2 are equivalent, that is, mutually absolutely continuous.

Assumption B.   We say that (G, b, σ, μ_0) satisfy Assumption **B** if $G = (V, E)$ is a countable locally finite connected graph and if $b = (b_v)_{v \in V}$, $\sigma = (\sigma_v)_{v \in V}$ and $\mu_0 \in \mathcal{P}((\mathbb{R}^d)^V)$ satisfy the following:

(**B**.1) The initial law μ_0 is a 2MRF on $(\mathbb{R}^d)^V$. Moreover, there exists a product measure $\mu_0^* = \prod_{v \in V} \lambda_v \in \mathcal{P}((\mathbb{R}^d)^V)$ such that $\mu_0[A] \sim \mu_0^*[A]$ for each finite set $A \subset V$. Further, the initial law μ_0 satisfies

$$\sup_{v \in V} \int_{(\mathbb{R}^d)^V} |x_v|^2 \, \mu_0(dx_V) < \infty. \tag{2.4}$$

(**B**.2) The drift coefficients $(b_v)_{v \in V}$ satisfy Assumption (**A**.2), for some constants $(C_T)_{T>0}$.
(**B**.3) The diffusion matrices $(\sigma_v)_{v \in V}$ satisfy Assumption (**A**.3).
(**B**.4) The SDE system (2.1) is unique in law, and this law is denoted by $P = P^{\mu_0} \in \mathcal{P}(\mathcal{C}^V)$.

Remark 2.5. Using Assumption (**A**.3b) if the graph is finite or Assumption (**B**.3) if the graph is infinite, we may define for any initial law $\nu \in \mathcal{P}((\mathbb{R}^d)^V)$ the measure $P^{*,\nu} \in \mathcal{P}(\mathcal{C}^V)$ to be the law of the unique weak solution of the SDE system

$$dX_v(t) = \sigma_v(t, X_v) \, dW_v(t), \quad v \in V, \quad (X_v(0))_{v \in V} \sim \nu. \tag{2.5}$$

Note in particular that if we take $\nu = \mu_0^*$, where μ_0^* is a product measure as in Assumption (**A**.1) or (**B**.1), then P^{*,μ_0^*} too is a product measure.

We show in Lemma 5.2 that existence of a solution to (2.1) follows automatically from Assumptions (**B**.1–3). However, it is worth commenting on the uniqueness condition in Assumption (**B**.4). The following proposition shows that a suitable Lipschitz condition is enough to guarantee uniqueness; its proof is standard and hence relegated to Appendix A. Recall in the following that $\|x\|_{*,t} = \sup_{s \in [0,t]} |x(s)|$ for $x \in \mathcal{C}$.

Proposition 2.6. *Suppose Assumptions (**B**.1–3) hold, and $(b_v)_{v \in V}$ and $(\sigma_v)_{v \in V}$ are uniformly Lipschitz in the sense that for each $T > 0$ there exist $K_T, \bar{K}_T < \infty$ such that*

$$|b_v(t, x, y_{N_v(G)}) - b_v(t, x', y'_{N_v(G)})| \leq K_T \left(\|x - x'\|_{*,t} + \frac{1}{|N_v(G)|} \sum_{u \in N_v(G)} \|y_u - y'_u\|_{*,t} \right),$$

$$|\sigma_v(t, x) - \sigma_v(t, x')| \leq \bar{K}_T \|x - x'\|_{*,t},$$

*for all $v \in V$, $t \in [0, T]$, $x, x' \in \mathcal{C}$, and $y_{N_v(G)}, y'_{N_v(G)} \in \mathcal{C}^{N_v(G)}$. Then pathwise uniqueness holds for the SDE system (2.1). In particular, Assumption (**B**.4) holds.*

We now state our second main result.

Theorem 2.7. *Suppose $(G = (V, E), b, \sigma, \mu_0)$ satisfy Assumption **B**, and let $(X_v)_{v \in V}$ be the unique in law solution of the SDE system (2.1) with initial law μ_0. Then, for each $t > 0$, $(X_v[t])_{v \in V}$ is a 2MRF on $\mathcal{C}_t^V$. Moreover, $(X_v)_{v \in V}$ is a 2MRF on $\mathcal{C}^V$.*

The proof of Theorem 2.7 is given in Section 5.2, with preparatory results established in Sections 4 and 5.1. The factorization result used in the finite graph case is no longer applicable in the infinite graph case, and thus the proof employs a completely different approach, involving a rather subtle approximation argument, which is outlined in Section 4.1.

2.3. **Gibbs measures on path space.** Our final results interpret our SDE system in the spirit of Gibbs measures, for which we introduce the following notation. Given a Polish space $\mathcal{X}$, a graph $G = (V, E)$, a random $\mathcal{X}^V$-valued element $(Y_v)_{v \in V}$ with law $\nu \in \mathcal{P}(\mathcal{X}^V)$, and two disjoint sets $A, B \subset V$, we write $\nu[A \,|\, B]$ to denote a version of the regular conditional law of Y_A given Y_B. Precisely, we view $\nu[A \,|\, B]$ as a measurable map (kernel) from $\mathcal{X}^B$ to $\mathcal{P}(\mathcal{X}^A)$. Note that ν is a 2MRF if and only if $\nu[A \,|\, V \backslash A](x_{V \backslash A}) = \nu[A \,|\, \partial^2 A](x_{\partial^2 A})$ for ν-almost every $x \in \mathcal{X}^V$ and every finite set A. We make use of the following terminology of Gibbs measures (see [11] or [23] for further discussion of this classical framework).

Definition 2.8. Given a Polish space $\mathcal{X}$, graph $G = (V, E)$, and $\gamma \in \mathcal{P}(\mathcal{X}^V)$, define $\mathcal{G}_2(\gamma)$ as the set of 2MRFs $\nu \in \mathcal{P}(\mathcal{X}^V)$ such that, for each finite set $A \subset V$, we have $\nu[A] \sim \gamma[A]$ and also $\nu[A \,|\, \partial^2 A] = \gamma[A \,|\, \partial^2 A]$, almost everywhere with respect to $\gamma[\partial^2 A]$.

In other words, $\mathcal{G}_2(\gamma)$ is the set of (second-order, infinite volume) Gibbs measures corresponding to the *specification* $\{\gamma[A \,|\, \partial^2 A] : A \subset V \text{ finite}\}$. Note that if γ is itself a 2MRF then $\mathcal{G}_2(\gamma)$ is nonempty, as it contains γ itself. Moreover, it is straightforward to check that, if γ and ν are 2MRFs, then $\nu \in \mathcal{G}_2(\gamma)$ if and only if $\gamma \in \mathcal{G}_2(\nu)$.

Recall that, by Assumption (**B**.4), the SDE system (2.1) is well-posed starting from any initial distribution. Recall also from Section 1.2 that for $P, Q \in \mathcal{P}(\mathcal{C}^V)$ we write P_t and Q_t for their projections onto $\mathcal{C}_t^V$. The following bijection result is proved in Section 6.

Theorem 2.9. *Suppose (G, b, σ, μ_0) satisfy Assumption **B**. Let $P^{\mu_0} \in \mathcal{P}(\mathcal{C}^V)$ be the law of the solution of the SDE system* (2.1) *with initial law μ_0 and define*

$$\mathcal{M}_{\mathrm{init}}(\mu_0) := \left\{ \nu_0 \in \mathcal{G}_2(\mu_0) : \sup_{v \in V} \int_{\mathbb{R}^d} |x_v|^2 \, \nu_0(dx) < \infty \right\},$$

and

$$\mathcal{M}_{\mathrm{path}}(\mu_0) := \left\{ Q \in \mathcal{P}(\mathcal{C}^V) : Q_t \in \mathcal{G}_2(P_t^{\mu_0}) \; \forall t \geq 0, \; \sup_{v \in V} \int_{\mathcal{C}^V} |x_v(0)|^2 Q(dx) < \infty \right\}.$$

Then it holds that

$$\mathcal{M}_{\mathrm{init}}(\mu_0) = \left\{ Q \circ (X_V(0))^{-1} : Q \in \mathcal{M}_{\mathrm{path}}(\mu_0) \right\}. \tag{2.6}$$

Moreover, the map $Q \mapsto Q \circ (X_V(0))^{-1}$ defines a bijection between $\mathcal{M}_{\mathrm{path}}(\mu_0)$ and $\mathcal{M}_{\mathrm{init}}(\mu_0)$. In particular, if $Q \in \mathcal{P}(\mathcal{C}^V)$ satisfies $Q_t \in \mathcal{G}_2(P_t^{\mu_0})$ for all $t \geq 0$ and also $Q \circ (X_V(0))^{-1} = \mu_0$, then $Q = P^{\mu_0}$.

In fact, we will show in the proof of Theorem 2.9 that the bijection $Q \mapsto Q \circ (X_V(0))^{-1}$ between the sets $\mathcal{M}_{\mathrm{path}}(\mu_0)$ and $\mathcal{M}_{\mathrm{init}}(\mu_0)$ has inverse given by $\nu_0 \mapsto P^{\nu_0}$, where P^{ν_0} denotes the law of the solution of the SDE (2.1) with initial law ν_0, and we note that this SDE is unique in law by Assumption (**B**.4). Additionally, if $\mu_0(K^V) = 1$ for some compact set $K \subset \mathbb{R}^d$, then (recalling that membership in $\mathcal{G}_2(\cdot)$ requires absolute continuity) (2.6) can be rewritten as

$$\mathcal{G}_2(\mu_0) = \left\{ Q \circ (X_V(0))^{-1} : Q \in \mathcal{P}(\mathcal{C}^V), \, Q_t \in \mathcal{G}_2(P_t^{\mu_0}) \; \forall t \geq 0 \right\}.$$

We conclude this section with the following simple corollary of Theorems 2.7 and 2.9.

Corollary 2.10. *Suppose (G, b, σ, μ_0) satisfy Assumption **B** and P^{μ_0} represents the unique law of the SDE* (2.1). *If $\mathcal{G}_2(\mu_0)$ is a singleton, then the set $\mathcal{M}_{\mathrm{path}}(\mu_0)$ defined in Theorem 2.9 is equal to the singleton $\{P^{\mu_0}\}$, and hence, P^{μ_0} is completely characterized by its specifications $P_t^{\mu_0}[A|\partial^2 A], t \geq 0$, for finite $A \subset V$.*

Proof. Since Assumption (**B**.1) ensures that μ_0 is a 2MRF with finite second moment, the set on the left-hand side of (2.6) always contains μ_0. Thus, if $\mathcal{G}_2(\mu_0)$ is a singleton, then by Theorem 2.9 the set $\mathcal{M}_{\text{path}}(\mu_0)$ is also a singleton. On the other hand, by Theorem 2.7, for each $t \geq 0$ it holds that $P_t^{\mu_0}$ is a 2MRF and thus $P_t^{\mu_0} \in \mathcal{G}_2(P_t^{\mu_0})$. Hence, $P^{\mu_0} \in \mathcal{M}_{\text{path}}(\mu_0)$.   □

3. Interacting diffusions on a finite graph

In Section 3.1 (specifically Proposition 3.2) a useful characterization of a (positive) 2MRF is derived in an abstract setting. This is then used in Section 3.2 to prove Theorem 2.4; along the way Proposition 2.3 is also established. In Sections 3.3.1 and 3.3.2 this characterization is applied to demonstrate via explicit examples that the path-space 2MRF property established in Theorem 2.4 (and hence, Theorem 2.7) is sharp, in the sense that it cannot be improved to a 1MRF property, and in the sense that there is in general no 1MRF or 2MRF property for the time-t marginal laws.

3.1. **Clique factorizations.** We start by studying the relationship between random fields and factorization properties of their joint density with respect to a given reference measure. Throughout this section, we work with a fixed finite graph $G = (V, E)$, as well as a fixed Polish space $\mathcal{X}$, the state space. Recall the definition of the diameter $\text{diam}(A)$ of a set $A \subset V$, 1-cliques and 2-cliques of a graph, and 1st-order and 2nd-order MRFs given in Section 1.2.

First, we recall a well-known theorem often attributed to Hammersley-Clifford, which can be found in various forms in [11, Theorem 2.30] and [18, Proposition 3.8 and Theorem 3.9], the latter covering our precise setting.

Proposition 3.1 (Hammersley-Clifford). *Assume the graph $G = (V, E)$ is finite. Assume $\nu \in \mathcal{P}(\mathcal{X}^V)$ is absolutely continuous with respect to a product measure $\nu^* = \prod_{v \in V} \theta_v \in \mathcal{P}(\mathcal{X}^V)$ for some $\theta_v \in \mathcal{P}(\mathcal{X})$, $v \in V$. Consider the following statements:*

(1) ν is a 1MRF.
(2) The density of ν with respect to ν^* factorizes in the form

$$\frac{d\nu}{d\nu^*}(x) = \prod_{K \in \text{cl}_1(\text{G})} f_K(x_K), \qquad x \in \mathcal{X}^V,$$

for some measurable functions $f_K : \mathcal{X}^K \to \mathbb{R}_+$, for $K \in \text{cl}_1(\text{G})$.
Then (2) implies (1). If also $d\nu/d\nu^*$ is strictly positive, then (1) implies (2).

We next formulate an analogue for a 2MRF.

Proposition 3.2 (Second-order Hammersley-Clifford). *Assume the graph $G = (V, E)$ is finite. Assume $\nu \in \mathcal{P}(\mathcal{X}^V)$ is absolutely continuous with respect to a product measure $\nu^* = \prod_{v \in V} \theta_v \in \mathcal{P}(\mathcal{X}^V)$ for some $\theta_v \in \mathcal{P}(\mathcal{X})$, $v \in V$. Consider the following statements:*

(1) ν is a 2MRF.
(2) The density of ν with respect to ν^* factorizes in the form

$$\frac{d\nu}{d\nu^*}(x) = \prod_{K \in \text{cl}_2(\text{G})} f_K(x_K), \qquad x \in \mathcal{X}^V, \tag{3.1}$$

for some measurable functions $f_K : \mathcal{X}^K \to \mathbb{R}_+$, for $K \in \text{cl}_2(\text{G})$.
Then (2) implies (1). If also $d\nu/d\nu^*$ is strictly positive, then (1) implies (2).

Proof. Define the *square graph* $G^2 = (V, E')$ by connecting any two vertices of distance 2. That is, let

$$E' := \{(u, v) \in V^2 : 1 \leq d(u, v) \leq 2\},$$

where d is the graph distance on G. It is straightforward to check the following properties:

(i) The 1-cliques of G^2 are precisely the 2-cliques of G. That is, $\mathrm{cl}_2(G) = \mathrm{cl}_1(G^2)$.

(ii) We have $\partial_{G^2} A = \partial^2_G A$ for any set $A \subset V$.

It follows from (ii) that the statement (1) is equivalent to

(1') ν is a 1MRF relative to the graph G^2.

On the other hand, it follows from (i) that (2) is equivalent to

(2') The density of ν with respect to ν^* factorizes in the form

$$\frac{d\nu}{d\nu^*}(x) = \prod_{K \in \mathrm{cl}_1(\mathrm{G}^2)} f_K(x_K), \qquad x \in \mathcal{X}^V,$$

for some measurable functions $f_K : \mathcal{X}^K \to \mathbb{R}_+$, $K \in \mathrm{cl}_1(\mathrm{G}^2)$.

The equivalence of (1') and (2') follows from Proposition 3.1. $\square$

The 2MRF property is the more intuitive, but the second property of Proposition 3.2 will be quite useful in the analysis as well. Hence, we give it a name.

Definition 3.3. We say that $\nu \in \mathcal{P}(\mathcal{X}^V)$ *has a 2-clique factorization with respect to ν^** if the density $d\nu/d\nu^*$ can be written in the form (3.1).

Remark 3.4. For a finite graph $G = (V, E)$ and Polish space $\mathcal{X}$, the following *cutset* characterization of 1MRF's on $\mathcal{X}^V$ is well known: An $\mathcal{X}^V$-valued random element $(Y_v)_{v \in V}$ is a 1MRF if and only if Y_A is conditionally independent of Y_B given Y_S for any disjoint sets $A, B, S \subset V$ with the property that every path starting in A and ending in B contains at least one vertex of S. Given the correspondence between a 2MRF on a graph and a 1MRF on the square graph (established in the proof of Proposition 3.2), this is easily seen to imply the following *cutset* characterization of 2MRFs: An $\mathcal{X}^V$-valued random element $(Y_v)_{v \in V}$ is a 2MRF if and only if Y_A is conditionally independent of Y_B given Y_S for any disjoint sets $A, B, S \subset V$ with the property that every path starting in A and ending in B contains at least two adjacent vertices of S.

3.2. **Proof of the second-order Markov random field property for a finite graph.** We now present the proofs of Proposition 2.3 and Theorem 2.4. Throughout this section, we work with a fixed finite graph $G = (V, E)$ and consider the canonical measurable space $\mathcal{C}^V = (\mathcal{C}^V, \mathcal{B}(\mathcal{C}^V))$, and let $(X_v)_{v \in V} : \mathcal{C}^V \to \mathcal{C}^V$ denote the canonical processes, that is, $X_v((x_u)_{u \in V}) = x_v$ for $x = (x_u)_{u \in V} \in \mathcal{C}^V$, for $v \in V$. Let $\mu_0, \mu_0^* \in \mathcal{P}((\mathbb{R}_d)^V)$ be as in Assumption (**A**.1), and let $P^* = P^{*,\mu_0^*} \in \mathcal{P}(\mathcal{C}^V)$ denote the law of the unique solution of the driftless SDE system (2.5) starting from initial law μ_0^* (the well-posedness of which is given by Assumption (**A**.3b)). Recall that μ_0^* and thus P^* are both product measures. Then, recalling that $dX_v(t) = \sigma_v(t, X_v) \, dW_v(t)$ for $v \in V$, define the following local martingale (under P^*):

$$M_v(t) := \int_0^t (\sigma_v \sigma_v^\top)^{-1} b_v(s, X_v, X_{N_v(G)}) \cdot dX_v(s), \qquad t \geq 0, \tag{3.2}$$

where we use the shorthand notation $(\sigma_v \sigma_v^\top)^{-1} b_v(s, x_v, x_{N_v(G)})$ to denote the map

$$\mathbb{R}_+ \times \mathcal{C}^V \ni (s, x) \mapsto (\sigma_v(s, x_v) \sigma_v^\top(s, x_v))^{-1} b_v(s, x_v, x_{N_v(G)}) \in \mathbb{R}^d. \tag{3.3}$$

Also, given any continuous local martingale M, we let $\mathcal{E}(M)$ denote the stochastic exponential

$$\mathcal{E}_t(M) := \exp\left(M(t) - \frac{1}{2}[M](t) \right), \qquad t \geq 0, \tag{3.4}$$

where $[M]$ denotes the (optional) quadratic variation process of M.

Let $(f_K)_{K\in \mathrm{cl}_2(G)}$ be as in Assumption (**A**.1). For each $t > 0$, define the measure $P_t \in \mathcal{P}(\mathcal{C}_t^V)$ by

$$\frac{dP_t}{dP_t^*} := \frac{d\mu_0}{d\mu_0^*}(X_V(0))\mathcal{E}_t\left(\sum_{v\in V} M_v\right),$$
$$= \prod_{K\in \mathrm{cl}_2(G)} f_K(X_K(0)) \prod_{v\in V} \mathcal{E}_t(M_v), \tag{3.5}$$

with $\mathcal{E}(M)$ and P^* as defined in the previous paragraph. Note that $W_v := \int_0^{\cdot} \sigma_v^{-1}(s, X_v)\,dX_v(s)$, $v \in V$, are independent d-dimensional Brownian motions under P^* by Remark 2.5. Therefore the stochastic exponentials appearing in (3.5) are true P^*-martingales due to the form of M_v in (3.2), the linear growth assumption (**A**.2) on the drifts and the non-degeneracy of σ_v; see Lemma B.1 with $\mathbb{Q} = P^*$, $X = (X_v)_{v\in V}$ and $f_v(t,x) = \sigma_v^{-1}b_v(t, x_v, x_{N_v(G)})$, $v \in V$. Further, observe that $(M_v)_{v\in V}$ are orthogonal under P^*. So Girsanov's theorem [13, Corollary 3.5.2] implies that under P_t, $\tilde{W}_v := W_v - \int_0^{\cdot} \sigma_v^{-1}(s, X_v)b_v(s, X_v, X_{N_v(G)})\,ds$, $v \in V$, are independent d-dimensional standard Brownian motions on $[0, t]$. From this it follows that under P_t, X solves the SDE (2.1) on $[0, t]$, and the same argument also shows that the restriction to $[0, t]$ of any solution to (2.1) must have law P_t on $\mathcal{C}_t^V$. Thus, we have uniqueness in law. Weak existence follows from Kolmogorov's extension theorem [12, Theorem 6.16] on observing that $\{P_t, t \geq 0\}$ form a consistent family in the sense that P_s is the restriction of P_t to $\mathcal{C}_s^V$ for each $t > s > 0$ (due to the martingale property of $\frac{dP_t}{dP_t^*}$). This completes the proof of Proposition 2.3.

On the other hand, the fact that for each $t > 0$, $(X_v[t])_{v\in V}$ is a 2MRF on C_t^V follows from (3.5) on applying Proposition 3.2 with $\mathcal{X} = \mathcal{C}_t$ and $\mu^* = P_t^*$, noting that P_t^* is a product measure on $\mathcal{C}_t^V$ and that, for each $v \in V$, $\{v\} \cup N_v(G)$ is a 2-clique and M_v of (3.2) is $X_{\{v\}\cup N_v(G)}$-measurable. This proves the first assertion of Theorem 2.4. For the second assertion of Theorem 2.4, denote by $P = P^{\mu_0} \in \mathcal{P}(\mathcal{C}^V)$ the law of the unique solution of the SDE system (2.1) with initial law μ_0. Fix a finite set $A \subset V$ and bounded continuous functions f, g, h on $\mathcal{C}^A, \mathcal{C}^{\partial^2 A}, \mathcal{C}^{V\setminus(A\cup\partial^2 A)}$, respectively. Fix $t > 0$ and let $\mathcal{F}_t := \sigma\{X_{\partial^2 A}[t]\}$ and $\mathcal{F}_\infty := \sigma\{X_{\partial^2 A}\}$. Below, with some abuse of notation, for any $B \subset V$, we will also interpret elements $y \in \mathcal{C}_t^B$ as elements of $\mathcal{C}^B$ by simply setting $y(s) = y(t)$ for $s \geq t$. Note that with this identification, for any $x \in \mathcal{C}^V$ and $B \subset V$, $x_B[t] \to x_B$ in $\mathcal{C}^B$ as $t \to \infty$. Then, noting that $\sigma(\cup_{t>0}\mathcal{F}_t) = \mathcal{F}_\infty$, invoking the martingale convergence theorem (in the third equality below), and using the fact that $P_t = P \circ (X_V[t])^{-1}$ is a 2MRF on $\mathcal{C}_t^V$ for each t (in the second equality below), we have

$$\mathbb{E}^P\left[f(X_A)g(X_{\partial^2 A})h(X_{V\setminus(A\cup\partial^2 A)})\right]$$
$$= \lim_{s\to\infty}\lim_{t\to\infty}\mathbb{E}^P\left[f(X_A[s])g(X_{\partial^2 A}[t])h(X_{V\setminus(A\cup\partial^2 A)}[s])\right]$$
$$= \lim_{s\to\infty}\lim_{t\to\infty}\mathbb{E}^P\left[\mathbb{E}^P\left[f(X_A[s])\mid \mathcal{F}_t\right]g(X_{\partial^2 A}[t])\mathbb{E}^P\left[h(X_{V\setminus(A\cup\partial^2 A)}[s])\mid \mathcal{F}_t\right]\right]$$
$$= \lim_{s\to\infty}\mathbb{E}^P\left[\mathbb{E}^P\left[f(X_A[s])\mid \mathcal{F}_\infty\right]g(X_{\partial^2 A})\mathbb{E}^P\left[h(X_{V\setminus(A\cup\partial^2 A)}[s])\mid \mathcal{F}_\infty\right]\right]$$
$$= \mathbb{E}^P\left[\mathbb{E}^P\left[f(X_A)\mid \mathcal{F}_\infty\right]g(X_{\partial^2 A})\mathbb{E}^P\left[h(X_{V\setminus(A\cup\partial^2 A)})\mid \mathcal{F}_\infty\right]\right], \tag{3.6}$$

where we have also made repeated use of the boundedness and continuity of f, g, h and the bounded convergence theorem. This shows that X_A and $X_{V\setminus(A\cup\partial^2 A)}$ are conditionally independent given $X_{\partial^2 A}$ under P, that is, P is a 2MRF on $\mathcal{C}^V$. This completes the proof. □

3.3. **Illustrative examples.** We now provide examples to show that the 2MRF property cannot in general be strengthened.

3.3.1. *The failure of the first-order MRF property for trajectories.* In general, P_t fails to be a first-order Markov random field on $\mathcal{C}_t^V$ for any $t > 0$, even if the initial states are i.i.d. To see why, notice that the density dP_t/dP_t^* given by (3.5) does not in general admit a clique factorization. Indeed, for $v \in V$ and $t > 0$, we recall the definition of M_v from (3.2) and $\mathcal{E}_t(M_v)$ from (3.4), which we write in full as

$$\mathcal{E}_t(M_v) = \exp\left(\int_0^t (\sigma_v \sigma_v^\top)^{-1} b_v(s, X_v, X_{N_v(G)}) \cdot dX_v(s) \right.$$
$$\left. - \frac{1}{2} \int_0^t \left\langle b_v(s, X_v, X_{N_v(G)}), (\sigma_v \sigma_v^\top)^{-1} b_v(s, X_v, X_{N_v(G)}) \right\rangle ds \right).$$

Noting that $\{v\} \cup N_v(G)$ is a 2-clique but not a 1-clique, this reveals why one cannot hope for a factorization over 1-cliques. For example, consider the "nice" case where $\sigma_v \equiv I$ and $b_v(s, x_v, x_{N_v(G)}) = \sum_{u \in N_v(G)} (x_u(s) - x_v(s))$. (Equivalently, $b_v(s, x_v, x_{N_v(G)}) = \nabla_{x_v} h(x)$ is of gradient-type with potential $h(x) = -\frac{1}{2} \sum_{(u,v) \in E} |x_u - x_v|^2$, where E is the edge set of G.) Then the first term in the above exponential splits nicely into a sum of pairwise interactions $\sum_{u \in N_v(G)} \int_0^t (X_u(s) - X_v(s)) \cdot dX_v(s)$, but the second term becomes

$$-\frac{1}{2} \sum_{u,w \in N_v(G)} \int_0^t \langle X_u(s) - X_v(s), X_w(s) - X_v(s) \rangle ds.$$

It is this term which fails to factorize further over 1-cliques as opposed to 2-cliques and thus precludes the first-order Markov property whenever $d\mu_0/d\mu_0^*$ is strictly positive due to Proposition 3.1.

To informally provide a different (but arguably more intuitive) perspective on why the first-order Markov property for past histories fails, consider the case when G is a line segment of length $\ell = 3$, labelling the vertices $-1, 0, 1$. Then, although the driving Brownian motions are all independent and the dynamics of each of the two extreme vertices only depend on its own state and the state of the center vertex, at any time t, conditioning on the past history of the states of the center vertex, does not make $X_{-1}(t)$ independent of $X_1(t)$ because the conditioning correlates the Brownian motions W_{-1} and W_1 on the interval $[0, t]$. This happens because the past history of X_0 is influenced by both W_{-1} and W_1 via X_{-1} and X_1. On the other hand, to see why the 2MRF property nevertheless does hold, note that if G were a line segment of length 4, labeling the vertices $\{-2, -1, 1, 2\}$, then conditioning on the history of the states of the two center vertices -1 and 1 no longer correlates the Brownian motions W_{-2} and W_2 since the dynamics of each of the conditioned vertices depends on a different driving Brownian motion. Thus, although the conditioning changes the distribution of W_{-2} and W_2 (for instance, they need no longer be Brownian motions), they remain independent, and hence $X_{-2}(t)$ is conditionally independent of $X_2(t)$ in this case.

Remark 3.5. There are certain situations in which P_t is, in fact, a 1MRF for each $t > 0$ (even though we know from the above examples that this is not in general the case). For example, suppose that for every $v \in V$, there exists a clique K_v of G with $v \in K_v \subset N_v(G)$ such that $b_v(t, x_v, x_{N_v(G)}) = \widetilde{b}_v(t, x_v, x_{K_v})$ depends on $x_{N_v(G)}$ only through x_{K_v}. Suppose also that $d\mu_0/d\mu_0^*$ admits a 1-clique factorization. Then, recalling (3.5), note that for each v the martingale M_v is measurable with respect to X_{K_v}, and deduce from Proposition 3.1 that P_t is a first-order Markov random field. For a concrete example that has the above form, consider the case when G is a triangular lattice with $V = \{\emptyset, 0, 1, \ldots, m\}$, for some $m \in \mathbb{N}$, with the central vertex $\emptyset$ having the neighborhood $N_\emptyset(G) = \{0, \ldots, m\}$ and for each $v \in V \setminus \{\emptyset\}$, $N_v(G) := \{\emptyset, v+1, v-1\}$, where the

vertices are to be interpreted mod $m + 1$. Further, suppose the initial conditions are i.i.d. and that for some $c \in \mathbb{R}$, $b_v(t, x_v, x_{N_v(G)}) = c(x_\emptyset + x_{v+1})$ for $v \in V \setminus \{\emptyset\}$ and $b_\emptyset(t, x_\emptyset, x_{N_\emptyset(G)}) = cx_\emptyset$. Then this provides a specific example with $K_v = \{\emptyset, v+1\} \subset N_v(G)$ for $v \in V \setminus \{\emptyset\}$ and $K_\emptyset = \emptyset$. In a similar spirit, the directed cycle graph model of [9] provides another example.

3.3.2. *The failure of MRF properties for time-t marginals.* It is natural to wonder if and when the time-t marginals $P_t \circ X(t)^{-1} \in \mathcal{P}((\mathbb{R}^d)^V)$ remain a first- or second-order Markov random field, given that this property is true at time 0, or even given i.i.d. initial conditions. This question is related to propagation of Gibbsianness and Gibbs-non-Gibbs transitions that have been studied in the literature, which is discussed in greater detail in Remark 3.6.

Here, we provide a simple example where both the first-order and second-order Markov property fail for time-t marginals. In fact, in this simple model we will see that there is no non-trivial conditional independence structure. Consider the segment with 5 vertices: $G = (V, E)$ given by $V = \{1, 2, 3, 4, 5\}$ and $E = \{(i, i+1) : i = 1, 2, 3, 4\}$, and consider the SDE system

$$
\begin{aligned}
dX_1(t) &= (X_2(t) - 2X_1(t)) \, dt + dW_1(t), \\
dX_i(t) &= (X_{i-1}(t) + X_{i+1}(t) - 2X_i(t)) \, dt + dW_i(t), \qquad i = 2, 3, 4, \\
dX_5(t) &= (X_4(t) - 2X_5(t)) \, dt + dW_5(t),
\end{aligned}
\tag{3.7}
$$

with $X_i(0) = 0$ for each i. Once again, note that the drift here is of gradient type with potential $h(x) = \sum_{i=1}^4 x_i x_{i+1} - \sum_{i=1}^5 x_i^2$. Letting $\boldsymbol{X}(t)$ denote the column vector $(X_1(t), \ldots, X_5(t))$ and similarly for $\boldsymbol{W}(t)$, we may write this in vector form as

$$
d\boldsymbol{X}(t) = L\boldsymbol{X}(t) \, dt + d\boldsymbol{W}(t),
\tag{3.8}
$$

where $L = A - 2I$ is the adjacency matrix A of the graph minus twice the identity I:

$$
L = \begin{pmatrix}
-2 & 1 & 0 & 0 & 0 \\
1 & -2 & 1 & 0 & 0 \\
0 & 1 & -2 & 1 & 0 \\
0 & 0 & 1 & -2 & 1 \\
0 & 0 & 0 & 1 & -2
\end{pmatrix}.
$$

The solution of the SDE (3.8) is given by

$$
\boldsymbol{X}(t) = e^{Lt} \int_0^t e^{-Ls} \, d\boldsymbol{W}(s), \qquad t > 0.
$$

Noting that L is symmetric and invertible, we deduce that $\boldsymbol{X}(t)$ is jointly Gaussian with mean zero and covariance matrix

$$
\mathbb{E}[\boldsymbol{X}(t)\boldsymbol{X}(t)^\top] = \int_0^t e^{2Ls} \, ds = \frac{1}{2}L^{-1}(e^{2Lt} - I).
\tag{3.9}
$$

This covariance matrix can easily be computed explicitly by noting that the tridiagonal Toeplitz matrix A is explicitly diagonalizable. To spare the reader any tedium, we provide only some pertinent snapshots. At time $t = 2$ the covariance matrix is

$$
\mathbb{E}[\boldsymbol{X}(2)\boldsymbol{X}(2)^\top] = \begin{pmatrix}
0.3611 & 0.2388 & 0.1435 & 0.0767 & 0.0324 \\
0.2388 & 0.5046 & 0.3156 & 0.1759 & 0.0767 \\
0.1435 & 0.3156 & 0.5370 & 0.3156 & 0.1435 \\
0.0767 & 0.1759 & 0.3156 & 0.5046 & 0.2388 \\
0.0324 & 0.0767 & 0.1435 & 0.2388 & 0.3611
\end{pmatrix}.
\tag{3.10}
$$

Using the well known formula for conditional measures of joint Gaussians, we compute from this that

$$\text{Cov}(X_1(t), X_3(t)|X_2(t)) = \begin{pmatrix} 0.2481 & -0.0058 \\ -0.0058 & 0.3397 \end{pmatrix},$$

which reveals that $X_1(t)$ and $X_3(t)$ are not conditionally independent given $X_2(t)$. Hence, $(X_i(t))_{i \in G}$ is not a first-order Markov random field. Similarly, by computing

$$\text{Cov}(X_1(t), X_4(t)|X_2(t), X_3(t)) = \begin{pmatrix} 0.2480 & -0.0030 \\ -0.0030 & 0.3189 \end{pmatrix},$$

we see that $X_1(t)$ and $X_4(t)$ are not conditionally independent given $(X_2(t), X_3(t))$. Hence, $(X_i(t))_{i \in G}$ is not a second-order Markov random field.

In fact, in this example, there is no non-trivial conditional independence structure, in the sense that there are no two vertices i, j such that $X_i(t)$ and $X_j(t)$ are conditionally independent given $\{X_k(t) : k \in G \setminus \{i, j\}\}$ for some $t > 0$. This can be read off from the the the so-called *precision matrix*, which is simply the inverse of the covariance matrix, $Q(t) := (\mathbb{E}[\boldsymbol{X}(t)\boldsymbol{X}(t)^\top])^{-1}$. As is well known and can easily be seen from the form of the multivariate Gaussian density, the precision matrix reveals the conditional independence structure (see, e.g., [18, Proposition 5.2]), in the following sense: For $t > 0$ define the graph $\widetilde{G}(t) = (V, \widetilde{E}(t))$ with the same vertex set V but with $(i, j) \in \widetilde{E}(t)$ if and only if $Q_{i,j}(t) \neq 0$. Then $\boldsymbol{X}(t)$ is a (first-order) Markov random field with respect to the graph $\widetilde{G}(t)$. In our example, $\widetilde{G}(t)$ is the complete graph for each $t > 0$, and this Markov property is vacuous. (Note, however, that $Q(t) \to 2L$ as $t \to \infty$ because L is negative definite, and the unique invariant measure of this diffusion is a first-order Markov random field with respect to the original graph G.)

A variation on this example gives rise to another interesting phenomenon. Suppose we modify the example by replacing the diagonal entries of L with zeros, i.e., remove all the $-2X$ terms from the drifts in (3.7). Then the covariance matrix is again invertible, and now $Q_{1,4}(t) = Q_{2,5}(t) = 0$ for all $t > 0$, where we continue with the notation of the previous paragraph. That is, $\widetilde{G}(t)$ is not the complete graph, but rather the complete graph with the edges $(1, 4)$ and $(2, 5)$ removed, for each $t > 0$. In particular, $X_1(t)$ and $X_4(t)$ are conditionally independent given $(X_2(t), X_3(t), X_5(t))$, for each $t > 0$.

Remark 3.6. As mentioned in the introduction, one motivation for studying such conditional independence questions is that (a stronger version of) the MRF structure of interacting SDEs can lead to an autonomous "local characterization" of the dynamics at a vertex and its neighborhood, as developed in the quite different setting of unimodular Galton-Watson trees in [17]. From this perspective, it would be of interest to investigate if there are non-trivial special cases when the first-order or second-order MRF property for time-t marginals propagates. A different but related question that has been studied in the literature is propagation of Gibbsianness for an infinite system of interacting real-valued diffusions indexed by $\mathbb{Z}^d$. Specifically, the work [7] considers a collection of interacting diffusions, indexed by $\mathbb{Z}^d$, with identity covariance and a drift that is the gradient of a Hamiltonian function associated with a certain interaction potential Φ, and with an initial distribution that is also a Gibbs measure (as in Section 2.3) with respect to a Gibbsian specification (in the sense of [11, Chapter 2]) associated with another interaction potential Φ_0, where both interaction potentials Φ and Φ_0 are assumed to be of finite range and satisfy certain smoothness conditions. It is shown in [7] that when either t or the interaction strength is sufficiently small, the time-t marginals are strongly Gibbsian, that is, associated with Gibbsian specifications that have an absolutely summable, though not necessarily finite range, interaction potential. Extensions of these results to the case of interacting real-valued diffusions

on $\mathbb{Z}^d$ with non-Markovian drifts with finite memory (again with finite range interactions and identity covariance) were later obtained in [24] and [25]. The restrictions on the time and interaction strength in these works arise from the fact that perturbative arguments are used. However, in general for moderate interaction strengths and moderate times, the time-t marginals can fail to be Gibbsian (see, e.g., [33], as well as the survey [32], which also discusses related results for spin systems).

4. Finite-graph approximations for Markov random fields

In this section we establish some important preparatory results that are used in the proof of Theorem 2.7, which extends the finite graph results of Theorem 2.4 to the infinite graph setting. Fix (G, b, σ, μ_0) that satisfy Assumption **B** and suppose $G = (V, E)$ is countably infinite. Recall that $P = P^{\mu_0} \in \mathcal{P}(\mathcal{C}^V)$ and $P^{*,\mu_0} \in \mathcal{P}(\mathcal{C}^V)$ denote the unique law of the SDE systems (2.1) and (2.5), respectively, both with initial laws μ_0, which are well-posed by Assumptions (**B**.4) and (**B**.3). To show that $P_t = P_t^{\mu_0}$ forms a 2MRF on $\mathcal{C}_t^V$, we can no longer apply the clique factorization arguments used for finite graphs because the formula (3.5) does not extend to infinite graphs. Even worse, the density dP_t/dP_t^* therein does not exist, and it seems impossible to establish directly that by projecting to a finite set $A \subset V$ we have a density $dP_t[A]/dP_t^*[A]$ that admits a 2-clique factorization. Instead, we approximate the measure on the infinite graph by 2MRFs on a growing sequence of finite graphs, arguing that the desired 2MRF property passes to the limit. To highlight some of the subtleties that arise in such an approximation argument, and to better motivate the other results established in this section, we first desribe the approximating sequence of measures in Section 4.1. Then in the subsequent two sections we establish some general properties of finite-graph 2MRFs to be used in the proof of Theorem 2.7 in Section 5.2, which are also of independent interest.

4.1. **Construction of the approximating sequence of SDEs.** We fix (G, b, σ, μ_0) that satisfy Assumption **B**. As in Section 3.2, we will work with the canonical measure space $\mathcal{C}^V = (\mathcal{C}^V, \mathcal{B}(\mathcal{C}^V), P^{*,\mu_0})$, and let $(X_v)_{v \in V} : \mathcal{C}^V \to \mathcal{C}^V$ again denote the canonical processes. Also, recall from Section 1.2 that given any measurable space $\mathcal{X}$, measure $\mu \in \mathcal{P}(\mathcal{X}^V)$ and subset $U \subset V$, $\mu[U] \in \mathcal{P}(\mathcal{X}^U)$ denotes the restriction of μ to the set $\mathcal{X}^U$.

Let $(V_n)_{n \in \mathbb{N}}$ be an arbitrary sequence of subsets of V such that $\bigcup_n V_n = V$, and let $G_n = (V_n, E_n)$, for some edge set E_n to be specified later. Also, for each $n \in \mathbb{N}$ and $v \in V_n$, let $b_v^n : \mathbb{R}_+ \times \mathcal{C} \times \mathcal{C}^{N_v(G_n)} \mapsto \mathbb{R}^d$ be any progressively measurable map that satisfies the same conditions as b_v in Assumption (**B**.2). Fix $t > 0$ and for each n, define $P_t^n = P_t^{\mu_0,n} \in \mathcal{P}(\mathcal{C}_t^V)$ by

$$\frac{dP_t^n}{dP_t^{*,\mu_0}} = \frac{dP_t^{\mu_0,n}}{dP_t^{*,\mu_0}} = \prod_{v \in V_n} \mathcal{E}_t \left(\int_0^{\cdot} (\sigma_v \sigma_v^\top)^{-1} b_v^n(s, X_v, X_{N_v(G_n)}) \cdot dX_v(s) \right), \tag{4.1}$$

where, as before, $(\sigma_v \sigma_v^\top)^{-1} b_v^n(s, x, x_{N_v(G_n)})$ denotes the map (3.3). We can apply Lemma B.1 with $\mathbb{Q} = P^{*,\mu_0}$, $X = (X_v)_{v \in V_n}$ and $f(t, x) = (\sigma_v^{-1} b_v^n(t, x_v, x_{N_v(G_n)}))_{v \in V_n}$, to conclude that the stochastic exponential in (4.1) is a true P^{*,μ_0}-martingale, due to the linear growth, non-degeneracy and boundedness properties of b_v^n and σ_v in Assumptions (**B**.2) and (**B**.3). Hence, the family $(P_t^n)_{t>0}$ is consistent in the sense that the restriction of P_t^n to $\mathcal{C}_s^V$ is precisely P_s^n for each $t > s > 0$. Thus, by the Kolmogorov extension theorem, $(P_t^n)_{t>0}$ uniquely determines a probability measure P^n on $\mathcal{C}^V$. Now, from (4.1), (2.5) and Girsanov's theorem [13, Corollary 3.5.2], it follows that under P^n the canonical process solves the SDE system

$$\begin{aligned} dX_v(t) &= b_v^n(t, X_v, X_{N_v(G_n)}) \, dt + \sigma_v(t, X_v) \, dW_v(t), \quad v \in V_n, \\ dX_v(t) &= \sigma_v(t, X_v) \, dW_v(t), \quad v \in V \setminus V_n, \end{aligned} \tag{4.2}$$

with $(X_v(0))_{v \in V} \sim \mu_0$, where $(W_v)_{v \in V}$ are independent Brownian motions under P^n. Note that for $v \in V_n$, the third argument of b_v^n looks only at the states in $N_v(G_n)$, and thus b_v^n depends only on the states of vertices in G_n. Thus, $P^n[V_n]$ is precisely the law of the finite-graph SDE system (2.1) with inputs $(G, (b_v^n)_{v \in V}, (\sigma_v)_{v \in V}, \mu_0[V_n])$.

In order to implement our approximation argument we would like to choose G_n and $(b_v^n)_{v \in V}$ such that both $P^n \to P$ and each $P_t^n[V_n]$ is a 2MRF. In order to have $P^n \to P$ we should naturally choose V_n increasing to V and b_v^n to behave like b_v for most v. But the 2MRF property is more delicate. It would follow from the finite-graph result of Theorem 2.4 that $P_t^n[V_n]$ is a 2MRF on $\mathcal{C}^{V_n}$ only if $\mu_0[V_n]$ were a 2MRF on $(\mathbb{R}^d)^{V_n}$. But $\mu_0[V_n]$ is not necessarily a 2MRF for arbitrary V_n (e.g., with G_n the induced subgraph), even though μ_0 is a 2MRF on the full graph G by assumption; in other words, the 2MRF property is not in general preserved under projections, as illustrated in Example 4.1 below. However, in Section 4.2 we show that for any Markov random field on an infinite graph $G = (V, E)$, it is possible to identify a suitable increasing sequence of vertices $(V_n)_{n \in \mathbb{N}}$ and associated graph $G_n = (V_n, E_n)$ for each $n \in \mathbb{N}$ that is a slight modification of the induced subgraph on V_n, such that the desired projection property holds. Then, in Section 4.3 we prove some results on preservation of a class of conditional distributions of 2MRFs under restriction to induced subgraphs. The above results are combined with tightness and convergence estimates for the approximating sequence $\{P^n\}_{n \in \mathbb{N}}$ obtained in Section 5.1 to complete the proof of Theorem 2.7 in Section 5.2.

4.2. **Projections of Markov random fields.** We first provide a simple example to illustrate that the restriction of an MRF to an induced subgraph need not remain an MRF.

Example 4.1. Suppose G is a finite two-dimensional lattice, with vertex set V identified with $\{-n, \ldots, n\}^2$ and the usual nearest-neighbor edge set, and let $(Y_v)_{v \in G}$ be a 1MRF on $\mathbb{R}^V$. Consider the line subgraph $H = \{(i, 0) : i = -n, \ldots, n\}$ in G, and consider the restriction $(Y_v)_{v \in H}$ of the 1MRF to H. Note that every path in H that starts in $A := \{(0, 0)\}$ and ends in $B = \{(2, 0)\}$ must traverse through the vertex of $S := \{(1, 0)\}$. Thus, by the cutset characterization of 1MRF's given in Remark 3.4, for $(Y_v)_{v \in H}$ to be an 1MRF on H, $Y_{(0,0)}$ must be conditionally independent of $Y_{(2,0)}$ given $Y_{(1,0)}$. However, by the same cutset characterization, it is clear that this conditional independence cannot be deduced from the 1MRF property of $(Y_v)_{v \in G}$ on G since there are paths in G that start in A and end in B that are disjoint from S. Similarly, if we assume $(Y_v)_{v \in G}$ is a 2MRF, the configuration $(Y_v)_{v \in H}$ can fail to be a 2MRF.

This example does suggest, however, that we can restore the MRF property by enlarging the edge set of the induced subgraph to reflect the lost connectivity. The following lemma gives one way to do this which is certainly not the only way, but it serves our purpose. For a random element $(Y_v)_{v \in V}$ of $\mathcal{X}^V$ with law $\nu \in \mathcal{P}(\mathcal{X}^V)$, and for a set $A \subset V$, recall that we write $\nu[A]$ to denote the law of Y_A, the coordinates in A.

Lemma 4.2. *Fix a rooted graph $G = (V, E, \varnothing)$ and $n \geq 4$. Define $V_n := \{v \in V : d(v, \varnothing) \leq n\}$ and $U_n := V_n \setminus V_{n-2}$, where d denotes the graph distance. Define a graph $G_n = (V_n, E_n)$, where*

$$E_n := \{(u, v) \in V_n \times V_n : (u, v) \in E\} \cup \{(u, v) \in U_n \times U_n, u \neq v\}.$$

(i) For any $A \subset V_{n-3}$, it holds that $\partial_G^2 A = \partial_{G_n}^2 A$. Also, for any $A' \subset V_{n-2}$, $\partial_G^2 A' \subset \partial_{G_n}^2 A'$.
(ii) If $K \in \mathrm{cl}_2(G)$ satisfies $K \subset V_n$, then $K \in \mathrm{cl}_2(G_n)$.
(iii) If a $\mathcal{X}^V$-valued random variable $(Y_v)_{v \in V}$ is a 2MRF with respect to G, then $(Y_v)_{v \in V_n}$ is a 2MRF with respect to G_n.

(iv) *Suppose V is finite and the law ν of Y_V admits the following 2-clique factorization with respect to a product measure $\nu^* = \prod_{v \in V} \theta_v \in \mathcal{P}(\mathcal{X}^V)$ for some $\theta_v \in \mathcal{P}(\mathcal{X})$,*

$$\frac{d\nu}{d\nu^*}(x_V) = \prod_{K \in \mathrm{cl}_2(G)} f_K(x_K),$$

for some measurable functions $f_K : \mathcal{X}^K \to \mathbb{R}_+$, for $K \in \mathrm{cl}_2(G)$. Then $\nu[V_n]$ admits a 2-clique factorization

$$\frac{d\nu[V_n]}{d\nu^*[V_n]}(x_{V_n}) = \prod_{K \in \mathrm{cl}_2(G_n)} f_K^0(x_K),$$

for some measurable functions $f_K^0 : \mathcal{X}^K \to \mathbb{R}_+$, for $K \in \mathrm{cl}_2(G_n)$, which additionally satisfy the consistency condition $f_K^0 \equiv f_K$ for $K \in \mathrm{cl}_2(G)$ such that $K \subset V_{n-3}$.

Proof.

(i) From the definition of E_n it follows quickly that (a) for $A' \subset V_{n-2}$, $\partial_G A' = \partial_{G_n} A'$, and (b) for $A'' \subset V_{n-1}$, $\partial_G A'' \subset \partial_{G_n} A''$. Iterate these observations to prove the claims.

(ii) Let d_G and d_{G_n} denote the graph distance in G and G_n, repsectively. From the definition of E_n, it is straightforward to argue that $d_{G_n} \leq d_G$ on $V_n \times V_n$. Indeed, for any $u, v \in V_n$ and any path from u to v in G, there is a path from u to v in G_n which is not longer. This implies for every $u, v \in V_n$, $d_G(u, v) \leq 2$ implies $d_{G_n}(u, v) \leq 2$, which proves property (ii).

(iii) Let $(Y_v)_{v \in V}$ be a 2MRF with respect to G. Let $A \subset V_n$, $B = V_n \setminus (A \cup \partial_{G_n}^2 A)$, and $S := \partial_{G_n}^2 A$. Assuming without loss of generality that A and B are nonempty, we must prove that Y_A and Y_B are conditionally independent given Y_S. First notice that one cannot have both $A \cap U_n \neq \emptyset$ and $B \cap U_n \neq \emptyset$, as this would imply $d_{G_n}(A, B) \leq 1$, contradicting the definition of B. Therefore we must have either $A \cap U_n = \emptyset$, $B \cap U_n = \emptyset$, or both.

Case 1: Suppose $A \cap U_n = \emptyset$. This means $A \subset V_{n-2}$ and hence $\partial_G^2 A \subset S$ by (i). Since Y_A and $Y_{V \setminus (A \cup \partial_G^2 A)}$ are conditionally independent given $Y_{\partial_G^2 A}$, we then have conditional independence of Y_A and Y_B given Y_S. Indeed, this uses the elementary fact that if (Z_1, Z_2, Z_3, Z_4) are random variables with Z_1 conditionally independent of (Z_2, Z_3) given Z_4, then Z_1 is conditionally independent of Z_2 given (Z_3, Z_4).

Case 2: Suppose $B \cap U_n = \emptyset$. This means $B \subset V_{n-2}$ and hence, again by (i), $\partial_G^2 B \subset \partial_{G_n}^2 B$. Also note that $\partial_{G_n}^2 B \subset S$ (since otherwise $A \cap \partial_{G_n}^2 B \neq \emptyset$, which contradicts the definition of B). Since the 2MRF property with respect to G implies Y_B and $Y_{V \setminus (B \cup \partial_G^2 B)}$ are independent conditioned on $Y_{\partial_G^2 B}$, we then have conditional independence of Y_B and Y_A given Y_S. Since $A \subset V_n$ was arbitrary, this proves that $(Y_v)_{v \in V_n}$ is a 2MRF with respect to G_n.

(iv) Let $\mathcal{K}_n$ denote the set of $K \in \mathrm{cl}_2(G)$ such that $K \subset V_n$. Recalling that ν^* is a product measure, using the assumed clique factorization of ν, we can then write

$$\frac{d\nu[V_n]}{d\nu^*[V_n]}(x_{V_n}) = \int_{\mathcal{X}^{V \setminus V_n}} \prod_{K \in \mathrm{cl}_2(G)} f_K(x_K)\, \nu^*[V \setminus V_n](dx_{V \setminus V_n})$$

$$= \prod_{K \in \mathcal{K}_n} f_K(x_K) \int_{\mathcal{X}^{V \setminus V_n}} \prod_{K \in \mathrm{cl}_2(G) \setminus \mathcal{K}_n} f_K(x_K)\, \nu^*[V \setminus V_n](dx_{V \setminus V_n}).$$

Now note that any $K \in \mathrm{cl}_2(G) \setminus \mathcal{K}_n$ is not contained in V_n, and as a 2-clique it can have no neighbors in V_{n-2}. Recalling that $U_n = V_n \setminus V_{n-2}$, we see that the integral expression is

x_{U_n}-measurable; that is, there is a measurable function $g_n : \mathcal{X}^{U_n} \to \mathbb{R}_+$ such that

$$g_n(x_{U_n}) = \int_{\mathcal{X}^{V \setminus V_n}} \prod_{K \in \mathrm{cl}_2(G) \setminus \mathcal{K}_n} f_K(x_K)\, \nu^*[V \setminus V_n](dx_{V \setminus V_n}).$$

Note that $U_n \in \mathrm{cl}_2(G_n)$ by definition of G_n. Since clearly $\mathcal{K}_n \subset \mathrm{cl}_2(G_n)$, we find that the expression

$$\frac{d\nu[V_n]}{d\nu^*[V_n]}(x_{V_n}) = \prod_{K \in \mathcal{K}_n} f_K(x_K) g_n(x_{U_n})$$

exhibits a 2-clique factorization of $\nu[V_n]$ over the graph G_n satisfying the desired consistency condition.

<div align="right">□</div>

4.3. Conditional distributions of second-order Markov random fields. First, in Lemma 4.3, given a 2MRF with respect to a graph, and another 2MRF on a subgraph, or more generally given MRFs on two overlapping graphs, we identify conditions under which the conditional distributions of a subset in the intersection (given its complement) coincide for both 2MRFs. This will be used to establish, for a suitable choice of b^n, a certain consistency condition for the sequence of approximating measures $\{P^n\}_{n \in \mathbb{N}}$ used in the proof of Theorem 2.7. Let us briefly recall a notation we introduced more carefully just before Theorem 2.9: For $\nu \in \mathcal{P}(\mathcal{X}^V)$ and $A, B \subset V$ we write $\nu[A \,|\, B]$ for the conditional law of the A-coordinates given the B-coordinates.

Lemma 4.3. *Let $G = (V_G, E_G)$ and $H = (V_H, E_H)$ be finite graphs, and assume $V^* \subset V_G \cap V_H$ satisfies*

$$E_G \cap (V^* \times V^*) = E_H \cap (V^* \times V^*). \tag{4.3}$$

Moreover, let $A \subset V^*$ satisfy $\partial_G^2 A \subset V^*$ and $\partial_H^2 A \subset V^*$. Then $\partial_H^2 A = \partial_G^2 A =: \partial^2 A$, and it holds that

$$\{K \in \mathrm{cl}_2(G) : K \cap A \neq \emptyset\} = \{K \in \mathrm{cl}_2(H) : K \cap A \neq \emptyset\} =: \mathcal{K}_A. \tag{4.4}$$

Next, let $\nu^H \in \mathcal{P}(\mathcal{X}^{V_H})$ and $\nu^G \in \mathcal{P}(\mathcal{X}^{V_G})$, and suppose there exists a product measure $\nu^* = \prod_{v \in V_G \cup V_H} \theta_v \in \mathcal{P}(\mathcal{X}^{V_G \cup V_H})$ for some $\theta_v \in \mathcal{P}(\mathcal{X})$, $v \in V_G \cup V_H$, such that the densities factorize as

$$\frac{d\nu^H}{d\nu^*[V_H]}(x_{V_H}) = \prod_{K \in \mathrm{cl}_2(H)} f_K^H(x_K), \qquad \frac{d\nu^G}{d\nu^*[V_G]}(x_{V_G}) = \prod_{K \in \mathrm{cl}_2(G)} f_K^G(x_K),$$

for measurable functions $(f_K^H : \mathcal{X}^K \mapsto \mathbb{R}_+)_{K \in \mathrm{cl}_2(H)}$ and $(f_K^G : \mathcal{X}^K \mapsto \mathbb{R}_+)_{K \in \mathrm{cl}_2(G)}$ satisfying $f_K^H \equiv f_K^G$ for all $K \in \mathcal{K}_A$. Then $\nu^H[A \,|\, \partial^2 A] = \nu^G[A \,|\, \partial^2 A]$, almost surely with respect to $\nu^*[\partial^2 A]$.

Note that Lemma 4.3 is vacuously true if V^* is empty.

Proof of Lemma 4.3. Let $A \subset V^*$ satisfy $\partial_G^2 A \subset V^*$ and $\partial_H^2 A \subset V^*$. It is immediate from (4.3) that $\partial_H^2 A = \partial_G^2 A$, and we write simply $\partial^2 A$ for this set. To check (4.4), note that if $K \in \mathrm{cl}_2(G)$ intersects A, then $K \subset A \cup \partial^2 A \subset V^*$. By (4.3) the edge sets of G and H agree when restricted to V^*, and we deduce that $K \in \mathrm{cl}_2(H)$; this proves $\subset$ in (4.4), but the reverse inclusion follows by the same argument. Note that, with $\mathcal{K}_A$ as defined in (4.4), we have also shown that

$$K \in \mathcal{K}_A \Rightarrow K \subset V^*. \tag{4.5}$$

Let us work in the rest of the proof on the canonical probability space $(\mathcal{X}^{V_G \cup V_H}, \mathcal{B}(\mathcal{X}^{V_G \cup V_H}), \nu^*)$, with $\mathbb{E}$ denoting expectation on this space, and all equations are understood to hold ν^*-almost

surely. Let $I = (I_v)_{v \in V_G \cup V_H}$ denote the identity map on $\mathcal{X}^{V_G \cup V_H}$. By Proposition 3.2, ν^H is a 2MRF, and so $\nu^H[A \,|\, \partial^2 A](I_{\partial^2 A}) = \nu^H[A \,|\, V_H \backslash A](I_{V_H \backslash A})$ a.s. Hence,

$$\frac{d\nu^H[A \,|\, \partial^2 A](I_{\partial^2 A})}{d\nu^*[A]}(I_A) = \frac{\frac{d\nu^H}{d\nu^*[V_H]}(I_{V_H})}{\mathbb{E}\left[\frac{d\nu^H}{d\nu^*[V_H]}(I_{V_H}) \,|\, I_{V_H \backslash A}\right]}$$

$$= \frac{\prod_{K \in \mathrm{cl}_2(H)} f_K^H(I_K)}{\mathbb{E}[\prod_{K \in \mathrm{cl}_2(H)} f_K^H(I_K) \,|\, I_{V_H \backslash A}]},$$

where we emphasize that the expectation in the denominator is with respect to independent random variables $(I_v)_{v \in V_G \cup V_H}$.

The key observation is that if $K \in \mathrm{cl}_2(H)$ does not intersect A, then the term $f_K^H(I_K)$ factors out of the conditional expectation and cancels. Hence, with $\mathcal{K}_A$ as in (4.4), we see that

$$\frac{d\nu^H[A \,|\, \partial^2 A](I_{\partial^2 A})}{d\nu^*[A]}(I_A) = \frac{\prod_{K \in \mathcal{K}_A} f_K^H(I_K)}{\mathbb{E}[\prod_{K \in \mathcal{K}_A} f_K^H(I_K) \,|\, I_{V_H \backslash A}]}. \tag{4.6}$$

Since $I_{V_H \backslash V^*}$ is independent of I_{V^*}, in view of (4.5), we may equivalently condition on $I_{V^* \backslash A}$ in the denominator of the term on the right-hand side of (4.6) to obtain

$$\frac{d\nu^H[A \,|\, \partial^2 A](I_{\partial^2 A})}{d\nu^*[A]}(I_A) = \frac{\prod_{K \in \mathcal{K}_A} f_K^H(I_K)}{\mathbb{E}[\prod_{K \in \mathcal{K}_A} f_K^H(I_K) \,|\, I_{V^* \backslash A}]}.$$

Repeating the same arguments that led us to this point, we also find that

$$\frac{d\nu^G[A \,|\, \partial^2 A](I_{\partial^2 A})}{d\nu^*[A]}(I_A) = \frac{\prod_{K \in \mathcal{K}_A} f_K^G(I_K)}{\mathbb{E}[\prod_{K \in \mathcal{K}_A} f_K^G(I_K) \,|\, I_{V^* \backslash A}]}.$$

Recalling that $f_K^H \equiv f_K^G$ for $K \in \mathcal{K}_A$ by assumption, the proof is complete. $\qquad\square$

The last lemma allows us to deduce the following insensitivity result that shows that given a finite graph $G = (V, E)$ and associated SDE (2.1), the conditional law of trajectories of particles in a set $A \subset V$ given the trajectories of particles at the double-boundary $\partial^2 A$ of the set does not depend on the graph structure outside of $A \cup \partial^2 A$.

Proposition 4.4. *Let $G = (V_G, E_G)$ and $H = (V_H, E_H)$ be finite graphs, and assume $V^* \subset V_G \cap V_H$ satisfies (4.3). Let $A \subset V^*$ satisfy $\partial_G^2 A \subset V^*$ and $\partial_H^2 A \subset V^*$, so that Lemma 4.3 ensures that $\partial_H^2 A = \partial_G^2 A =: \partial^2 A$ and that (4.4) holds (defining $\mathcal{K}_A$ as therein). Suppose $(G, b^G, \sigma^G, \mu_0^G)$ and $(H, b^H, \sigma^H, \mu_0^H)$ both satisfy Assumption **A**, and let $P^G \in \mathcal{P}(\mathcal{C}^{V_G})$ and $P^H \in \mathcal{P}(\mathcal{C}^{V_H})$ be the corresponding unique laws of the SDE described in (2.1). Further, suppose the following consistency conditions hold:*

(i) We have

$$b_v^H \equiv b_v^G, \quad \text{for } v \in A \cup \partial^2 A, \tag{4.7}$$

$$\sigma_v^H \equiv \sigma_v^G, \quad \text{for } v \in V_G \cap V_H. \tag{4.8}$$

(ii) There is a product measure $\mu_0^* = \prod_{v \in V_G \cup V_H} \lambda_v \in \mathcal{P}((\mathbb{R}^d)^{V_G \cup V_H})$ for some $\lambda_v \in \mathcal{P}(\mathbb{R}^d), v \in V_G \cup V_H$, such that μ_0^G and μ_0^H admit 2-clique factorizations:

$$\frac{d\mu_0^G}{d\mu_0^*[V_G]}(x_{V_G}) = \prod_{K \in \mathrm{cl}_2(G)} f_K^G(x_K), \qquad \frac{d\mu_0^H}{d\mu_0^*[V_H]}(x_{V_H}) = \prod_{K \in \mathrm{cl}_2(H)} f_K^H(x_K), \tag{4.9}$$

for some measurable functions $(f_K^G : (\mathbb{R}^d)^K \mapsto \mathbb{R}_+)_{K \in \mathrm{cl}_2(G)}$ and $(f_K^H : (\mathbb{R}^d)^K \mapsto \mathbb{R}_+)_{K \in \mathrm{cl}_2(H)}$ that satisfy the consistency condition $f_K^G \equiv f_K^H$ for every $K \in \mathcal{K}_A$.

Then $P_t^G[A \,|\, \partial^2 A] = P_t^H[A \,|\, \partial^2 A]$ for each $t > 0$, both in the sense of $P_t^H[\partial^2 A]$-almost sure and $P_t^G[\partial^2 A]$-almost sure equality.

Proof. As in (2.5), let $P^* \in \mathcal{P}(\mathcal{C}^{V_G \cup V_H})$ be the unique law of the solution $X = (X_v)_{v \in V_G \cup V_H}$ of the driftless SDE

$$dX_v(t) = \sigma_v^G(t, X_v)\, dW_v(t), \quad v \in V_G, \qquad dX_v(t) = \sigma_v^H(t, X_v)\, dW_v(t), \quad v \in V_H \setminus V_G,$$

initialized with $X(0) \sim \mu_0^*$. Again working on the canonical probability space $(\mathcal{C}^{V_G \cup V_H}, \mathcal{B}(\mathcal{C}^{V_G \cup V_H}), P^*)$, define the martingales $M^H = (M_v^H)_{v \in H}$:

$$M_v^H(t) := \int_0^t (\sigma_v^H (\sigma_v^H)^\top)^{-1} b_v^H(s, X_v, X_{N_v(H)}) \cdot dX_v(s), \quad v \in V_H,$$

with $M^G = (M_v^G)_{v \in V_G}$, defined analogously, as in (3.2). Using (3.5) and (4.9) we can write

$$\frac{dP_t^H}{dP_t^*[V_H]} = \prod_{K \in \mathrm{cl}_2(H)} f_K^H(X_K(0)) \prod_{v \in V_H} \mathcal{E}_t(M_v^H),$$

$$\frac{dP_t^G}{dP_t^*[V_G]} = \prod_{K \in \mathrm{cl}_2(G)} f_K^G(X_K(0)) \prod_{v \in V_G} \mathcal{E}_t(M_v^G).$$

with $\mathcal{E}_t$ defined as in (3.4). Note that if $v \in A \cup \partial A$, then we have $N_v(H) = N_v(G)$; indeed, this is due to (4.3) and the inclusions $\partial_G A \subset V^*$ and $\partial_H A \subset V^*$. Thus, by the consistency conditions (4.7) and (4.8) along with the expressions above for M^H and M^G, we have $\mathcal{E}_t(M_v^H) = \mathcal{E}_t(M_v^G)$ for $v \in A \cup \partial A$. Applying Lemma 4.3 with $\mathcal{X} = \mathcal{C}_t$, $\nu^* = P^*$, $\nu^H = P_t^H$ and $\nu^G = P_t^G$, it follows from the consistency conditions (i) and (ii) that $P_t^G[A \,|\, \partial^2 A] = P_t^H[A \,|\, \partial^2 A]$ holds in the sense of $P_t^*[\partial^2 A]$-almost sure equality. Since both $P_t^H[\partial^2 A]$ and $P_t^G[\partial^2 A]$ are absolutely continuous with respect to $P_t^*[\partial^2 A]$, the claim follows. $\square$

5. MARKOV RANDOM FIELD PROPERTY FOR INFINITE-DIMENSIONAL DIFFUSIONS

Fix a countably infinite connected graph $G = (V, E)$, and let (G, b, σ, μ_0) be as in Assumption **B**. As usual, let $P = P^{\mu_0}$ and $P_t = P_t^{\mu_0}$ denote the unique law of the SDE (2.1) and its projection, and let P^{*, μ_0} be the law of the canonical SDE system (2.5) started from initial law μ_0. In this section we will prove Theorem 2.7, that is, the 2MRF property for P_t and P. We will also use the same canonical space $(\mathcal{C}^V, \mathcal{B}(\mathcal{C}^V), P^{*, \mu_0})$ and canonical processes $(X_v)_{v \in V}$ as in Section 4.1.

Throughout, choose an arbitrary vertex $\o$ in V to be the root, and let $G_n = (V_n, E_n)$ and $U_n = V_n \setminus V_{n-2}$ be as defined in Lemma 4.2. Also, set

$$b_v^n = b_v, \quad v \in V_{n-2}, \qquad b_v^n = 0, \quad v \in U_n, \tag{5.1}$$

(The family $\{b_v^n : n \geq 3, v \in U_n\}$ is arbitrary and set to zero for convenience, but more generally must merely be measurable and uniformly bounded.) Let $\{P^n\}_{n \in \mathbb{N}}$ and $\{P_t^n\}_{n \in \mathbb{N}}$ be the corresponding approximating sequence of measures and its projections, as defined in Section 4.1. We first establish tightness and convergence results for $\{P_t^n\}_{n \in \mathbb{N}}$ in Section 5.1 and finally present the proof of Theorem 2.7 in Section 5.2.

5.1. **Tightness and convergence results.** In the following, let $H(\cdot \,|\, \cdot)$ denote relative entropy, defined for $\nu \ll \mu$ by

$$H(\nu|\mu) = \int \frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu}\, d\mu,$$

and $H(\nu|\mu) = \infty$ for $\nu \not\ll \mu$. Recall also our notation $\|x\|_{*,t} := \sup_{0 \leq s \leq t} |x_s|$ for the truncated supremum norm.

Lemma 5.1. *Suppose Assumption **B** holds. For each $t > 0$ and each finite set $A \subset V$, we have*

$$\sup_n \sup_{v \in V_n} \mathbb{E}^{P^n} \left[\|X_v\|_{*,t}^2 \right] < \infty, \tag{5.2}$$

$$\sup_n H \left(P_t^n[A] \,\big|\, P_t^{*,\mu_0}[A] \right) < \infty, \tag{5.3}$$

$$\sup_n H \left(P_t^{*,\mu_0}[A] \,\big|\, P_t^n[A] \right) < \infty. \tag{5.4}$$

Proof. Fix $t > 0$. We begin with a standard estimate. Recall the definition of b_v^n from (5.1), apply Itô's formula to the SDE (4.2), and use the linear growth of b_v from Assumption (**B**.2) along with the uniform boundedness of σ_v from Assumption (**B**.3) to conclude that, for each $n \in \mathbb{N}$ and $v \in V_n$,

$$\mathbb{E}^{P^n} \left[\|X_v\|_{*,t}^2 \right] \leq C \mathbb{E}^{P^n} \left[|X_v(0)|^2 + \int_0^t |b_v^n(s, X_v, X_{N_v(G)})|^2 \, ds + \int_0^t |\sigma_v(s, X_v(s))|^2 \, ds \right]$$

$$\leq C \mathbb{E}^{P^n} \left[1 + |X_v(0)|^2 + \int_0^t \left(\|X_v\|_{*,s}^2 + \frac{1}{|N_v(G)|} \sum_{u \in N_v(G)} \|X_u\|_{*,s}^2 \right) ds \right],$$

where $C \in (0, \infty)$ is a constant that can change from line to line but does not depend on n or v. This implies that

$$\sup_{v \in V_n} \mathbb{E}^{P^n} \left[\|X_v\|_{*,t}^2 \right] \leq C \left(1 + \sup_{v \in V} \mathbb{E}^{P^n} \left[|X_v(0)|^2 \right] + \int_0^t \sup_{v \in V_n} \mathbb{E}^{P^n} \left[\|X_v\|_{*,s}^2 \right] ds \right),$$

where we have used the inclusion $V_n \subset V$. Apply Gronwall's inequality to find

$$\sup_{v \in V_n} \mathbb{E}^{P^n} \left[\|X_v\|_{*,t}^2 \right] \leq C \left(1 + \sup_{v \in V} \int_{(\mathbb{R}^d)^V} |x_v|^2 \, \mu_0(dx_V) \right). \tag{5.5}$$

The right-hand side is finite by Assumption (**B**.1), and so (5.2) follows.

Given a finite subset $A \subset V$, define $Q_t^n \in \mathcal{P}(\mathcal{C}_t^V)$ by

$$\frac{dQ_t^n}{dP_t^{*,\mu_0}} = \prod_{v \in V_n \setminus A} \mathcal{E}_t(M_v^n), \quad \text{where} \quad M_v^n := \int_0^{\cdot} (\sigma_v \sigma_v^{\top})^{-1} b_v^n(s, X_v, X_{N_v(G_n)}) \cdot dX_v(s), \ v \in V.$$

Due to Assumptions (**B**.2) and (**B**.3), and the definition of P^{*,μ_0} from Remark 2.5, we can apply Lemma B.1 with $\mathbb{Q} = P^{*,\mu_0}$, $X = (X_v)_{v \in V_n}$ and $f(t, x) = (\mathbf{1}_{\{v \in V_n \setminus A\}} \sigma_v^{-1} b_v^n(t, x_v, x_{N_v(G_n)}))_{v \in V_n}$, to conclude that $\frac{dQ_t^n}{dP_t^{*,\mu_0}}$ is a true P^{*,μ_0}-martingale. It then follows from Girsanov's theorem [13, Corollary 3.5.2] and the uniqueness in law of the driftless SDE (2.3) that $Q_t^n[A] = P_t^{*,\mu_0}[A]$. By a similar argument,

$$\frac{dP_t^n}{dP_t^{*,\mu_0}} = \prod_{v \in V_n} \mathcal{E}_t(M_v^n), \quad \text{and thus} \quad \frac{dP_t^n}{dQ_t^n} = \prod_{v \in A \cap V_n} \mathcal{E}_t(M_v^n),$$

where, in case $A \cap V_n = \emptyset$, we interpret the empty product as 1. Once again invoking the linear growth of b, the boundedness of σ_v, (5.5), the fact that $Q_t^n[A] = P^{*,\mu}[A]$ and Remark 2.5, note that Girsanov's theorem also shows that for every $v \in A \cap V_n$, under P^n, $M_v^n - [M_v^n]$ is a martingale and $[M_v^n](t) = \int_0^t |\sigma_v^{-1} b_v^n(s, X_v, X_{N_v(G_n)})|^2 \, ds$.

Recall the well known *data processing inequality* of relative entropy, which states that $H(\nu \circ f^{-1} \,|\, \nu' \circ f^{-1}) \leq H(\nu \,|\, \nu')$ for any two probability measures ν and ν' on a Polish space E, and

any measurable function $f : E \to E'$ to another Polish space E'; for a proof see, for instance, [1, Section 4.2] or [28, Appendix E]. Using this data processing inequality, we obtain

$$
\begin{aligned}
H(P_t^n[A] \,|\, P_t^{*,\mu_0}[A]) &= H(P_t^n[A] \,|\, Q_t^n[A]) \\
&\leq H(P_t^n \,|\, Q_t^n) \\
&= \sum_{v \in A \cap V_n} \mathbb{E}^{P^n} \left[M_v^n(t) - \frac{1}{2}[M_v^n](t) \right] \\
&= \frac{1}{2} \sum_{v \in A \cap V_n} \mathbb{E}^{P^n} \left[\int_0^t |\sigma_v^{-1} b_v^n(s, X_v, X_{N_v(G_n)})|^2 \, ds \right] \\
&\leq C \sum_{v \in A \cap V_n} \mathbb{E}^{P^n} \left[1 + \|X_v\|_{*,t}^2 + \frac{1}{|N_v(G_n)|} \sum_{u \in N_v(G_n)} \|X_u\|_{*,t}^2 \right] \\
&\leq C|A| \left(1 + 2 \sup_{v \in V_n} \mathbb{E}^{P^n} \left[\|X_v\|_{*,t}^2 \right] \right).
\end{aligned}
$$

Therefore (5.3) follows from (5.2).

Noting that due to the identity $Q_t^n[A] = P^{*,\mu}[A]$ and Remark 2.5, under Q^n, $(X_v)_{v \in A}$ is driftless and M_v^n is a martingale. Therefore, noting that

$$
\frac{dQ_t^n}{dP_t^n} = \prod_{v \in A \cap V_n} \exp \left(-M_v^n(t) + \frac{1}{2} \int_0^t |\sigma_v^{-1} b_v^n(s, X_v, X_{N_v(G_n)})|^2 \, ds \right),
$$

another application of the the data processing inequality of relative entropy yields

$$
\begin{aligned}
H(P_t^{*,\mu_0}[A] \,|\, P_t^n[A]) &= H(Q_t^n[A] \,|\, P_t^n[A]) \\
&\leq H(Q_t^n \,|\, P_t^n) \\
&= \frac{1}{2} \sum_{v \in A \cap V_n} \mathbb{E}^{Q^n} \left[\int_0^t |\sigma_v^{-1} b_v^n(s, X_v, X_{N_v(G_n)})|^2 \, ds \right] \\
&\leq C \sum_{v \in A \cap V_n} \mathbb{E}^{Q^n} \left[1 + \|X_v\|_{*,t}^2 + \frac{1}{|N_v(G_n)|} \sum_{u \in N_v(G_n)} \|X_u\|_{*,t}^2 \right] \\
&\leq C|A| \left(1 + 2 \sup_{v \in V_n} \mathbb{E}^{Q^n} \left[\|X_v\|_{*,t}^2 \right] \right).
\end{aligned}
$$

The same argument that was used to obtain (5.5) can also be used to show that (5.5) holds with P^n replaced by Q^n. Therefore $\sup_n \sup_{v \in V_n} \mathbb{E}^{Q^n} \left[\|X_v\|_{*,t}^2 \right] < \infty$ by Assumption (**B**.1), and hence the last display implies (5.4). $\qquad\square$

The next lemma will be used to show both that the *existence* of a weak solution to the infinite SDE system (2.1) holds automatically under Assumption **B**, and also that it arises as the limit of finite-graph systems. Recall that $P \in \mathcal{P}(\mathcal{C}^V)$ denotes the law of the solution of (2.1). Recall also, as discussed at the beginning of this section, that we work throughout with a sequence $G_n = (V_n, E_n)$ as in Lemma 4.2, with b^n and P^n defined as in (5.1) and the sentences thereafter.

Lemma 5.2. *Suppose Assumption **B** holds. Then $P^n \to P$ weakly on $\mathcal{C}^V$. Moreover, for any finite set $A' \subset V$, any $t > 0$, and any bounded measurable function $\psi : \mathcal{C}_t^{A'} \to \mathbb{R}$, we have*

$$
\lim_{n \to \infty} \mathbb{E}^{P^n}[\psi(X_{A'}[t])] = \mathbb{E}^P[\psi(X_{A'}[t])].
$$

Proof. Fix $t > 0$. The entropy bound of (5.3) shows that $\{P_t^n[A']\}_{n\in\mathbb{N}}$ are precompact in the weak* topology induced on $\mathcal{P}(\mathcal{C}_t^{A'})$ by the bounded measurable functions on $\mathcal{C}_t^{A'}$ [4, Lemma 6.2.16]. In particular, this sequence is tight, and since this holds for every finite set A' and every $t > 0$ we deduce that the entire sequence $\{P^n\}_{n\in\mathbb{N}}$ is tight in $\mathcal{C}^V$. Note also that for sufficiently large n it holds under P^n that the processes

$$\int_0^s \sigma_v^{-1}(r, X_v)\, dX_v(r) - \int_0^s \sigma_v^{-1}(r, X_v) b_v(r, X_v, X_{N_v(G)})\, dr, \quad s \geq 0, \quad v \in V_{n-2}, \qquad (5.6)$$

are independent standard Wiener processes, due to the consistency condition for the b_v^n's and the identity $N_v(G_n) = N_v(G)$ valid for $v \in V_{n-2}$.

Now let $Q \in \mathcal{P}(\mathcal{C}^V)$ be any weak (in the usual sense) subsequential limit of $\{P^n\}_{n\in\mathbb{N}}$. That is, let $\{n_k\}_{k\in\mathbb{N}}$ be a subsequence such that $n_k \to \infty$ and $P^{n_k} \to Q$ weakly as $k \to \infty$. The aforementioned precompactness in the weak* topology implies that

$$\lim_{k\to\infty} \mathbb{E}^{P^{n_k}}[\psi(X_{A'}[t])] = \mathbb{E}^Q[\psi(X_{A'}[t])],$$

for any finite set $A' \subset V$, any $t > 0$, and any bounded measurable function ψ on $\mathcal{C}_t^{A'}$. We conclude that, under Q, the processes in (5.6) are independent Wiener processes, for $v \in V$. This shows that Q is the law of a weak solution of the SDE system (2.1), which we know to be unique by assumption (**B**.5). Hence, $Q = P$. $\qquad\square$

5.2. **Proof of the second-order Markov random field property on the infinite graph.**

Proof of Theorem 2.7. Fix $(G = (V, E), b, \sigma, \mu_0)$ and $X = (X_v)_{v\in V}$ as in the statement of the theorem. For $n \geq 4$, consider the sequence of graphs $G_n = (V_n, E_n), n \in \mathbb{N}$ constructed from G as in Lemma 4.2. We first note that due to the fact that μ_0 is a 2MRF by Assumption (**B**.1), part (iii) of Lemma 4.2, with $\mathcal{X} = \mathbb{R}^d, \nu = \mu_0, \nu^* = \mu_0^*$, ensures that $\mu_0[V_n]$ is a 2MRF with respect to the graph G_n. Moreover, since $d\mu_0[V_n]/d\mu_0^*[V_n]$ is strictly positive by Assumption (**B**.1), Proposition 3.2 shows that $\mu_0[V_n]$ admits a 2-clique factorization with respect to the product measure $\mu_0^*[V_n]$ for each n. Hence, $\mu_0[V_n]$ satisfies Assumption (**A**.1), which when combined with the definition of $b^n = (b_v^n)_{v\in V_n}$ in (5.1) and the fact that b, σ satisfy Assumptions (**B**.2) and (**B**.3), shows that $(G_n, b^n, (\sigma_v)_{v\in V_n}, \mu_0[V_n])$ satisfy Assumption **A**. Since $P^n[V_n]$ is the law of the SDE (4.2) on the finite graph G_n, it is a 2MRF by Theorem 2.4.

Now, fix two finite sets $A, B \subset V$ with B disjoint of $A \cup \partial^2 A$, where throughout, we use ∂^2 to denote ∂_G^2. Let n_0 denote the smallest integer greater than or equal to 4 for which $A \cup \partial^2 A \cup B \subset V_{n_0-3}$, and let $n \geq n_0$. Then, part (iv) of Lemma 4.2, again with $\mathcal{X} = \mathbb{R}^d$, $\nu = \mu_0$, and $\nu^* = \mu_0^*$, ensures that $\mu_0[V_n]$ and $\mu_0[V_{n_0}]$ admit 2-clique factorizations which are consistent in the sense that the corresponding measurable functions $f_K^{G_n}$ and $f_K^{G_{n_0}}$ agree for every $K \in \text{cl}_2(G_{n_0})$ that intersects A (equivalently, for every $K \in \text{cl}_2(G)$ that intersects A). Since $b_v^n = b_v^{n_0} = b_v$ for all $v \in A \cup \partial^2 A$ by (5.1), and since $A \cup \partial_G^2 A \subset V_{n_0-3}$, we may apply Proposition 4.4, with $G = G_n$, $H = G_{n_0}$, $V^* = V_{n_0-3}$, $\mu_0^{G_k} = \mu_0[V_k]$, and $(b_v^{G_k}, \sigma_v^{G_k}) = (b_v, \sigma_v)$ for $v \in G_k$ and $k \in \{n_0, n\}$, to deduce that $P_t^n[A\,|\,\partial^2 A] = P_t^{n_0}[A\,|\,\partial^2 A]$ for all $n \geq n_0$. In other words, this implies that given a bounded continuous function $f : \mathcal{C}_t^A \to \mathbb{R}$, there exists a measurable function $\varphi : \mathcal{C}_t^{\partial^2 A} \to \mathbb{R}$ (that does not depend on n) such that

$$\varphi(X_{\partial^2 A}[t]) = \mathbb{E}^{P^n}[f(X_A[t])\,|\,X_{\partial^2 A}[t]], \quad P^n - a.s., \text{ for } n \geq n_0. \qquad (5.7)$$

Now, fix additional bounded continuous functions $g : \mathcal{C}_t^{\partial^2 A} \to \mathbb{R}$ and $h : \mathcal{C}_t^B \to \mathbb{R}$. For $t > 0$, taking the conditional expectation with respect to $X_{V_n\setminus A}[t]$ inside the expectation on the left-hand side below and using the 2MRF property of P^n we have

$$\mathbb{E}^{P^n}[f(X_A[t])g(X_{\partial^2 A}[t])h(X_B[t])] = \mathbb{E}^{P^n}[\mathbb{E}^{P^n}[f(X_A[t])\,|\,X_{\partial^2 A}]g(X_{\partial^2 A}[t])h(X_B[t])].$$

When combined with (5.7), this implies

$$\mathbb{E}^{P^n}[f(X_A[t])g(X_{\partial^2 A}[t])h(X_B[t])] = \mathbb{E}^{P^n}[\varphi(X_{\partial^2 A}[t])g(X_{\partial^2 A}[t])h(X_B[t])].$$

Using the second part of Lemma 5.2 for the finite set $A' = A \cup \partial^2 A \cup B$, and for both $\psi(y_{A'}) := f(y_A)g(y_{\partial^2 A})h(y_B)$, and $\psi(y_{A'}) = \varphi(y_{\partial^2 A})g(y_{\partial^2 A})h(y_B)$ for $y_{A'} \in \mathcal{C}_t^{A'}$, we may pass to the limit $n \to \infty$ and denote $P = P^{\mu_0}$ to get

$$\mathbb{E}^{P}[f(X_A[t])g(X_{\partial^2 A}[t])h(X_B[t])] = \mathbb{E}^{P}[\varphi(X_{\partial^2 A}[t])g(X_{\partial^2 A}[t])h(X_B[t])].$$

This at once shows both that

$$\mathbb{E}^{P}[f(X_A[t]) \,|\, X_{\partial^2 A}[t]] = \varphi(X_{\partial^2 A}[t]) = \mathbb{E}^{P^n}[f(X_A[t]) \,|\, X_{\partial^2 A}[t]],$$

for all bounded continuous f and $n \geq n_0$, which proves Proposition 5.4 below, and also that $X_A[t]$ and $X_B[t]$ are conditionally independent given $X_{\partial^2 A}[t]$ under P. The latter proves the first statement in Theorem 2.7, except for the fact that we have only proven this conditional independence when A and B are finite. Because $B \subset V \setminus (A \cup \partial^2 A)$ was an arbitrary finite set and $\{X_B[t] : B \subset A \cup \partial^2 A\}$ generates the same σ-field as $X_{A \cup \partial^2 A}[t]$, we deduce that that P_t is a 2MRF. The second statement follows from the same argument as in (3.6) using the 2MRF property of $P_t = P_t^{\mu_0}$. This completes the proof. $\square$

We recapitulate two results that were established in the course of the proof, which may be of independent interest, and which are used in the proof of Theorem 2.9 in the next section.

Remark 5.3. Note that the first paragraph of the proof above shows that if (G, b, σ, μ_0) satisfy Assumption **B** and for $G_n = (V_n, E_n)$, $n \in \mathbb{N}$, is as in Lemma 4.2, and b^n, P^n, P_t^n, $n \in \mathbb{N}$, $t > 0$, are as defined at the beginning of Section 5, then $P_t^n[V_n]$ is a 2MRF for each $n \in \mathbb{N}$ and $t > 0$.

Proposition 5.4. *Suppose (G, b, σ, μ_0) satisfy Assumption **B**, and let $G_n = (V_n, E_n), n \in \mathbb{N}$, be the sequence of graphs constructed from G as in Lemma 4.2. Fix $t > 0$, and let $P_t = P_t^{\mu_0}$ be the law of the unique weak solution to the SDE (2.1) with initial law μ_0. Then, for $n \geq 3$, and $A \subset V_{n-3}$, $\partial_G^2 A \subset V_n$ and, P_t-almost surely,*

$$P_t[A \,|\, \partial_G^2 A] = P_t^n[A \,|\, \partial_G^2 A].$$

6. Proof of Gibbs measure properties

In this section we prove Theorem 2.9. Recall the definition of P^{*,μ_0} as the law of the solution of (2.5) initialized at μ_0.

Proof of Theorem 2.9. Let (G, b, σ, μ_0) satisfy Assumption **B**, and let P^{μ_0} be the unique solution of the SDE system (2.1) with initial law μ_0. We work again on the canonical space $(\mathcal{C}^V, \mathcal{B}(\mathcal{C}^V), P^{*,\mu_0})$, with $X_V = (X_v)_{v \in V}$ denoting the canonical process. Define the sets $\mathcal{M}_{\text{init}} = \mathcal{M}_{\text{init}}(\mu_0)$ and $\mathcal{M}_{\text{path}} = \mathcal{M}_{\text{path}}(\mu_0)$ as in the statement of the theorem. For any $\nu_0 \in \mathcal{M}_{\text{init}}$, the SDE system (2.1) is well-posed starting from ν_0, and we let $P^{\nu_0} \in \mathcal{P}(\mathcal{C}^V)$ denote the law of this solution. The proof of the theorem is broken down into five claims.

Claim 1.   Suppose (G, b, σ, μ_0) satisfies Assumption **B**. If $\nu_0 \in \mathcal{M}_{\text{init}}$, then (G, b, σ, ν_0) also satisfies Assumption **B**, and for every finite set $A \subset V$ and $t > 0$ we have $P_t^{\nu_0}[A] \sim P_t^{\mu_0}[A]$.

Proof of Claim 1. First, suppose $\nu_0 \in \mathcal{M}_{\text{init}}$. By the definition of $\mathcal{M}_{\text{init}}$, ν_0 has a finite second moment. Moreover, for each finite set $A \subset V$, we have $\nu_0[A] \sim \mu_0^*[A]$ since $\nu_0 \in \mathcal{G}_2(\mu_0)$ implies (by Definition 2.8) that $\nu_0[A] \sim \mu_0[A]$, and Assumption (**B**.1) ensures that $\mu_0[A] \sim \mu_0^*[A]$. Thus ν_0 satisfies Assumption (**B**.1).

Now let $t > 0$, and let $A \subset V$ be finite. From Lemma 5.2 it follows that $P^{n,\mu_0} \to P^{\mu_0}$ weakly. It then follows from (5.3), (5.4) and the lower semicontinuity of relative entropy that $P_t^{\mu_0}[A] \ll$

$P_t^{*,\mu_0}[A]$ and $P_t^{*,\mu_0}[A] \ll P_t^{\mu_0}[A]$. Therefore $P_t^{\mu_0}[A] \sim P_t^{*,\mu_0}[A]$, and similarly $P_t^{\nu_0}[A] \sim P_t^{*,\nu_0}[A]$. Finally, note that $P^{*,\mu_0}[A]$ (resp. $P^{*,\nu_0}[A]$) is the law of the solution of the SDE system

$$dX_v(t) = \sigma_v(t, X_v)\, dW_v(t), \qquad v \in A,$$

with initial law $\mu_0[A]$ (resp. $\nu_0[A]$), and it follows from $\nu_0[A] \sim \mu_0[A]$ that $P_t^{*,\mu_0}[A] \sim P_t^{*,\nu_0}[A]$. Putting it together, we have $P_t^{\nu_0}[A] \sim P_t^{*,\nu_0}[A] \sim P_t^{*,\mu_0}[A] \sim P_t^{\mu_0}[A]$. $\square$

Claim 2. For any $Q \in \mathcal{M}_{\text{path}}$, we have $Q_0 := Q \circ (X_V(0))^{-1} \in \mathcal{M}_{\text{init}}$.

Proof of Claim 2. The proof of this claim is straightforward: fix $Q \in \mathcal{M}_{\text{path}}$, and set $Q_0 := Q \circ (X_V(0))^{-1}$. Then by the definition of $\mathcal{M}_{\text{path}}$, we have $Q_t \in \mathcal{G}_2(P_t^{\mu_0})$ for all $t \geq 0$ and $\sup_{v \in V} \int_{\mathbb{R}^d} |x_v|^2\, Q_0(dx) < \infty$. Taking $t = 0$ gives $Q_0 \in \mathcal{G}_2(\mu_0)$, where we have used the elementary fact that $P_0^{\mu_0} = P^{\mu_0} \circ (X_V(0))^{-1} = \mu_0$. Thus Q_0 belongs to $\mathcal{M}_{\text{init}}$. $\square$

Claim 3. If $\nu_0 \in \mathcal{M}_{\text{init}}$ then $P^{\nu_0} \in \mathcal{M}_{\text{path}}$.

Proof of Claim 3. Fix $\nu_0 \in \mathcal{M}_{\text{init}}$. Then by the first assertion of Claim 1, for every finite set $A \subset V$ and $t \geq 0$, $P_t^{\nu_0}[A] \sim P_t^{\mu_0}[A]$. So it only remains to show that for every $t > 0$,

$$P_t^{\nu_0}[A \,|\, \partial^2 A] = P_t^{\mu_0}[A \,|\, \partial^2 A], \quad \text{for finite } A \subset V. \tag{6.1}$$

First, recall that Claim 1 also shows that (G, b, σ, ν_0) satisfies Assumption **B**. Next, let $G_n = (V_n, E_n)$ be the increasing sequence of finite graphs defined in Lemma 4.2, and let $P^{\mu_0,n}, P^{\nu_0,n} \in \mathcal{P}(\mathcal{C}^V)$ denote the law of the solution of the corresponding SDE system (4.2) with initial laws $\mu_0[V_n]$ and $\nu_0[V_n]$, respectively. Throughout this proof, the boundary operator ∂ is always with respect to the infinite graph G. Fix $A \subset V$ finite, and fix n large enough that $A \subset V_{n-3}$, recalling that V_n was defined in Section 4.1. By Proposition 5.4, we have both

$$\begin{aligned}
P_t^{\mu_0}[A|\partial^2 A] &= P_t^{\mu_0,n}[A|\partial^2 A], \\
P_t^{\nu_0}[A|\partial^2 A] &= P_t^{\nu_0,n}[A|\partial^2 A].
\end{aligned} \tag{6.2}$$

By Lemma 4.2(iii), $\nu_0[V_n]$ is a 2MRF. Also since $\nu_0[V_n] \sim \mu_0^*[V_n]$ implies $d\nu_0[V_n]/d\mu_0^*[V_n] > 0$, by Proposition 3.2 there is a 2-clique factorization

$$\frac{d\nu_0[V_n]}{d\mu_0^*[V_n]}(x_{V_n}) = \prod_{K \in \text{cl}_2(G_n)} g_K^n(x_K), \tag{6.3}$$

for some measurable functions $g_K^n : (\mathbb{R}^d)^K \to \mathbb{R}_+$. Similarly $\mu_0[V_n]$ admits a 2-clique factorization,

$$\frac{d\mu_0[V_n]}{d\mu_0^*[V_n]}(x_{V_n}) = \prod_{K \in \text{cl}_2(G_n)} f_K^n(x_K), \tag{6.4}$$

for some measurable functions $f_K^n : (\mathbb{R}^d)^K \to \mathbb{R}_+$. Note that these 2-clique factorizations are not unique. We claim (and justify below) that f_K^n and g_K^n can be chosen to be consistent, i.e., so that

$$f_K^n \equiv g_K^n, \quad \text{for all } K \in \text{cl}_2(G_n) \text{ with } K \cap A \neq \emptyset. \tag{6.5}$$

This will ultimately allow us to apply Proposition 4.4. To see this, let $I_{V_n} = (I_v)_{v \in V_n}$ denote the canonical random variable on the probability space $(\mathbb{R}^d)^{V_n}$. That is, $I_v : (\mathbb{R}^d)^{V_n} \to \mathbb{R}^d$ is defined by $I_v((x_u)_{u \in V_n}) := x_v$ for each $v \in V_n$. Define

$$\widehat{f}^n(I_{U_n}) = \mathbb{E}^{\mu_0^*}\left[\frac{d\mu_0[V_n]}{d\mu_0^*[V_n]} \,\bigg|\, I_{U_n}\right], \qquad \widehat{g}^n(I_{U_n}) = \mathbb{E}^{\mu_0^*}\left[\frac{d\nu_0[V_n]}{d\mu_0^*[V_n]} \,\bigg|\, I_{U_n}\right].$$

Recalling that $U_n = V_n \setminus V_{n-2}$, and using (6.3), we have

$$\frac{d\nu_0[V_{n-2} \,|\, U_n]}{d\mu_0^*[V_{n-2} \,|\, U_n]} = \frac{\frac{d\nu_0[V_n]}{d\mu_0^*[V_n]}}{\widehat{g}^n(I_{U_n})} = \frac{1}{\widehat{g}^n(I_{U_n})} \prod_{K \in \mathrm{cl}_2(G_n)} g_K^n(I_K).$$

Applying the same argument to μ_0 rather than ν_0 and using (6.4), we also obtain

$$\frac{d\mu_0[V_{n-2} \,|\, U_n]}{d\mu_0^*[V_{n-2} \,|\, U_n]} = \frac{\frac{d\mu_0[V_n]}{d\mu_0^*[V_n]}}{\widehat{f}^n(I_{U_n})}. \tag{6.6}$$

Further, recognizing that $U_n = \partial^2 V_{n-2}$, since $\nu_0 \in \mathcal{G}_2(\mu_0)$ we see that

$$\nu_0[V_{n-2}|U_n] = \mu_0[V_{n-2}|U_n].$$

Combining the last three displays, we find

$$\frac{d\mu_0[V_n]}{d\mu_0^*[V_n]}(I_{V_n}) = \frac{d\nu_0[V_{n-2} \,|\, U_n]}{d\mu_0^*[V_{n-2} \,|\, U_n]}\widehat{f}^n(I_{U_n}) = \frac{\widehat{f}^n(I_{U_n})}{\widehat{g}^n(I_{U_n})} \prod_{K \in \mathrm{cl}_2(G_n)} g_K^n(I_K).$$

Comparing this and (6.6) with (6.3) and (6.4) and noting that $U_n \in \mathrm{cl}_2(G_n)$, we can thus take $f_K^n \equiv g_K^n$ in (6.4) for $K \in \mathrm{cl}_2(G_n) \setminus \{U_n\}$ and $f_{U_n}^n \equiv g_{U_n}^n \widehat{f}^n/\widehat{g}^n$. This proves the consistency claim in (6.5); indeed, since $A \subset V_{n-3}$, we know that U_n does not intersect $A \cup \partial A$.

Let $\mathcal{K} = \{K \in \mathrm{cl}_2(G_n) : K \cap A \neq \emptyset\}$. Using the consistency property (6.5), we can finally conclude from Proposition 4.4 (with $V_G = V_H = V_n$, $b_v^H = b_v^G = b_v$, $\sigma_v^H = \sigma_v^H = \sigma_v$, $\mu_0^H = \mu_0$ and $\mu_0^G = \nu_0$) that $P_t^{\nu_0,n}[A \,|\, \partial^2 A] = P_t^{\mu_0,n}[A \,|\, \partial^2 A]$. Recalling (6.2), this completes the proof of (6.1). $\qquad\square$

Together, Claims 2 and 3 prove (2.6). We now prove the last assertion of the theorem. For this we will frequently apply the martingale representation theorem; for completeness, we provide a full justification of its repeated application in Remark 6.1, but defer it to the end of the proof so as not to interrupt the flow of the main arguments.

Claim 4. If $Q \in \mathcal{P}(\mathcal{C}^V)$ satisfies $Q_t \in \mathcal{G}_2(P_t^{\mu_0})$ for all $t \geq 0$ and also $Q \circ (X_V(0))^{-1} = \mu_0$, then $Q = P^{\mu_0}$.

Proof of Claim 4. As in the proof of Claim 3, we let $G_n = (V_n, E_n)$ be the increasing sequence of finite graphs defined in Lemma 4.2 and let the boundary operator ∂ always be with respect to the infinite graph G. Also, let $P^n = P^{\mu_0,n} \in \mathcal{P}(\mathcal{C}^V)$ denote the law of the solution of the corresponding SDE system (4.2) with initial law $\mu_0[V_n]$. Now, fix a finite set $A \subset V$ and $T \in (0, \infty)$. Let n_0 denote the smallest integer such that $A \cup \partial^2 A \subset V_{n_0-3}$. Define the martingales

$$M_v^n(t) = \int_0^t (\sigma_v \sigma_v^\top)^{-1} b_v^n(s, X_v, X_{N_v(G_n)}) \cdot dX_v(s), \qquad n \in \mathbb{N}, v \in V_n.$$

Due to Assumptions (**B**.2) and (**B**.3), it follows from Lemma B.1 (with $\mathbb{Q} = P^{*,\mu_0}$, $X = (X_v)_{v \in V_n}$ and $f(t,x) = (\sigma_v^{-1} b_v^n(t, x_v, x_{N_v(G_n)}))_{v \in V_n}$) that $\mathcal{E}(M_v^n)$ is a P^{*,μ_0}-martingale. Thus, by Girsanov's theorem [13, Corollary 3.5.2], we may write

$$\frac{dP_t^n[V_n]}{dP_t^{*,\mu_0}[V_n]} = \prod_{v \in V_n} \mathcal{E}_t(M_v^n). \tag{6.7}$$

Now, by applying Remark 5.3 to (G, b, σ, μ_0) and $(G, 0, \sigma, \mu_0)$, respectively, it follows that the measures $P_t^n[V_n]$ and $P_t^{*,\mu_0}[V_n]$ are 2MRFs with respect to G_n. Hence, for $n \geq n_0$,

$$
\begin{aligned}
\frac{dP_t^n[A \mid \partial^2 A]}{dP_t^{*,\mu_0}[A \mid \partial^2 A]} &= \frac{dP_t^n[A \mid V_n \backslash A]}{dP_t^{*,\mu_0}[A \mid V_n \backslash A]} \\
&= \frac{dP_t^n[V_n]}{dP_t^{*,\mu_0}[V_n]} \bigg/ \mathbb{E}^{P^{*,\mu_0}} \left[\left. \frac{dP_t^n[V_n]}{dP_t^{*,\mu_0}[V_n]} \right| X_{V_n \backslash A}[t] \right] \\
&= \prod_{v \in V_n} \mathcal{E}_t(M_v^n) \bigg/ \mathbb{E}^{P^{*,\mu_0}} \left[\left. \prod_{v \in V_n} \mathcal{E}_t(M_v^n) \right| X_{V_n \backslash A}[t] \right].
\end{aligned}
$$

For $v \in V_n \backslash (A \cup \partial A)$, $\mathcal{E}_t(M_v^n)$ is measurable with respect to $X_{V_n \backslash A}[t]$ and thus factors out of the conditional expectation and cancels. Thus,

$$
\frac{dP_t^n[A \mid \partial^2 A]}{dP_t^{*,\mu_0}[A \mid \partial^2 A]} = \prod_{v \in A \cup \partial A} \mathcal{E}_t(M_v^n) \bigg/ \mathbb{E}^{P^{*,\mu_0}} \left[\left. \prod_{v \in A \cup \partial A} \mathcal{E}_t(M_v^n) \right| X_{V_n \backslash A}[t] \right]. \qquad (6.8)
$$

Because $Q_t \in \mathcal{G}_2(P_t^{\mu_0})$ by assumption, we have $Q_t[A \mid \partial^2 A] = P_t^{\mu_0}[A \mid \partial^2 A]$. By Proposition 5.4, we have $P_t^{\mu_0}[A \mid \partial^2 A] = P_t^{\mu_0,n}[A \mid \partial^2 A]$, and it follows that the density $dQ_t[A \mid \partial^2 A]/dP_t^{*,\mu_0}[A \mid \partial^2 A]$ is given by the same expression (6.8).

Now take $A = V_{n-2}$, and note that $U_n := V_n \setminus V_{n-2} = \partial^2 V_{n-2}$. Because $Q_t[U_n] \sim P_t^{*,\mu_0}[U_n]$ by assumption, and because both Q and P^{*,μ_0} start from the same initial state distribution μ_0, we may use the martingale representation theorem (specifically, apply Remark 6.1 below with $\xi = dQ_T[U_n]/dP_T^{*,\mu_0}[U_n]$, which clearly satisfies $\mathbb{E}^{P^{*,\mu_0}}[\xi] = 1$) to find progressively measurable functions $r_v^n : [0, T] \times \mathcal{C}^{U_n} \to \mathbb{R}^d, v \in U_n$, which are $dt \otimes dP^{*,\mu_0}$-square-integrable such that in terms of the associated X_{U_n}-adapted continuous martingales $R_v^n(t) = \int_0^t r_v^n(s, X_{U_n}) \cdot dX_v(s)$, $t \in [0, T]$, $v \in U_n$, we can write for $t \in [0, T]$,

$$
\frac{dQ_t[U_n]}{dP_t^{*,\mu_0}[U_n]} = \prod_{v \in U_n} \mathcal{E}_t(R_v^n).
$$

Noting that $(X_v)_{v \in U_n}$ are orthogonal martingales, it follows that the martingales $(R_v^n)_{v \in U_n}$ are also orthogonal; that is, the covariation process $[R_v^n, R_u^n]$ is identically zero for $v \neq u$. Thus, since $U_n \cap V_{n-2} = \emptyset$, $U_n \cup V_{n-2} = V_n$ and $V_{n-1} = V_{n-2} \cup \partial V_{n-2}$, applying (6.8) with $A = V_{n-2}$ we have

$$
\begin{aligned}
\frac{dQ_t[V_n]}{dP_t^{*,\mu_0}[V_n]} &= \frac{dQ_t[V_{n-2} \mid U_n]}{dP_t^{*,\mu_0}[V_{n-2} \mid U_n]} \frac{dQ_t[U_n]}{dP_t^{*,\mu_0}[U_n]} \\
&= \prod_{v \in U_n} \mathcal{E}_t(R_v^n) \prod_{v \in V_{n-1}} \mathcal{E}_t(M_v^n) \bigg/ \mathbb{E}^{P^{*,\mu_0}} \left[\left. \prod_{v \in V_{n-1}} \mathcal{E}_t(M_v^n) \right| X_{U_n}[t] \right].
\end{aligned}
$$

The process in the denominator is a positive martingale (being the optional projection of a martingale) adapted to the filtration of X_{U_n} and thus, again using the martingale representation theorem (this time applying Remark 6.1 below with $\xi = \mathbb{E}^{P^{*,\mu_0}} \left[\left. \prod_{v \in V_{n-1}} \mathcal{E}_T(M_v^n) \right| X_{U_n}[T] \right]$ and invoking (6.7) to conclude that $\mathbb{E}^{P^{*,\mu_0}}[\xi] = \mathbb{E}^{P^{*,\mu_0}}[dP_t^n[V_{n-1}]/dP_t^{*,\mu_0}[V_{n-1}]] = 1$), there exist $dt \otimes dP^{*,\mu_0}$-square-integrable, progressively measurable functions $\widetilde{r}_v^n : [0, T] \times \mathcal{C}^{U_n} \to \mathbb{R}^d, v \in U_n$, and associated X_{U_n}-adapted continuous martingales $\widetilde{R}_v^n(t) = \int_0^t \widetilde{r}_v^n(s, X_{U_n}) \cdot dX_v(s)$, $v \in U_n$,

such that

$$\mathbb{E}^{P^{*,\mu_0}}\left[\prod_{v \in V_{n-1}} \mathcal{E}_t(M_v^n) \,\middle|\, X_{U_n}[t]\right] = \prod_{v \in U_n} \mathcal{E}_t(\widetilde{R}_v^n),$$

Now note that, for any continuous martingales Z and $(Z_i)_{i \in I}$, with I a finite index set, we have the identities $1/\mathcal{E}(Z) = \mathcal{E}(-Z)e^{[Z]}$ and

$$\prod_{i \in I} \mathcal{E}(Z_i) = \exp\left(\sum_{i \in I} Z_i - \frac{1}{2}\sum_{i \in I}[Z_i]\right)$$

$$= \exp\left(\sum_{i \in I} Z_i - \frac{1}{2}\Big[\sum_{i \in I} Z_i\Big] + \frac{1}{2}\sum_{i,j \in I, i \neq j}[Z_i, Z_j]\right)$$

$$= \mathcal{E}\left(\sum_{i \in I} Z_i\right)\exp\left(\frac{1}{2}\sum_{i,j \in I, i \neq j}[Z_i, Z_j]\right),$$

where $[Z_i, Z_j]$ denotes the covariation process. Hence,

$$\frac{dQ_t[V_n]}{dP_t^{*,\mu_0}[V_n]} = \prod_{v \in U_n} \mathcal{E}_t(R_v^n)\mathcal{E}_t(-\widetilde{R}_v^n)\exp([\widetilde{R}_v^n](t)) \prod_{v \in V_{n-1}} \mathcal{E}_t(M_v^n)$$

$$= \prod_{v \in V_{n-2}} \mathcal{E}_t(M_v^n) \prod_{v \in V_{n-1}\setminus V_{n-2}} \mathcal{E}_t(M_v^n + R_v^n - \widetilde{R}_v^n)\exp\left([M_v^n - \widetilde{R}_v^n, R_v^n - \widetilde{R}_v^n](t)\right)$$

$$\cdot \prod_{v \in V_n\setminus V_{n-1}} \mathcal{E}_t(R_v^n - \widetilde{R}_v^n)\exp\left([\widetilde{R}_v^n, \widetilde{R}_v^n - R_v^n](t)\right).$$

Recalling the orthogonality properties of R_v^n and $\widetilde{R}_v^n$ mentioned above, we see that we can write $Z_t := dQ_t[V_n]/dP_t^{*,\mu_0}[V_n]$ in the form $Z(t) = \mathcal{E}_t(N)e^{A(t)}$, where N is a continuous square-integrable martingale and $A := \sum_{v \in V_{n-1}\setminus V_{n-2}}[M_v^n - \widetilde{R}_v^n, R_v^n - \widetilde{R}_v^n] + \sum_{v \in V_n\setminus V_{n-1}}[\widetilde{R}_v^n, \widetilde{R}_v^n - R_v^n]$ is a square-integrable and a.s. absolutely continuous with $A(0) = 0$. Since Z is a martingale, we necessarily have $A \equiv 0$; indeed, Itô's formula gives $dZ(t) = Z(t)(dN(t) + dA(t))$, and for Z to be a martingale we must have $dA(t) = 0$. It follows that

$$\frac{dQ_t[V_n]}{dP_t^{*,\mu_0}[V_n]} = \prod_{v \in V_{n-2}} \mathcal{E}_t(M_v^n) \prod_{v \in V_{n-1}\setminus V_{n-2}} \mathcal{E}_t(M_v^n + R_v^n - \widetilde{R}_v^n) \prod_{v \in V_n\setminus V_{n-1}} \mathcal{E}_t(R_v^n - \widetilde{R}_v^n).$$

Since Z is a P^{*,μ_0}-martingale, Girsanov's theorem [13, Corollary 3.5.2] can be applied, using the definition of $(M_v^n)_{v \in \mathbb{V}_n}$, to deduce that $Q_t[V_n]$ is the law of a solution $(X_v^n[t])_{v \in V_n}$ of the SDE system (perhaps on an auxiliary probability space)

$$dX_v^n(s) = b_v(s, X_v^n, X_{N_v(G)}^n)\,ds + \sigma(s, X_v^n)\,dB_v(s), \qquad\qquad \text{for } v \in V_{n-2},$$

$$dX_v^n(s) = \left((r_v^n - \widetilde{r}_v^n)(s, X_{U_n}^n) + b_v^n(s, X_v^n, X_{N_v(G_n)}^n)\right)ds + \sigma(s, X_v^n)\,dB_v(s), \text{ for } v \in V_{n-1}\setminus V_{n-2},$$

$$dX_v^n(s) = (r_v^n - \widetilde{r}_v^n)(s, X_{U_n}^n)\,ds + \sigma(s, X_v^n)\,dB_v(s), \qquad\qquad \text{for } v \in V_n\setminus V_{n-1},$$

where $(B_v)_{v \in V_n}$ are independent Brownian motions.

Define $X_v^n \equiv 0$ for $v \notin V_n$. Since the sets V_n increase to V, it is easily shown as in Lemma 5.2 that, as $n \to \infty$, $(X_v^n[T])_{v \in V}$ converges in law in $\mathcal{C}_T^V$ to a solution of the infinite SDE system (2.1) with initial distribution μ_0, restricted to the interval $[0, T]$. Recalling that $(X_v^n(0))_{v \in V} \sim \mu_0$ and that the solution to the infinite SDE system is unique in law by Assumption (**B**.4), we conclude

that $(X_v^n[T])_{v \in V}$ converges in law to $P_T^{\mu_0}$. But $X_{V_n}^n[T]$ has law $Q_T[V_n]$ by construction, which implies $X_V^n[T]$ converges in law to Q_T. Therefore $Q_T = P_T^{\mu_0}$. Since $T \in (0, \infty)$ was arbitrary, $Q = P^{\mu_0}$, which completes the proof of Claim 4. $\qquad\square$

To complete the proof of the theorem, it only remains to establish the bijection between the two sets in (2.6). However, we now show that this is a simple consequence of the last claim.

Claim 5. The map $Q \mapsto Q \circ (X_V(0))^{-1}$ defines a bijection between the sets $\mathcal{M}_{\text{path}}$ and $\mathcal{M}_{\text{init}}$.

Proof of Claim 5. Let $Q \in \mathcal{M}_{\text{path}}$, and set $\nu_0 := Q \circ (X_V(0))^{-1}$. By Claim 2, ν_0 belongs to $\mathcal{M}_{\text{init}}$, and by Claim 3, P^{ν_0} lies in $\mathcal{M}_{\text{path}}$. Since trivially $P^{\nu_0} \circ (X_V(0))^{-1} = \nu_0$, to prove the claim it suffices to prove that $Q = P^{\nu_0}$. By Claim 1, (G, b, σ, ν_0) satisfies Assumption **B**, and thus Claim 4 applies with ν_0 in place of μ_0. That is, by applying Claim 4 to ν_0 instead of μ_0, we deduce that if $Q \in \mathcal{P}(\mathcal{C}^V)$ satisfies $Q_t \in \mathcal{G}_2(P_t^{\nu_0})$ for all $t \geq 0$ and also $Q \circ (X_V(0))^{-1} = \nu_0$, then $Q = P^{\nu_0}$. By definition of $\mathcal{M}_{\text{path}}$ we have $Q_t \in \mathcal{G}_2(P_t^{\mu_0})$ for all $t \geq 0$, and it follows from (6.1), which was established in the proof of Claim 3, that $\mathcal{G}_2(P_t^{\nu_0}) = \mathcal{G}_2(P_t^{\mu_0})$. This proves Claim 5, which completes the proof of Theorem 2.9. $\qquad\square$

Remark 6.1. We sketch here the argument behind the use of the martingale representation theorem in the proof of Theorem 2.9 above. Recall that by Assumption (**A**.3b) the SDE system $dX_v(t) = \sigma_v(t, X_v) \, dW_v(t)$, $v \in U_n$, with initial law μ_0 is unique in law, with the law of the solution $X = (X_v)_{v \in U_n}$ given by $P^{*,\mu_0}[U_n]$. This implies uniqueness of the associated martingale problem (cf. [13, Corollary 5.4.9]), which is known to imply that the solution has the predictable representation property (cf. [27, Theorem V.25.1] or [30, Theorem 2.7]), in the following sense: with $\mathcal{F}_t^X := \sigma(X(s) : s \leq t)$, $t \geq 0$, for $T < \infty$ and an $\mathcal{F}_T^X$-measurable random variable $\xi > 0$ with $\mathbb{E}[\xi] = 1$, the martingale $Z(t) = \mathbb{E}[\xi \, | \, \mathcal{F}_t^X] > 0, t \in [0, T]$, can be represented as $Z(t) = 1 + \int_0^t \varphi(s, X) \cdot dX(s)$ for some predictable process $\varphi \colon [0, T] \times \mathcal{C}_T \to \mathbb{R}$ satisfying $\int_0^T |\varphi(t, X)|^2 \, dt < \infty$ a.s., recalling that σ_v is uniformly bounded and nondegenerate. Then, for $t \in [0, T]$, setting $\psi(t, X) = \varphi(t, X)/Z(t)$, by Itô's formula, we have $d \log Z(t) = \psi(t, X) \cdot dX(t) - \frac{1}{2}\psi(t, X)^\top d[X](t)\psi(t, X)$. Hence, $Z(t) = \mathcal{E}_t(\int_0^{\cdot} \psi(s, X) \cdot dX(s))$, $t \in [0, T]$.

APPENDIX A. PROOF OF PATHWISE UNIQUENESS UNDER LIPSCHITZ ASSUMPTIONS

Proof of Proposition 2.6. Let $(X_v)_{v \in V}$ and $(\widetilde{X}_v)_{v \in V}$ denote two solutions driven by the same Wiener processes and starting from the same initial states.

Fix $T < \infty$. For each $v \in V$ and $t \in [0, T]$, by Itô's formula, the boundedness of σ (see Assumption (**B**.3)), the assumed Lipschitz condition on the drift and diffusion coefficients we have

$$
\mathbb{E}\left[\|X_v - \widetilde{X}_v\|_{*,t}^2\right] \leq 2t\mathbb{E}\left[\int_0^t \left|b_v(s, X_v(s), X_{N_v(G)}(s)) - b_v(s, \widetilde{X}_v(s), \widetilde{X}_{N_v(G)}(s))\right|^2 ds\right]
$$

$$
+ 8\mathbb{E}\left[\int_0^t \left|\sigma_v(s, X_v(s)) - \sigma_v(s, \widetilde{X}_v(s))\right|^2 ds\right]
$$

$$
\leq 4tK_T^2\mathbb{E}\left[\int_0^t \left(\|X_v - \widetilde{X}_v\|_{*,s}^2 + \frac{1}{|N_v(G)|}\sum_{u \in N_v(G)} \|X_u - \widetilde{X}_u\|_{*,s}^2\right) ds\right]
$$

$$
+ 8\bar{K}_T^2\mathbb{E}\left[\int_0^t \|X_v - \widetilde{X}_v\|_{*,s}^2 \, ds\right].
$$

Hence, using the constants K_T and $\bar{K}_T$ from the statement of the proposition,

$$\sup_{v \in V} \mathbb{E}\left[\|X_v - \widetilde{X}_v\|_{*,t}^2\right] \leq 8(tK_T^2 + \bar{K}_T^2) \int_0^t \sup_{v \in V} \mathbb{E}\left[\|X_v - \widetilde{X}_v\|_{*,s}^2\right] ds.$$

Complete the proof using Gronwall's inequality. $\square$

Appendix B. Justification for applying Girsanov's theorem

In this section we state a result that justifies our repeated application of Girsanov's theorem under the condition that the drift is progressively measurable and has linear growth. Lemma B.1 below is in fact a path-dependent multi-dimensional version of [14, Theorems 5.1 and 8.1]. A simpler proof is provided here for completeness.

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$ be a filtered probability space supporting a $\mathbb{F}$-Brownian motion W of dimension m as well as an $\mathbb{F}$-adapted process X of dimension d such that X satisfies the SDE

$$dX(t) = \sigma(t, X)\, dW(t), \quad X(0) \sim \mu, \tag{B.1}$$

where $\mu \in \mathcal{P}(\mathbb{R}^d)$ and $\sigma : \mathbb{R}_+ \times \mathcal{C} \to \mathbb{R}^{d \times m}$ is bounded and progressively measurable. Also, let $\mathbb{E}$ denote expectation with respect to $\mathbb{Q}$. Fix a progressively measurable $f : \mathbb{R}_+ \times \mathcal{C} \mapsto \mathbb{R}^m$, and define the stochastic integral

$$M_t := \int_0^t f(s, X) \cdot dW_s, \qquad t \in [0, \infty),$$

which is well defined (and a local martingale) due to the linear growth condition (B.2) imposed on f in the lemma below. Recall in what follows that $\|x\|_{*,t} = \sup_{s \in [0,t]} |x(s)|$.

Lemma B.1. *Under the above setting, suppose for each $T \in (0, \infty)$ there exists $C_T < \infty$ such that*

$$|f(t, x)| \leq C_T \left(1 + \|x(s)\|_{*,t}\right), \tag{B.2}$$

for all $t \in [0, T]$ and $x \in \mathcal{C}$. Then the stochastic exponential $\{\mathcal{E}_t(M)\}_{t \geq 0}$ defined in (3.4) is a true $\mathbb{Q}$-martingale.

Proof. Since $\{\mathcal{E}_t(M)\}_{t \geq 0}$ is always a $\mathbb{Q}$-supermartingale, it suffices to show that $\mathbb{E}[\mathcal{E}_T(M)|X(0) = x] = 1$ for each $x \in \mathbb{R}^d$ and $T \in (0, \infty)$. So fix $T \in (0, \infty)$ and assume without loss of generality that $X(0) = x \in \mathbb{R}^d$ in (B.1). Since σ is bounded, X is a martingale and, from standard concentration inequalities for martingales (see, e.g., [31, Lemma 2.1]), we can find some $C > 0$ such that $\mathbb{Q}(\|X - x\|_{*,T} \geq a) \leq \exp(-Ca^2)$ for each $a > 0$. It then follows from the equivalence between sub-Gaussian tails and finite square-exponential moments (see, e.g., [2, Section 2.3]) that there exists $c > 0$ such that $\mathbb{E}[\exp(c\|X\|_{*,T}^2)] < \infty$. Now taking $0 = t_0 < t_1 < \cdots < t_{n(T)} = T$ with $t_n - t_{n-1} \leq c/C_T^2$, and using the linear growth condition (B.2) on f, we have

$$\mathbb{E}\left[\exp\left(\frac{1}{2} \int_{t_{n-1}}^{t_n} |f(s, X)|^2\, ds\right)\right] \leq \mathbb{E}\left[\exp\left((t_n - t_{n-1})C_T^2(1 + \|X\|_{*,T}^2)\right)\right] < \infty.$$

It then follows from [13, Corollary 3.5.14] that $\{\mathcal{E}_t(M)\}_{t \geq 0}$ is a true $\mathbb{Q}$-martingale. $\square$

References

[1] S.M. Ali and S.D. Silvey, *A general class of coefficients of divergence of one distribution from another*, Journal of the Royal Statistical Society. Series B. Methodological **28** (1966), 131–142.

[2] S. Boucheron, G. Lugosi, and P. Massart, *Concentration inequalities: A nonasymptotic theory of independence*, Oxford university press, 2013.

[3] P. Cattiaux, S. Roelly, and H. Zessin, *Une approche Gibbsienne des diffusions Browniennes infini-dimensionnelles*, Probability Theory and Related Fields **104** (1996), 147–179.

[4] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, Vol. 38, Springer Science & Business Media, 2009.

[5] D. Dereudre, *Interacting Brownian particles and Gibbs fields on pathspaces*, ESAIM:Probability and Statistics **7** (2003), 251–277.

[6] D. Dereudre and S. Roelly, *On Gibbsianness of infinite-dimensional diffusions*, Markov Processes and Related Fields **10** (2004), 395–410.

[7] _____, *Propagation of Gibbsianness for infinite-dimensional gradient Brownian diffusions*, Journal of Statistical Physics **121** (2005), no. 3-4, 511–551.

[8] _____, *Path-dependent infinite-dimensional SDE with non-regular drift: An existence result*, Ann. Inst. H. Poincaré Probab. Statist. **53** (2017), no. 2, 641–657.

[9] N. Detering, J.-P. Fouque, and T. Ichiba, *Directed chain stochastic differential equations*, Stochastic Processes and their Applications **130** (2020), no. 4, 2519–2551.

[10] J.D. Deuschel, *Infinite-dimensional diffusion processes as Gibbs measures on $C[0,1]^{Z^d}$*, Probability Theory and Related Fields **76** (1987), 325–340.

[11] H.-O. Georgii, *Gibbs measures and phase transitions*, Vol. 9, Walter de Gruyter, 2011.

[12] O. Kallenberg, *Foundations of modern probability*, 2nd ed., Springer Science & Business Media, 2006.

[13] I. Karatzas and S.E. Shreve, *Brownian motion and stochastic calculus*, Graduate Texts in Mathematics, Springer New York, 1991.

[14] F. Klebaner and R. Liptser, *When a stochastic exponential is a true martingale. extension of the Beneš method*, Theory of Probability and Its Applications **58** (2014), no. 1, 38–62.

[15] C. Külske and F. Redig, *Loss without recovery of Gibbsianness during diffusion of continuous spins*, Probability Theory and Related Fields **135** (2006), 428–456.

[16] D. Lacker, K. Ramanan, and R. Wu, *Local weak convergence and propagation of ergodicity for sparse networks of interacting processes*, arXiv preprint arXiv:1904.02585 (2020).

[17] _____, *Marginal dynamics of interacting diffusions on unimodular Galton-Watson trees*, arXiv preprint arXiv:2009.11667 (2020).

[18] S.L. Lauritzen, *Graphical models*, Vol. 17, Clarendon Press, 1996.

[19] E. Luçon and W. Stannat, *Mean field limit for disordered diffusions with singular interactions*, Annals of Applied Probability **24** (2014), no. 5, 1946–1993.

[20] G.S. Medvedev, *The continuum limit of the Kuramoto model on sparse random graphs*, Commun. Math. Sci. **17** (2019), 883–898.

[21] R.A. Minlos, S. Roelly, and H. Zessin, *Gibbs states on space-time*, Potential Analysis **13** (2000), 367–408.

[22] S. Nadtochiy and M. Shkolnikov, *Mean field systems on networks, with singular interaction through hitting times*, Annals of Probability **48** (2020), no. 3, 1520–1556.

[23] F. Rassoul-Agha and T. Seppäläinen, *A course on large deviations with an introduction to Gibbs measures*, Vol. 162, American Mathematical Soc., 2015.

[24] F. Redig, S. Roelly, and W. Ruszel, *Short-time Gibbsianness for infinite-dimensional diffusions with space-time interaction*, Journal of Statistical Physics **138** (2010), 1124–1144.

[25] S. Roelly and W.M. Ruszel, *Propagation of Gibbsianness for infinite-dimensional diffusions with space-time interaction*, Markov Proc. Rel. Fields **20** (2014), 653–574.

[26] S. Roelly and H. Zessin, *Une caractérisation des mesures de Gibbs sur $C(0,1)^{Z^d}$ par le calcul des variations stochastiques*, Annales de l'I. H. P., section B **29** (1993), 327–338.

[27] L.C.G. Rogers and D. Williams, *Diffusions, Markov processes and martingales: Volume 2, Itô calculus*, Vol. 2, Cambridge University Press, 2000.

[28] P. Ménard S. Gerchinovitz and G. Stoltz, *Fano's inequality for random variables*, Statistical Science **35** (2020), no. 2, 178–201.

[29] D.W. Stroock and S.R.S. Varadhan, *Multidimensional diffusion processes*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 233, Springer-Verlag, Berlin-New York, 1979. MR532498

[30] D.W. Stroock and M. Yor, *On extremal solutions of martingale problems*, Annales scientifiques de l'école normale supérieure, 1980, pp. 95–164.

[31] S. van de Geer, *Exponential inequalities for martingales, with application to maximum likelihood estimation for counting processes*, Ann. Statist. **23** (1995), no. 5, 1779–1801.

[32] A.C.D. Van Enter, C. Külske, A.A. Opoku, and W.M. Ruszel, *Gibbs–non-Gibbs properties for n-vector lattice and mean-field models*, Brazilian Journal of Probability and Statistics **24** (2010), no. 2, 226–255.

[33] A.C.D. Van Enter and W.M. Ruszel, *Gibbsianness versus Non-Gibbsianness of time-evolved planar rotor models*, Stochastic Processes and their Applications **119** (2009), no. 6, 1866–1888.

Department of Industrial Engineering and Operations Research, Columbia University, New York, NY

Division of Applied Mathematics, Brown University, Providence, RI

Department of Mathematics, Iowa State University, Ames, IA
Email address: daniel.lacker@columbia.edu, kavita_ramanan@brown.edu, ruoyu@iastate.edu