arXiv:2003.13180v2 [math.PR] 6 Oct 2020

GRAPHON MEAN FIELD SYSTEMS

ERHAN BAYRAKTAR, SUMAN CHAKRABORTY, AND RUOYU WU

Abstract. We consider heterogeneously interacting diffusive particle systems and their large population limit. The interaction is of mean field type with weights characterized by an underlying graphon. A law of large numbers result is established as the system size increases and the underlying graphons converge. The limit is given by a graphon mean field system consisting of independent but heterogeneous nonlinear diffusions whose probability distributions are fully coupled. Well-posedness, continuity and stability of such systems are provided. We also consider a not-so-dense analogue of the finite particle system, obtained by percolation with vanishing rates and suitable scaling of interactions. A law of large numbers result is proved for the convergence of such systems to the corresponding graphon mean field system.

Contents

1. Introduction

In this work we study mean field diffusive particle systems with heterogeneous interaction and their large population limit. The interaction is of mean field type and is characterized

Date: October 7, 2020.

2020 *Mathematics Subject Classification.* 05C80 60J60 60K35.

Key words and phrases. graphons, graphon particle systems, mean field interaction, heterogeneous interaction, networks, percolation.

through a step graphon. More precisely, denoting by X_i^n the state of the i-th particle,

$$\begin{aligned} X_i^n(t) = X_{\frac{i}{n}}(0) &+ \int_0^t \frac{1}{n}\sum_{j=1}^n \xi_{ij}^n b(X_i^n(s), X_j^n(s))\, ds \\ &+ \int_0^t \frac{1}{n}\sum_{j=1}^n \xi_{ij}^n \sigma(X_i^n(s), X_j^n(s))\, dB_{\frac{i}{n}}(s), \quad i \in \{1,\dots,n\}, \end{aligned} \tag{1.1}$$

where b and σ are some bounded and Lipschitz functions, $\{B_u : u \in [0,1]\}$ are i.i.d. d-dimensional Brownian motions, and $X_u(0)$ is a collection of independent $\mathbb{R}^d$-valued random variables, with probability distribution $\mu_u(0)$ for each $u \in [0,1]$, and independent of $\{B_u : u \in [0,1]\}$. Here ξ_{ij}^n determines the interaction between particles i and j, and depends on some step graphon G_n converging to a limiting graphon in the cut metric.

The classic mean field system with homogeneous interaction, which corresponds to $\xi_{ij}^n \equiv 1$ in (1.1), dates back to works of Boltzmann, Vlasov, McKean and others (see [21, 24, 30] and references therein). While the original motivation for the study came from statistical physics, similar models have arisen in many different application areas, including economics, chemical and biological systems, communication networks and social sciences (see e.g. [8] for an extensive list of references). Systems with inhomogeneity described by multi-type populations, where the interaction between two particles depends on their types, have been proposed in social sciences [15], statistical mechanics [14], neurosciences [1], and others [9, 25]. In recent years, there have been an increasing attention on the study of mean field systems on large networks, including [3, 5, 10, 16–18, 26], where the majority of focus is on Erdős-Rényi random graphs. Among these, [26] allows the edge probability between two nodes to depend on independent random media variables associated with these two nodes, and [17] analyzes the mean field game on Erdős-Rényi random graphs.

We extend the study of mean field models to a much larger class of graphs. To put our work in context, let us describe the class of graphs that we are going to consider. We consider sequences of dense graphs that converge to a limit in an appropriate sense (see [23] and references therein). Roughly speaking, this limit theory treats a graph $\mathbb{G}_n$ on n vertices as a function $G_n \colon I \times I \to \mathbb{R}$, where $I := [0,1]$. This function is what we call a graphon. Then $\mathbb{G}_n$ is said to converge to the function G if and only if G_n converges to G in "cut metric" (see Section 2 for the definition). We consider mean field models on such converging graph sequences. The motivation for considering such graph sequences is that aside from its theoretical implications, many important graph models (both random and deterministic) have been shown to converge to a limit. See [2, 6, 13, 22] for many such examples. Unfortunately, the graph limit theory only works for dense graphs (graphs with order of n^2 many edges). To extend our study to the not-so-dense settings we also consider bond percolated models of graphs (see [7]).

We should point out that the use of graphons to analyze heterogeneous interaction in game theory emerged recently (see e.g. [11, 12, 27]). Among these, [27] analyzes static graphon games and the convergence of the n-player game with interaction network sampled from a given graphon. Static graphon games are considered in [12] and the convergence of the n-player game with general interaction network that converges to a given graphon is obtained. The diffusive dynamics for the states of the particles, with constant diffusion coefficients, is considered in [11] for continuum graphon mean field games. However, [11] does not address the convergence problem of the finite particle system to the limiting problem they analyze.

The goal of this work is to study the asymptotics of the diffusive particle system (1.1) with heterogeneous interaction, and their not-so-dense analogue in (1.2) below. Our first main result is the existence, uniqueness, continuity, and stability property (Proposition 2.1 and Theorem 2.1) for the limiting graphon particle system (2.1), consisting of a continuum of independent but non-identical nonlinear diffusions. Among these, the stability property (Theorem 2.1(c)) in particular says that the system solution converges in a suitable sense provided that the underlying graphon converges in the cut metric. The proof makes use of a coupling argument but challenge is two fold: First, the cut metric is in general very weak, in that the convergence $G_n \to G$ does not necessarily imply the L^2-convergence of G_n as operators on $I \times I$, namely one may not have $\int_{I\times I}[G_n(x,y) - G(x,y)]^2\,dx\,dy \to 0$. However, one could alternatively view G_n as operators from $L^\infty(I)$ to $L^1(I)$ that are continuous with respect to the cut metric (see Remark 2.1). This observation is actually an important building block of many proofs in this work. Second, the interaction in the graphon particle system (2.1) does not match with such a choice of operator, unless the coefficients $b(x,y)$ and $\sigma(x,y)$ could be decomposed as the product of functions of each variable. For this, a truncation and approximation argument is applied to these coefficients, and the associated errors are carefully analyzed (see Section 5.2).

The second main result is the convergence of the n particle system (1.1) to the graphon particle system (2.1), for a sequence of convergent underlying step graphons (graphons with blockwise constant values; see (3.2)). A law of large numbers (LLN) is established in Theorem 3.1, which says that the empirical measure of n particles in (1.1) converges in probability to the averaged distribution of a continuum particles in (2.1). The proof again relies crucially on a truncation and approximation argument applied to the system coefficients (see Lemma 6.1 and Section 6.2). In Theorem 3.2, we also obtain a precise particle-wise uniform convergence rate, when the underlying step graphons are sampled from a given graphon with a certain continuity property.

Our third main result is the analysis of the not-so-dense analogue of (1.1):

$$X_i^n(t) = X_{\frac{i}{n}}(0) + \int_0^t \frac{1}{n\beta_n}\sum_{j=1}^n \xi_{ij}^n b(X_i^n(s), X_j^n(s))\,ds + \int_0^t \sigma(X_i^n(s))\,dB_{\frac{i}{n}}(s), \quad i \in \{1,\ldots,n\}, \tag{1.2}$$

where $\beta_n \in (0,1]$ is some sequence of numbers that may converge to 0, and $\{\xi_{ij}^n\}$ are independent Bernoulli random variables with possibly vanishing probabilities (of order β_n) associated with the underlying step graphon. Similar to mean field systems on Erdős-Rényi random graphs [5, 17, 26], the strength of interaction here is scaled by the order of the number of neighbors (see Remark 4.1 for more explanations). In Theorem 4.1, we show a LLN that the limit of such systems is again given by a graphon particle system, provided that the underlying step graphons converge and $\lim_{n\to\infty} n\beta_n = \infty$. The main challenge lies in the heterogeneity of the system and the average of interactions of order ξ_{ij}^n/β_n that is unbounded in n. The unbounded interaction ξ_{ij}^n/β_n is taken care of in [5,17,26] using exchangeability. However, due to the lack of exchangeability here, a new approach is needed. Indeed, besides the application of coupling, truncation and approximation arguments, the key ingredient in the proof of Theorem 4.1 is (7.4) in Lemma 7.2, which shows that the expected effect of unbounded interactions ξ_{ij}^n/β_n on the coupled difference $|X_j^n - X_{\frac{j}{n}}|$ is roughly the same as $\mathbb{E}[\xi_{ij}^n/\beta_n]\mathbb{E}|X_j^n - X_{\frac{j}{n}}|$, up to some constant multiples and negligible errors. The proof of Lemma 7.2 applies a collection of change of measure arguments separately to each pair (resp. triplet) of certain auxiliary particles and the edge (resp. edges) connecting them. For each change of measure, the Radon-Nikodym derivative and the difference among pre-limit, limiting and auxiliary particles are carefully analyzed. Due to the technical application of the Girsanov's Theorem, the diffusion coefficient in (1.2) is taken to be state-dependent only. Lastly, we also obtain a precise rate of convergence in Theorem 4.2, when the underlying step graphons are sampled from a given graphon with a certain continuity property.

1.1. **Organization.** The paper is organized as follows. In Section 2 we analyze the graphon particle system (2.1). The existence and uniqueness is proved in Proposition 2.1. The continuity and stability of the system is presented in Theorem 2.1 in Section 2.1. A concrete example is given in Section 2.2. In Section 3 we study the convergence of the n particle system (1.1). The LLN is given in Theorem 3.1, and a precise rate of convergence is given in Theorem 3.2 under conditions. In Section 4 we study the convergence of the n particle system with not-so-dense interaction (1.2). The LLN is given in Theorem 4.1, and a precise rate of convergence is given in Theorem 4.2 under conditions. Sections 5, 6 and 7 are devoted to the proofs of results in Section 2, 3 and 4, respectively.

We close this section by introducing some frequently used notation.

1.2. **Notation.** Given a Polish space $\mathbb{S}$, denote by $\mathcal{P}(\mathbb{S})$ the space of probability measures on $\mathbb{S}$ endowed with the topology of weak convergence. For $\mu \in \mathcal{P}(\mathbb{S})$ and a μ-integrable function $f\colon \mathbb{S} \to \mathbb{R}$, let $\langle f, \mu\rangle := \int_{\mathbb{S}} f(x)\,\mu(dx)$. For $f\colon \mathbb{S} \to \mathbb{R}$, let $\|f\|_\infty := \sup_{x\in\mathbb{S}} |f(x)|$. The probability law of a random variable X will be denoted by $\mathcal{L}(X)$. Fix $T \in (0,\infty)$ and all processes will be considered over the time horizon $[0,T]$. Denote by $\mathbb{C}([0,T] : \mathbb{S})$ the space of continuous functions from $[0,T]$ to $\mathbb{S}$, endowed with the uniform topology. Let $\mathcal{C}_d := \mathbb{C}([0,T] : \mathbb{R}^d)$ and $\|x\|_{*,t} := \sup_{0\le s\le t} |x_s|$ for $x \in \mathcal{C}_d$ and $t \in [0,T]$. We will use κ to denote various constants in the paper and $\kappa(m)$ to emphasize the dependence on some parameter m. Their values may change from line to line. Expectations under $\mathbb{P}$ will be denoted by $\mathbb{E}$. To simplify the notation, we will usually write $\mathbb{E}[X^2]$ as $\mathbb{E}X^2$.

2. Graphon particle systems

We follow the notation used in [23, Chapters 7 and 8]. Let $I := [0,1]$. Denote by $\mathcal{G}$ the space of all bounded symmetric measurable functions $G\colon I \times I \to \mathbb{R}$. A graphon G is an element of $\mathcal{G}$ with $0 \le G \le 1$.

The cut norm on $\mathcal{G}$ is defined by

$$\|G\|_\square := \sup_{S,T\in\mathcal{B}(I)} \left| \int_{S\times T} G(u,v)\,du\,dv \right|,$$

and the corresponding cut metric and cut distance are defined by

$$d_\square(G_1,G_2) := \|G_1 - G_2\|_\square, \quad \delta_\square(G_1,G_2) := \inf_{\varphi\in S_I} \|G_1 - G_2^\varphi\|_\square,$$

where S_I denotes the set of all invertible measure preserving maps $I \to I$, and $G^\varphi(u,v) := G(\varphi(u),\varphi(v))$.

Remark 2.1. *We will also view a graphon G as an operator from $L^\infty(I)$ to $L^1(I)$ with the operator norm*

$$\|G\| := \|G\|_{\infty\to1} := \sup_{\|g\|_\infty\le1} \|Gg\|_1 = \sup_{\|g\|_\infty\le1} \int_I \left| \int_I G(u,v)g(v)\,dv \right| du.$$

From [23, Lemma 8.11] *it follows that if $\|G_n - G\|_\square \to 0$ for a sequence of graphons G_n, then $\|G_n - G\| \to 0$.*

Given a graphon G and an initial distribution $\mu(0) := (\mu_u(0) \in \mathcal{P}(\mathbb{R}^d) : u \in I)$, consider the following graphon particle system:

$$X_u(t) = X_u(0) + \int_0^t \int_I \int_{\mathbb{R}^d} b(X_u(s), x) G(u, v)\, \mu_{v,s}(dx)\, dv\, ds \\ + \int_0^t \int_I \int_{\mathbb{R}^d} \sigma(X_u(s), x) G(u, v)\, \mu_{v,s}(dx)\, dv\, dB_u(s), \;\; \mu_{u,t} = \mathcal{L}(X_u(t)), \;\; u \in I. \quad (2.1)$$

As introduced in Section 1, here b and σ are bounded and Lipschitz functions, $\{B_u : u \in I\}$ are i.i.d. d-dimensional Brownian motions, and $X_u(0)$ is a collection of independent $\mathbb{R}^d$-valued random variables, with law $\mu_u(0)$ for each $u \in I$, and independent of $\{B_u : u \in I\}$, defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$. We make the following assumption on the initial distribution.

Condition 2.1. *The map $I \ni u \mapsto \mu_u(0) = \mathcal{L}(X_u(0)) \in \mathcal{P}(\mathbb{R}^d)$ is measurable. Moreover, $\sup_{u \in I} \mathbb{E}|X_u(0)|^2 < \infty$.*

The following proposition gives well-posedness for the system (2.1). For graphon mean field games with diffusive dynamics, under certain suitable conditions, existence and uniqueness were shown in [11, Theorem 2]. But we provide a proof in Section 5.1 for completeness.

Proposition 2.1. *Suppose Condition 2.1 holds. There exists a unique pathwise solution to the graphon particle system (2.1). Moreover, the map $I \ni u \mapsto \mu_u \in \mathcal{P}(\mathcal{C}_d)$ is measurable.*

Remark 2.2. *We note that processes $\{X_u\}$ in (2.1) are independent but not identically distributed nonlinear diffusions. In particular, the Fokker–Planck equations for $\{X_u\}$ are nonlinear and fully coupled. In general, each X_u may not be a McKean–Vlasov process, as the probability law μ_u plays a negligible role in its evolution.*

In order to analyze the collection of probability laws $\mu = (\mu_u : u \in I)$, denote the product space of probability measures

$$\mathcal{M} := \{\nu = (\nu_u : u \in I) \in [\mathcal{P}(\mathcal{C}_d)]^I \mid I \ni u \mapsto \nu_u \in \mathcal{P}(\mathcal{C}_d) \text{ is measurable}\}.$$

For the convenience of analysis (see e.g. Remark 2.3), we make use of the following Wasserstein-2 metrics

$$W_2(\mu, \nu) := \left(\inf \left\{ \mathbb{E}|X - \tilde{X}|^2 : \mathcal{L}(X) = \mu, \mathcal{L}(\tilde{X}) = \nu \right\} \right)^{1/2}, \; \mu, \nu \in \mathcal{P}(\mathbb{R}^d), \quad (2.2)$$

$$W_{2,t}(\mu, \nu) := \left(\inf \left\{ \mathbb{E}\|X - \tilde{X}\|_{*,t}^2 : \mathcal{L}(X) = \mu, \mathcal{L}(\tilde{X}) = \nu \right\} \right)^{1/2}, \; t \in [0, T], \; \mu, \nu \in \mathcal{P}(\mathcal{C}_d), \quad (2.3)$$

$$W_{2,t}^{\mathcal{M}}(\mu, \nu) := \sup_{u \in I} W_{2,t}(\mu_u, \nu_u), \; t \in [0, T], \; \mu, \nu \in \mathcal{M}. \quad (2.4)$$

Remark 2.3. *From (2.2), (2.3) and (2.4) clearly we have*

$$W_2(\mu, \nu) \geq \sup_f \left| \int_{\mathbb{R}^d} f(x)\, \mu(dx) - \int_{\mathbb{R}^d} f(x)\, \nu(dx) \right|, \quad \mu, \nu \in \mathcal{P}(\mathbb{R}^d),$$

$$W_{2,t}(\mu, \nu) \geq \sup_f \left| \int_{\mathbb{R}^d} f(x)\, \mu_t(dx) - \int_{\mathbb{R}^d} f(x)\, \nu_t(dx) \right|, \quad \mu, \nu \in \mathcal{P}(\mathcal{C}_d),$$

$$W_{2,t}^{\mathcal{M}}(\mu, \nu) \geq \sup_{u \in I} \sup_f \left| \int_{\mathbb{R}^d} f(x)\, \mu_{u,t}(dx) - \int_{\mathbb{R}^d} f(x)\, \nu_{u,t}(dx) \right|, \quad \mu, \nu \in \mathcal{M},$$

for each $t \in [0,T]$, where the supremum is taken over all $f\colon \mathbb{R}^d \to \mathbb{R}$ such that the integrals exist and $|f(x)-f(\tilde{x})| \le |x-\tilde{x}|$ for $x,\tilde{x} \in \mathbb{R}^d$.

2.1. **Continuity and stability of the system.** In this section we are interested in establishing the continuity and stability properties for the graphon particle system (2.1). We usually make the following assumption on the initial distribution $\mu(0)$ and the graphon G.

Condition 2.2. *There exists a finite collection of intervals $\{I_i : i = 1, \dots, N\}$ for some $N \in \mathbb{N}$, such that $\cup_{i=1}^N I_i = I$ and for each $i \in \{1,\dots,N\}$:*

(a) The map $I_i \ni u \mapsto \mu_u(0) \in \mathcal{P}(\mathbb{R}^d)$ is continuous with respect to the W_2 metric.
(b) For each $u \in I_i$, there exists a subset $A_u \subset I$ such that $\lambda_I(A_u) = 0$ and $G(u,v)$ is continuous at $(u,v) \in I \times I$ for each $v \in I \setminus A_u$, where λ_I denotes the Lebesgue measure on I.

Remark 2.4. *Condition 2.2(b) holds naturally if G is continuous, or if G is continuous when restricted to each block $I_i \times I_j$. In particular, it holds for graphons such as $G(u,v) = \mathbf{1}_{[0,\frac{1}{2}]^2}(u,v)$ and $G(u,v) = \mathbf{1}_{\{|u-v|\le \frac{1}{4}\}}(u,v)$.*

Sometimes we may work with a special class of $\mu(0)$ and G having certain Lipschitz properties.

Condition 2.3. *There exists some $\kappa \in (0,\infty)$ and a finite collection of intervals $\{I_i : i = 1,\dots,N\}$ for some $N \in \mathbb{N}$, such that $\cup_{i=1}^N I_i = I$ and*

$$W_2(\mu_{u_1}(0), \mu_{u_2}(0)) \le \kappa |u_1 - u_2|, \quad u_1,u_2 \in I_i, \quad i \in \{1,\dots,N\},$$
$$|G(u_1,v_1) - G(u_2,v_2)| \le \kappa(|u_1-u_2| + |v_1-v_2|),\ (u_1,v_1),(u_2,v_2) \in I_i \times I_j,\ i,j \in \{1,\dots,N\}.$$

The following theorem gives continuity and stability of the graphon particle system (2.1). The proof is given in Section 5.2.

Theorem 2.1. *(a) (Continuity) Suppose Conditions 2.1 and 2.2 hold. Then for each $i \in \{1,\dots,N\}$, the map $I_i \ni u \mapsto \mu_u \in \mathcal{P}(\mathcal{C}_d)$ is continuous with respect to the $W_{2,T}$ metric.*
(b) (Lipschitz continuity) Suppose Conditions 2.1 and 2.3 hold. Then there exists some $\kappa \in (0,\infty)$ such that $W_{2,T}(\mu_u,\mu_v) \le \kappa |u-v|$ whenever $u,v \in I_i$ for some $i \in \{1,\dots,N\}$.
(c) (Stability) Suppose Condition 2.1 holds. Let μ^G be the probability law of (2.1) associated with G. The map $G \mapsto \mu^G$ is continuous in the sense that $\int_I [W_{2,T}(\mu_u^{G_n}, \mu_u^G)]^2\, du \to 0$ if a sequence of graphons $G_n \to G$ in the cut metric.

Remark 2.5. *(a) Theorem 2.1 (a,b) will be needed in Sections 3 and 4 to analyze the convergence of n-particle system with graphon mean field interactions.*
(b) Theorem 2.1(c) implies that the solution law to (2.1) depends on the underlying graphon G in a continuous manner. This and Proposition 2.1 together guarantees that the analysis of (2.1) is "well-posed" according to Hadamard's principle (cf. [4, Page 368] and [20, Page 38]).

2.2. **Some special graphon particle systems.** In this subsection we will introduce a special graphon G associated to which the system (2.1) is more tractable. In particular, consider the special case where G is blockwise constant (which arises as a limit of the stochastic block model), that is, there exists a finite collection of intervals $\{I_i : i = 1,\dots,N\}$ and constants $\{p_{ij} = p_{ji} \in [0,1] : i,j = 1,\dots,N\}$ for some $N \in \mathbb{N}$, such that $\cup_{i=1}^N I_i = I$ and

$$G(u,v) = p_{ij}, \quad (u,v) \in I_i \times I_j, \quad i,j \in \{1,\dots,N\}.$$

Due to the homogeneity in this case, the system (2.1) could be written in terms of just N representatives $u_i \in I_i$, $i \in \{1, \dots, N\}$:

$$X_{u_i}(t) = X_{u_i}(0) + \int_0^t \sum_{j=1}^N |I_j| p_{ij} \left(\int_{\mathbb{R}^d} b(X_{u_i}(s), x)\, \mu_{u_j,s}(dx) \right) ds$$
$$+ \int_0^t \sum_{j=1}^N |I_j| p_{ij} \left(\int_{\mathbb{R}^d} \sigma(X_{u_i}(s), x)\, \mu_{u_j,s}(dx) \right) dB_{u_i}(s), \quad \mu_{u_i,t} = \mathcal{L}(X_{u_i}(t)),$$

where $|A|$ denotes the Lebesgue measure of $A \subset I$. Note that this is simply a finite collection of multi-type McKean–Vlasov processes.

3. Mean-field particle systems on dense graphs

In this section, we consider a sequence of n interacting diffusions (1.1) with the strength of interaction governed by ξ_{ij}^n associated with some kernel G_n:

$$X_i^n(t) = X_{\frac{i}{n}}(0) + \int_0^t \frac{1}{n} \sum_{j=1}^n \xi_{ij}^n b(X_i^n(s), X_j^n(s))\, ds$$
$$+ \int_0^t \frac{1}{n} \sum_{j=1}^n \xi_{ij}^n \sigma(X_i^n(s), X_j^n(s))\, dB_{\frac{i}{n}}(s), \quad i \in \{1, \dots, n\}. \tag{3.1}$$

Here the pathwise existence and uniqueness of the solution is guaranteed by the boundedness and Lipschitz property of b and σ.

We make the following assumption on the strength of interaction ξ_{ij}^n and the associated kernel G_n.

Condition 3.1. *G_n is a step graphon, that is,*

$$G_n(u, v) = G_n\left(\frac{\lceil nu \rceil}{n}, \frac{\lceil nv \rceil}{n} \right), \quad \textit{for } (u, v) \in I \times I. \tag{3.2}$$

Moreover, $G_n \to G$ in the cut metric and

(i) either $\xi_{ij}^n = G_n(\frac{i}{n}, \frac{j}{n})$,
(ii) or $\xi_{ij}^n = \xi_{ji}^n =$Bernoulli$(G_n(\frac{i}{n}, \frac{j}{n}))$ independently for $1 \le i \le j \le n$, and independent of $\{B_u, X_u(0) : u \in I\}$.

Remark 3.1. *In general, if $\delta_\square(G_n, G) \to 0$ for a sequence of step graphons, then it follows from* [23, Theorem 11.59] *that $\|G_n - G\|_\square \to 0$, after suitable relabeling. Therefore we assume in Condition* 3.1 *that the convergence of G_n to G is in the cut metric $d_\square$ instead of the cut distance $\delta_\square$.*

The following convergence holds for the system (3.1). The proof is given in Section 6.1

Theorem 3.1. *Suppose Conditions* 2.1, 2.2 *and* 3.1 *hold. Then*

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E} \|X_i^n - X_{\frac{i}{n}}\|_{*,T}^2 \to 0 \tag{3.3}$$

and

$$\mu^n \to \bar{\mu} \textit{ in } \mathcal{P}(\mathcal{C}_d) \textit{ in probability} \tag{3.4}$$

as $n \to \infty$, where

$$\mu^n := \frac{1}{n}\sum_{i=1}^n \delta_{X_i^n}, \quad \bar{\mu} := \int_I \mu_u \, du.$$

If the interaction ξ_{ij}^n is sampled from a common graphon G, one could get a uniform rate of convergence.

Condition 3.2. *Suppose*

(i) either $\xi_{ij}^n = G(\frac{i}{n}, \frac{j}{n})$,

(ii) or $\xi_{ij}^n = \xi_{ji}^n =$Bernoulli$(G(\frac{i}{n}, \frac{j}{n}))$ independently for $1 \le i \le j \le n$, and independent of $\{B_u, X_u(0) : u \in I\}$.

The proof of the following rate of convergence is given in Section 6.3.

Theorem 3.2. *Suppose Conditions* 2.1*,* 2.3 *and* 3.2 *hold. Then there exists some $\kappa \in (0, \infty)$ such that*

$$\max_{i=1,\dots,n} \mathbb{E}\|X_i^n - X_{\frac{i}{n}}\|_{*,T}^2 \le \frac{\kappa}{n}. \tag{3.5}$$

4. Mean-field particle systems on not-so-dense graphs

In this section we consider a sequence of n interacting diffusions (1.2) with the strength of interaction governed by ξ_{ij}^n associated with some kernel G_n in a not-so-dense manner:

$$X_i^n(t) = X_{\frac{i}{n}}(0) + \int_0^t \frac{1}{n\beta_n}\sum_{j=1}^n \xi_{ij}^n b(X_i^n(s), X_j^n(s))\, ds + \int_0^t \sigma(X_i^n(s))\, dB_{\frac{i}{n}}(s), \quad i \in \{1, \dots, n\}. \tag{4.1}$$

Here b is bounded and Lipschitz, σ is bounded, Lipschitz and invertible with bounded inverse, and $\beta_n \in (0, 1]$ is some sparsity parameter.

We make the following assumption on the strength of interaction ξ_{ij}^n and the associated kernel G_n.

Condition 4.1. *G_n is a step graphon, that is,* (3.2) *holds. $G_n(u, v) = G_n(\frac{\lceil nu \rceil}{n}, \frac{\lceil nv \rceil}{n})$ for $(u, v) \in I \times I$. Moreover, $G_n \to G$ in the cut metric and $\xi_{ij}^n = \xi_{ji}^n =$Bernoulli$(\beta_n G_n(\frac{i}{n}, \frac{j}{n}))$ independently for $1 \le i \le j \le n$, and independent of $\{B_u, X_u(0) : u \in I\}$.*

Remark 4.1. *(a) The interaction in* (4.1) *is locally mean-field, where the strength of interaction between particles i and j is ξ_{ij}^n scaled down by $n\beta_n$, the order of number of neighbors of i or j.*

(b) The second assumption in Condition 4.1 *is also known as bond percolation, which has been studied for converging graph sequences in* [7]*. If we take $\lim_{n\to\infty} \beta_n = 0$ then graphs that we obtain are not-so-dense, in that for a graph $\mathbb{G}_n$ with an order of n^2 edges, the percolated graph will have approximately an order of $n^2\beta_n$ edges. Therefore β_n can be interpreted as the global sparsity parameter.*

The limiting graphon particle system is given by

$$\begin{aligned} X_u(t) = X_u(0) &+ \int_0^t \int_I \int_{\mathbb{R}^d} b(X_u(s), x) G(u, v)\, \mu_{v,s}(dx)\, dv\, ds \\ &+ \int_0^t \sigma(X_u(s))\, dB_u(s), \quad \mu_{u,t} = \mathcal{L}(X_u(t)), \quad u \in I. \end{aligned}$$

This is a special case of (2.1) and hence Proposition 2.1 and Theorem 2.1 still hold.

The following theorem gives a LLN for the system (4.1). The proof is given in Section 7.1.

Theorem 4.1. *Suppose Conditions* 2.1, 2.2 *and* 4.1 *hold. Suppose* $\lim_{n\to\infty} n\beta_n = \infty$. *Then*

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\|X_i^n - X_{\frac{i}{n}}\|_{*,T}^2 \to 0 \tag{4.2}$$

and

$$\mu^n \to \bar{\mu} \text{ in } \mathcal{P}(\mathcal{C}_d) \text{ in probability}$$

as $n \to \infty$, *where*

$$\mu^n := \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i^n}, \quad \bar{\mu} := \int_I \mu_u \, du.$$

If the interaction ξ_{ij}^n is sampled from a common graphon G, one could get a precise rate of convergence.

Condition 4.2. $\xi_{ij}^n = \xi_{ji}^n$ *=Bernoulli*$(\beta_n G(\frac{i}{n}, \frac{j}{n}))$ *independently for* $1 \le i \le j \le n$, *and independent of* $\{B_u, X_u(0) : u \in I\}$.

The proof of the following rate of convergence is given in Section 7.1.

Theorem 4.2. *Suppose Conditions* 2.1, 2.3 *and* 4.2 *hold. Suppose* $\liminf_{n\to\infty} n\beta_n > 0$. *Then for each* $q \in (1, \infty)$ *there exists some* $\kappa(q) \in (0, \infty)$ *such that*

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\|X_i^n - X_{\frac{i}{n}}\|_{*,T}^2 \le \frac{\kappa(q)}{(n\beta_n)^{1/q}}. \tag{4.3}$$

5. Proofs for Section 2

In this section we prove Proposition 2.1 and Theorem 2.1.

5.1. Proof of Proposition 2.1.

Define the map $\mathcal{M} \ni \mu \mapsto \Phi(\mu) \in [\mathcal{P}(\mathcal{C}_d)]^I$ by $\Phi(\mu) := (\mathcal{L}(X_u^\mu) : u \in I)$, where X_u^μ is the solution of

$$\begin{aligned} X_u^\mu(t) = X_u(0) &+ \int_0^t \int_I \int_{\mathbb{R}^d} b(X_u^\mu(s), x) G(u, v) \, \mu_{v,s}(dx) \, dv \, ds \\ &+ \int_0^t \int_I \int_{\mathbb{R}^d} \sigma(X_u^\mu(s), x) G(u, v) \, \mu_{v,s}(dx) \, dv \, dB_u(s). \end{aligned} \tag{5.1}$$

Note that the pathwise uniqueness of $\{X_u^\mu : u \in I\}$ is guaranteed by the bounded and Lipschitz properties of b and σ. We claim that

$$\text{the pathwise existence of } \{X_u^\mu : u \in I\} \text{ holds and } \Phi(\mu) \in \mathcal{M} \text{ for } \mu \in \mathcal{M}. \tag{5.2}$$

The proof of (5.2) is deferred to the end.

Next we claim that

$$W_{2,t}^{\mathcal{M}}(\Phi(\mu), \Phi(\nu)) \le \kappa \int_0^t W_{2,s}^{\mathcal{M}}(\mu, \nu) \, ds, \quad \mu, \nu \in \mathcal{M}. \tag{5.3}$$

To see this, consider the coupling

$$
\begin{aligned}
X_u^\mu(t) &= X_u(0) + \int_0^t \int_I \int_{\mathbb{R}^d} b(X_u^\mu(s), x) G(u,v)\, \mu_{v,s}(dx)\, dv\, ds \\
&\quad + \int_0^t \int_I \int_{\mathbb{R}^d} \sigma(X_u^\mu(s), x) G(u,v)\, \mu_{v,s}(dx)\, dv\, dB_u(s), \\
X_u^\nu(t) &= X_u(0) + \int_0^t \int_I \int_{\mathbb{R}^d} b(X_u^\nu(s), x) G(u,v)\, \nu_{v,s}(dx)\, dv\, ds \\
&\quad + \int_0^t \int_I \int_{\mathbb{R}^d} \sigma(X_u^\nu(s), x) G(u,v)\, \nu_{v,s}(dx)\, dv\, dB_u(s).
\end{aligned}
$$

It then follows from the Holder's inequality and Burkholder-Davis-Gundy inequality that

$$
\begin{aligned}
&\mathbb{E}\|X_u^\mu - X_u^\nu\|_{*,t}^2 \\
&\leq \kappa \mathbb{E} \int_0^t \int_I \left| \int_{\mathbb{R}^d} b(X_u^\mu(s), x) G(u,v)\, \mu_{v,s}(dx) - \int_{\mathbb{R}^d} b(X_u^\nu(s), x) G(u,v)\, \nu_{v,s}(dx) \right|^2 dv\, ds \\
&\quad + \kappa \mathbb{E} \int_0^t \int_I \left| \int_{\mathbb{R}^d} \sigma(X_u^\mu(s), x) G(u,v)\, \mu_{v,s}(dx) - \int_{\mathbb{R}^d} \sigma(X_u^\nu(s), x) G(u,v)\, \nu_{v,s}(dx) \right|^2 dv\, ds.
\end{aligned}
$$

By adding and subtracting terms, we have

$$
\begin{aligned}
&\left| \int_{\mathbb{R}^d} b(X_u^\mu(s), x) G(u,v)\, \mu_{v,s}(dx) - \int_{\mathbb{R}^d} b(X_u^\nu(s), x) G(u,v)\, \nu_{v,s}(dx) \right|^2 \\
&\leq 2 \left| \int_{\mathbb{R}^d} [b(X_u^\mu(s), x) - b(X_u^\nu(s), x)] G(u,v)\, \mu_{v,s}(dx) \right|^2 \\
&\quad + 2 \left| \int_{\mathbb{R}^d} b(X_u^\nu(s), x) G(u,v)\, [\mu_{v,s} - \nu_{v,s}](dx) \right|^2 \\
&\leq \kappa |X_u^\mu(s) - X_u^\nu(s)|^2 + \kappa [W_{2,s}^{\mathcal{M}}(\mu,\nu)]^2,
\end{aligned}
$$

where the last line uses the bounded and Lipschitz property of b and Remark 2.3. The same estimate holds when b is replaced by σ in the last display. It then follows from Gronwall's inequality that

$$
\mathbb{E}\|X_u^\mu - X_u^\nu\|_{*,t}^2 \leq \kappa \int_0^t [W_{2,s}^{\mathcal{M}}(\mu,\nu)]^2\, ds.
$$

Therefore the claim (5.3) holds.

Using the claim (5.3), we can immediately get pathwise uniqueness for the solution of (2.1). The pathwise existence also follows from (5.3) and a standard contraction argument (see e.g. [30, Section I.1]). To be precise, taking $\nu = (\mathcal{L}(Y_u) : u \in I)$ where $Y_u(t) \equiv X_u(0)$ for $u \in I$ and $t \in [0,T]$, iterating (5.3) gives

$$
W_{2,T}^{\mathcal{M}}(\Phi^{k+1}(\nu), \Phi^k(\nu)) \leq \kappa^k \frac{T^k}{k!} W_{2,T}^{\mathcal{M}}(\Phi(\nu), \nu), \quad k \in \mathbb{N}.
$$

Using Condition 2.1 and the boundedness of b and σ, one clearly has $W_{2,T}^{\mathcal{M}}(\Phi(\nu),\nu) < \infty$. Therefore $\Phi^k(\nu)$ is a Cauchy sequence, and converges to a fixed point μ of Φ. This gives the existence in law of the solution of (2.1), which together with the pathwise uniqueness gives the pathwise existence of the solution of (2.1). Note that $u \mapsto [\Phi^k(\nu)]_u$ is measurable as

$\Phi^k(\nu) \in \mathcal{M}$. Since $\mathcal{P}(\mathcal{C}_d)$ is Polish and $\lim_{k\to\infty} W_{2,T}^{\mathcal{M}}(\Phi^k(\nu), \mu) = 0$, we have the measurability of $I \ni u \mapsto \mu_u \in \mathcal{P}(\mathcal{C}_d)$ (cf. [19, Theorem 4.2.2]).

Finally we verify the claim (5.2). Fix $\mu \in \mathcal{M}$. Let $\tilde{X}_u^0(t) \equiv X_u(0)$ for $t \in [0,T]$ and $u \in I$. For $n \in \mathbb{N}$, let

$$\begin{aligned} \tilde{X}_u^n(t) = \tilde{X}_u^{n-1}(0) &+ \int_0^t \int_I \int_{\mathbb{R}^d} b(\tilde{X}_u^{n-1}(s), x) G(u,v)\, \mu_{v,s}(dx)\, dv\, ds \\ &+ \int_0^t \int_I \int_{\mathbb{R}^d} \sigma(\tilde{X}_u^{n-1}(s), x) G(u,v)\, \mu_{v,s}(dx)\, dv\, dB_u(s). \end{aligned} \tag{5.4}$$

Since $\mu \in \mathcal{M}$, it follows that $\tilde{X}_u^n$ and integrals in (5.4) are well-defined for all $u \in I$ and $n \in \mathbb{N}$. We will prove by induction that

$$I \ni u \mapsto \mathcal{L}(\tilde{X}_u^n, B_u) \in \mathcal{P}(\mathcal{C}_d \times \mathcal{C}_d) \text{ is measurable for each } n = 0, 1, \dots \tag{5.5}$$

By construction and Condition 2.1, (5.5) holds for $n = 0$. Next suppose (5.5) holds up to $k-1$ for some $k \in \mathbb{N}$. To complete the proof of (5.5), it suffices to show that

$$I \ni u \mapsto \mathcal{L}(\tilde{X}_u^k(t_1), B_u(t_1), \dots, \tilde{X}_u^k(t_m), B_u(t_m)) \in \mathcal{P}((\mathbb{R}^d \times \mathbb{R}^d)^m)$$

is measurable for all $0 \le t_1 \le \cdots \le t_m \le T$ and $m \in \mathbb{N}$. It then suffices to show that

$$I \ni u \mapsto \mathbb{E}\left[\prod_{i=1}^m \Big(f_i(\tilde{X}_u^k(t_i)) g_i(B_u(t_i)) \Big)\right] \in \mathbb{R}$$

is measurable for all $0 \le t_1 \le \cdots \le t_m \le T$, $m \in \mathbb{N}$ and bounded and continuous functions $\{f_i, g_i : i = 1, \dots, m\}$ on $\mathbb{R}^d$. Now consider the following auxiliary processes

$$\begin{aligned} \tilde{X}_u^{k,\delta}(t) = \tilde{X}_u^{k-1}(0) &+ \int_0^t \int_I \int_{\mathbb{R}^d} b(\tilde{X}_u^{k-1}(\lfloor \tfrac{s}{\delta} \rfloor \delta), x) G(u,v)\, \mu_{v,\lfloor \frac{s}{\delta} \rfloor \delta}(dx)\, dv\, ds \\ &+ \int_0^t \int_I \int_{\mathbb{R}^d} \sigma(\tilde{X}_u^{k-1}(\lfloor \tfrac{s}{\delta} \rfloor \delta), x) G(u,v)\, \mu_{v,\lfloor \frac{s}{\delta} \rfloor \delta}(dx)\, dv\, dB_u(s), \end{aligned}$$

where $\delta \in (0,1)$. Clearly, $\tilde{X}_u^{k,\delta}(t)$ converges to $\tilde{X}_u^k(t)$ in probability as $\delta \to 0$ for each $u \in I$. So it suffices to prove that

$$I \ni u \mapsto \mathbb{E}\left[\prod_{i=1}^m \Big(f_i(\tilde{X}_u^{k,\delta}(t_i)) g_i(B_u(t_i)) \Big)\right] \in \mathbb{R}$$

is measurable for all $0 \le t_1 \le \cdots \le t_m \le T$, $m \in \mathbb{N}$ and bounded and continuous functions $\{f_i, g_i : i = 1, \dots, m\}$ on $\mathbb{R}^d$. Fix $t \in [0,T]$. Since (5.5) holds for $k-1$, it further suffices to show that

$$\tilde{X}_u^{k,\delta}(t) = h(u, \tilde{X}_u^{k-1}, B_u)$$

for some measurable function $h \colon I \times \mathcal{C}_d \times \mathcal{C}_d \to \mathbb{R}$. Noting that $\tilde{X}_u^{k,\delta}$ is given in terms of finite sums of $\tilde{X}_u^{k-1}$ and B_u, we have that $h(u,\cdot,\cdot)$ is continuous on $\mathcal{C}_d \times \mathcal{C}_d$ for each $u \in I$, and that $h(\cdot, x, w)$ is measurable on I for each $(x,w) \in \mathcal{C}_d \times \mathcal{C}_d$. Therefore h is measurable and this verifies (5.5) by induction. Using the bounded and Lipschitz Properties of b and σ, from (5.4) we can get

$$\sup_{u\in I} \mathbb{E}\|\tilde{X}_u^{n+1} - \tilde{X}_u^n\|_{*,t}^2 \le \kappa \int_0^t \sup_{u \in I} \mathbb{E}\|\tilde{X}_u^n - \tilde{X}_u^{n-1}\|_{*,s}^2\, ds.$$

Therefore $\{\tilde{X}_u^n : n \in \mathbb{N}\}$ is Cauchy and converges uniformly in $u \in I$ in probability to some X_u^μ that satisfies (5.1). Since $\mathcal{P}(\mathcal{C}_d)$ is Polish, the measurability of $u \mapsto \mathcal{L}(\tilde{X}_u^n)$ then guarantees that $I \ni u \mapsto \mathcal{L}(X_u^\mu) \in \mathcal{P}(\mathcal{C}_d)$ is measurable (cf. [19, Theorem 4.2.2]). So Φ is actually a well-defined map from $\mathcal{M}$ to $\mathcal{M}$. This verifies (5.2) and completes the proof of Proposition 2.1. ∎

5.2. **Proof of Theorem 2.1.** (a) (b) Fix $u_1, u_2 \in I$. Consider the following diffusions:

$$\begin{aligned}
\tilde{X}_{u_1}(t) &= \tilde{X}_{u_1}(0) + \int_0^t \int_I \int_{\mathbb{R}^d} b(\tilde{X}_{u_1}(s), x) G(u_1, v)\, \mu_{v,s}(dx)\, dv\, ds \\
&\quad + \int_0^t \int_I \int_{\mathbb{R}^d} \sigma(\tilde{X}_{u_1}(s), x) G(u_1, v)\, \mu_{v,s}(dx)\, dv\, dB(s), \\
\tilde{X}_{u_2}(t) &= \tilde{X}_{u_2}(0) + \int_0^t \int_I \int_{\mathbb{R}^d} b(\tilde{X}_{u_2}(s), x) G(u_2, v)\, \mu_{v,s}(dx)\, dv\, ds \\
&\quad + \int_0^t \int_I \int_{\mathbb{R}^d} \sigma(\tilde{X}_{u_2}(s), x) G(u_2, v)\, \mu_{v,s}(dx)\, dv\, dB(s).
\end{aligned}$$

Here B is a d-dimensional Brownian motion independent of $\{\tilde{X}_{u_1}(0), \tilde{X}_{u_2}(0)\}$, $\mathcal{L}(\tilde{X}_{u_1}(0)) = \mu_{u_1}(0)$, $\mathcal{L}(\tilde{X}_{u_2}(0)) = \mu_{u_2}(0)$, but $\tilde{X}_{u_1}(0)$ and $\tilde{X}_{u_2}(0)$ may not be independent. From Proposition 2.1 we have $\mathcal{L}(\tilde{X}_{u_1}) = \mu_{u_1}$ and $\mathcal{L}(\tilde{X}_{u_2}) = \mu_{u_2}$. Also note that

$$\begin{aligned}
&\mathbb{E}\|\tilde{X}_{u_1} - \tilde{X}_{u_2}\|_{*,t}^2 \\
&\le \kappa \mathbb{E}|\tilde{X}_{u_1}(0) - \tilde{X}_{u_2}(0)|^2 \\
&\quad + \kappa \mathbb{E} \int_0^t \int_I \int_{\mathbb{R}^d} \left| b(\tilde{X}_{u_1}(s), x) G(u_1, v) - b(\tilde{X}_{u_2}(s), x) G(u_2, v) \right|^2 \mu_{v,s}(dx)\, dv\, ds \\
&\quad + \kappa \mathbb{E} \int_0^t \int_I \int_{\mathbb{R}^d} \left| \sigma(\tilde{X}_{u_1}(s), x) G(u_1, v) - \sigma(\tilde{X}_{u_2}(s), x) G(u_2, v) \right|^2 \mu_{v,s}(dx)\, dv\, ds \\
&\le \kappa \mathbb{E}|\tilde{X}_{u_1}(0) - \tilde{X}_{u_2}(0)|^2 + \kappa \mathbb{E} \int_0^t |\tilde{X}_{u_1}(s) - \tilde{X}_{u_2}(s)|^2\, ds + \kappa \int_I |G(u_1, v) - G(u_2, v)|^2\, dv,
\end{aligned}$$

where the last line follows on adding and subtracting terms and using the bounded and Lipschitz properties of b and σ. It then follows from Gronwall's inequality that

$$[W_{2,T}(\mu_{u_1}, \mu_{u_2})]^2 \le \mathbb{E}\|\tilde{X}_{u_1} - \tilde{X}_{u_2}\|_{*,T}^2 \le \kappa \mathbb{E}|\tilde{X}_{u_1}(0) - \tilde{X}_{u_2}(0)|^2 + \kappa \int_I |G(u_1, v) - G(u_2, v)|^2\, dv.$$

Taking the infimum over $\mathcal{L}(\tilde{X}_{u_1}(0)) = \mu_{u_1}(0)$ and $\mathcal{L}(\tilde{X}_{u_2}(0)) = \mu_{u_2}(0)$, we have

$$[W_{2,T}(\mu_{u_1}, \mu_{u_2})]^2 \le \kappa [W_2(\mu_{u_1}(0), \mu_{u_2}(0))]^2 + \kappa \int_I |G(u_1, v) - G(u_2, v)|^2\, dv.$$

Part (a) and Part (b) then follow from Condition 2.2 and Condition 2.3, respectively.

(c) Fix $G_n \to G$ in the cut metric as $n \to \infty$. Let X^{G_n}, μ^{G_n} (resp. X^G, μ^G) be the solution of (2.1) associated with the graphon G_n (resp. G). Fix $t \in [0,T]$. Note that

$$\begin{aligned}\int_I \mathbb{E}\|X_u^{G_n} - X_u^G\|_{*,t}^2\,du \le \kappa \int_0^t \int_I \mathbb{E}\Big|\int_I \int_{\mathbb{R}^d} b(X_u^{G_n}(s),x)G_n(u,v)\,\mu_{v,s}^{G_n}(dx)\,dv \\ - \int_I \int_{\mathbb{R}^d} b(X_u^G(s),x)G(u,v)\,\mu_{v,s}^G(dx)\,dv\Big|^2\,du\,ds \\ + \kappa \int_0^t \int_I \mathbb{E}\Big|\int_I \int_{\mathbb{R}^d} \sigma(X_u^{G_n}(s),x)G_n(u,v)\,\mu_{v,s}^{G_n}(dx)\,dv \\ - \int_I \int_{\mathbb{R}^d} \sigma(X_u^G(s),x)G(u,v)\,\mu_{v,s}^G(dx)\,dv\Big|^2\,du\,ds.\end{aligned} \tag{5.6}$$

We will analyze the first integrand above for fixed $s \in [0,t]$, and the analysis for σ is similar. By adding and subtracting terms, we have

$$\begin{aligned}&\int_I \mathbb{E}\Big|\int_I \int_{\mathbb{R}^d} b(X_u^{G_n}(s),x)G_n(u,v)\,\mu_{v,s}^{G_n}(dx)\,dv - \int_I \int_{\mathbb{R}^d} b(X_u^G(s),x)G(u,v)\,\mu_{v,s}^G(dx)\,dv\Big|^2\,du \\ &\le \kappa \int_I \int_I \mathbb{E}\Big|\int_{\mathbb{R}^d}[b(X_u^{G_n}(s),x) - b(X_u^G(s),x)]G_n(u,v)\,\mu_{v,s}^{G_n}(dx)\Big|^2\,dv\,du \\ &\quad + \kappa \int_I \int_I \mathbb{E}\Big|\int_{\mathbb{R}^d} b(X_u^G(s),x)G_n(u,v)\,[\mu_{v,s}^{G_n} - \mu_{v,s}^G](dx)\Big|^2\,dv\,du \\ &\quad + \kappa \int_I \mathbb{E}\Big|\int_I \int_{\mathbb{R}^d} b(X_u^G(s),x)[G_n(u,v) - G(u,v)]\,\mu_{v,s}^G(dx)\,dv\Big|^2\,du \\ &=: \kappa\left(\mathcal{J}_s^{n,1} + \mathcal{J}_s^{n,2} + \mathcal{J}_s^{n,3}\right).\end{aligned} \tag{5.7}$$

Now we analyze each term $\mathcal{J}_s^{n,k}$, $k = 1,2,3$. Using the Lipschitz property of b we have

$$\mathcal{J}_s^{n,1} \le \kappa \int_I \mathbb{E}|X_u^{G_n}(s) - X_u^G(s)|^2\,du. \tag{5.8}$$

Using Remark 2.3 and the bounded and Lipschitz property of b we have

$$\mathcal{J}_s^{n,2} \le \kappa \int_I [W_{2,s}(\mu_v^{G_n}, \mu_v^G)]^2\,dv. \tag{5.9}$$

For the last term $\mathcal{J}_s^{n,3}$, first note that

$$\sup_{u\in I} \mathbb{E}\|X_u^G\|_{*,T} < \infty \tag{5.10}$$

by Condition 2.1 and the boundedness of b and σ. Fix $M \in (0,\infty)$ and write

$$b_M(x,y) := b(x,y)\mathbf{1}_{\{|x|\le M, |y|\le M\}}. \tag{5.11}$$

Since b is bounded and Lipschitz, it follows from [29, Corollary 2 of Theorem 3.1] that there exist some $m \in \mathbb{N}$ and polynomials

$$\tilde{b}_m(x,y) := \sum_{k=1}^m a_k(x)c_k(y)\mathbf{1}_{\{|x|\le M, |y|\le M\}}, \tag{5.12}$$

where a_k and c_k are polynomials for each $k = 1,\dots,m$, such that

$$|b_M(x,y) - \tilde{b}_m(x,y)| \le 1/M. \tag{5.13}$$

By the boundedness of b and adding and subtracting terms, we have

$$\begin{aligned}
\mathcal{J}_s^{n,3} &\le \kappa \int_I \mathbb{E}\left| \int_I \int_{\mathbb{R}^d} b(X_u^G(s),x)[G_n(u,v)-G(u,v)]\, \mu_{v,s}^G(dx)\, dv \right| du \\
&\le \kappa \int_I \int_I \mathbb{E}\left| \int_{\mathbb{R}^d} [b(X_u^G(s),x) - b_M(X_u^G(s),x)][G_n(u,v)-G(u,v)]\, \mu_{v,s}^G(dx) \right| dv\, du \\
&+ \kappa \int_I \int_I \mathbb{E}\left| \int_{\mathbb{R}^d} [b_M(X_u^G(s),x) - \tilde{b}_m(X_u^G(s),x)][G_n(u,v)-G(u,v)]\, \mu_{v,s}^G(dx) \right| dv\, du \\
&+ \kappa \int_I \mathbb{E}\left| \int_I \int_{\mathbb{R}^d} \tilde{b}_m(X_u^G(s),x)[G_n(u,v)-G(u,v)]\, \mu_{v,s}^G(dx)\, dv \right| du \\
&=: \kappa \sum_{k=1}^3 \mathcal{J}_s^{n,3,k}.
\end{aligned} \tag{5.14}$$

Next we analyze each term $\mathcal{J}_s^{n,3,k}$, $k=1,2,3$. For $\mathcal{J}_s^{n,3,1}$, using (5.10) and (5.11) we have

$$\mathcal{J}_s^{n,3,1} \le \kappa \int_I \int_I \mathbb{E} \int_{\mathbb{R}^d} [\mathbf{1}_{\{|X_u^G(s)|>M\}} + \mathbf{1}_{\{|x|>M\}}]\, \mu_{v,s}^G(dx)\, dv\, du \le \frac{\kappa}{M}. \tag{5.15}$$

For $\mathcal{J}_s^{n,3,2}$, using (5.13) we have

$$\mathcal{J}_s^{n,3,2} \le \frac{\kappa}{M}. \tag{5.16}$$

For $\mathcal{J}_s^{n,3,3}$, using the definition of $\tilde{b}_m$ in (5.12) we have

$$\begin{aligned}
\mathcal{J}_s^{n,3,3} &\le \kappa \sum_{k=1}^m \int_I \left| \int_I [G_n(u,v)-G(u,v)] \left[\int_{\mathbb{R}^d} c_k(x) \mathbf{1}_{\{|x|\le M\}} \mu_{v,s}^G(dx) \right] dv \right| du \\
&\le \kappa(M) \|G_n - G\|,
\end{aligned} \tag{5.17}$$

where $\kappa(M)$ is some constant that depends on M but not on n. Combining (5.6)–(5.9) and (5.14)–(5.17) with Gronwall's inequality, we have

$$\begin{aligned}
\int_I [W_{2,t}(\mu_u^{G_n}, \mu_u^G)]^2\, du &\le \int_I \mathbb{E}\|X_u^{G_n} - X_u^G\|_{*,t}^2\, du \\
&\le \kappa \left(\int_0^t \int_I [W_{2,s}(\mu_u^{G_n}, \mu_u^G)]^2\, du\, ds + \frac{1}{M} + \kappa(M)\|G_n - G\| \right).
\end{aligned}$$

It then follows from the Gronwall's inequality again that

$$\int_I [W_{2,t}(\mu_u^{G_n}, \mu_u^G)]^2\, du \le \kappa \left(\frac{1}{M} + \kappa(M)\|G_n - G\| \right).$$

Since $G_n \to G$ in the cut metric, from Remark 2.1 we have $\|G_n - G\| \to 0$ as $n \to \infty$. Therefore, by taking $\limsup_{n\to\infty}$ and then $\limsup_{M\to\infty}$ in the last display, we have the desired result. This gives part (c) and completes the proof of Theorem 2.1. ■

6. Proofs for Section 3

In this section we prove Theorems 3.1 and 3.2.

6.1. **Proof of Theorem 3.1.** For the first assertion, fix $t \in [0,T]$.

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n \mathbb{E}\|X_i^n - X_{\frac{i}{n}}\|_{*,t}^2 \\
&\le \kappa \int_0^t \left[\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left| \frac{1}{n}\sum_{j=1}^n \xi_{ij}^n b(X_i^n(s), X_j^n(s)) - \int_I \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v)\, \mu_{v,s}(dx)\, dv \right|^2 \right] ds \\
&\quad + \kappa \int_0^t \left[\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left| \frac{1}{n}\sum_{j=1}^n \xi_{ij}^n \sigma(X_i^n(s), X_j^n(s)) - \int_I \int_{\mathbb{R}^d} \sigma(X_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v)\, \mu_{v,s}(dx)\, dv \right|^2 \right] ds.
\end{aligned} \tag{6.1}$$

We will analyze the first integrand above for fixed $s \in [0,t]$, and the analysis for σ is similar. By adding and subtracting terms, we have

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left| \frac{1}{n}\sum_{j=1}^n \xi_{ij}^n b(X_i^n(s), X_j^n(s)) - \int_I \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v)\, \mu_{v,s}(dx)\, dv \right|^2 \\
&\le \frac{3}{n}\sum_{i=1}^n \mathbb{E}\left| \frac{1}{n}\sum_{j=1}^n \xi_{ij}^n \left(b(X_i^n(s), X_j^n(s)) - b(X_{\frac{i}{n}}(s), X_{\frac{j}{n}}(s)) \right) \right|^2 \\
&\quad + \frac{3}{n}\sum_{i=1}^n \mathbb{E}\left| \frac{1}{n}\sum_{j=1}^n \left(\xi_{ij}^n b(X_{\frac{i}{n}}(s), X_{\frac{j}{n}}(s)) - \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{j}{n})\, \mu_{\frac{j}{n},s}(dx) \right) \right|^2 \\
&\quad + \frac{3}{n}\sum_{i=1}^n \mathbb{E}\left| \frac{1}{n}\sum_{j=1}^n \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{j}{n})\, \mu_{\frac{j}{n},s}(dx) - \int_I \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v)\, \mu_{v,s}(dx)\, dv \right|^2 \\
&=: 3\left(\mathcal{T}_s^{n,1} + \mathcal{T}_s^{n,2} + \mathcal{T}_s^{n,3} \right).
\end{aligned} \tag{6.2}$$

For each term $\mathcal{T}_s^{n,k}$, $k = 1, 2, 3$, we have the following key estimates, whose proof will be given in Section 6.2.

Lemma 6.1. *Suppose the conditions in Theorem* 3.1 *hold. Then there exists some $\kappa \in (0,\infty)$ such that the following holds for each $s \in [0,T]$:*

$$\mathcal{T}_s^{n,1} \le \frac{\kappa}{n}\sum_{i=1}^n \mathbb{E}|X_i^n(s) - X_{\frac{i}{n}}(s)|^2, \tag{6.3}$$

$$\mathcal{T}_s^{n,2} \le \frac{\kappa}{n}, \tag{6.4}$$

$$\lim_{n\to\infty} \int_0^T \mathcal{T}_s^{n,3}\, ds = 0. \tag{6.5}$$

Completing the proof of Theorem 3.1: Combining (6.1)–(6.4) gives

$$\frac{1}{n}\sum_{i=1}^n \mathbb{E}\|X_i^n - X_{\frac{i}{n}}\|_{*,t}^2 \le \kappa \int_0^t \frac{1}{n}\sum_{i=1}^n \mathbb{E}\|X_i^n - X_{\frac{i}{n}}\|_{*,s}^2\, ds + \frac{\kappa}{n} + \kappa \int_0^t \mathcal{T}_s^{n,3}\, ds.$$

Using the Gronwall's inequality and (6.5) we have (3.3).

Next we prove the second assertion in Theorem 3.1. Let

$$\bar{\mu}^n := \frac{1}{n}\sum_{i=1}^n \delta_{X_{\frac{i}{n}}}, \quad \tilde{\mu}^n := \frac{1}{n}\sum_{i=1}^n \mu_{\frac{i}{n}}.$$

Denote by d_{BL} the bounded Lipschitz metric on $\mathcal{P}(\mathcal{C}_d)$, that is,

$$d_{BL}(\mu,\nu) := \sup_{f\in\mathbb{B}_1} |\langle f, \mu-\nu\rangle|, \quad \mu,\nu \in \mathcal{P}(\mathcal{C}_d),$$

where $\mathbb{B}_1$ is the collection of all Lipschitz functions on $\mathcal{C}_d$ that are bounded by 1 with Lipschitz constant also bounded by 1. Note that

$$\begin{aligned}\mathbb{E}d_{BL}(\mu^n,\bar{\mu}^n) &\le \mathbb{E}\left[\sup_{f\in\mathbb{B}_1}|\langle f,\mu^n-\bar{\mu}^n\rangle|\right] \le \mathbb{E}\left[\sup_{f\in\mathbb{B}_1}\frac{1}{n}\sum_{i=1}^n |f(X_i^n)-f(X_{\frac{i}{n}})|\right] \\ &\le \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \|X_i^n - X_{\frac{i}{n}}\|_{*,T}\right] \to 0\end{aligned}$$

by (3.3). For each bounded and continuous function f on $\mathcal{C}_d$, using the independence of $\{X_{\frac{i}{n}}\}$, we have

$$\mathbb{E}\left(\langle f,\bar{\mu}^n\rangle - \langle f,\tilde{\mu}^n\rangle\right)^2 = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n\Big(f(X_{\frac{i}{n}})-\mathbb{E}f(X_{\frac{i}{n}})\Big)\right)^2 \le \frac{\|f\|_\infty^2}{n} \to 0.$$

From Condition 2.2 and Theorem 2.1(a) we have

$$\langle f,\tilde{\mu}^n\rangle - \langle f,\bar{\mu}\rangle = \frac{1}{n}\sum_{i=1}^n \langle f,\mu_{\frac{i}{n}}\rangle - \int_I \langle f,\mu_u\rangle\,du \to 0.$$

Combining these three estimates gives (3.4) and completes the proof of Theorem 3.1. ∎

6.2. **Proof of Lemma 6.1.** Fix $s\in[0,T]$. For $\mathcal{T}_s^{n,1}$, using the Lipschitz property of b we have

$$\begin{aligned}\mathcal{T}_s^{n,1} &\le \kappa\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[\frac{1}{n}\sum_{j=1}^n\Big(|X_i^n(s)-X_{\frac{i}{n}}(s)|^2 + |X_j^n(s)-X_{\frac{j}{n}}(s)|^2\Big)\right] \\ &\le \kappa\frac{1}{n}\sum_{i=1}^n \mathbb{E}|X_i^n(s)-X_{\frac{i}{n}}(s)|^2.\end{aligned}$$

This gives (6.3).

For $\mathcal{T}_s^{n,2}$, using a weak LLN type argument, we have

$$\begin{aligned}
\mathcal{T}_s^{n,2} &= \frac{1}{n}\sum_{i=1}^n \frac{1}{n^2}\sum_{j=1}^n\sum_{k=1}^n \mathbb{E}\Bigg[\Bigg(\xi_{ij}^n b(X_{\frac{i}{n}}(s), X_{\frac{j}{n}}(s)) - \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{j}{n})\,\mu_{\frac{j}{n},s}(dx)\Bigg) \\
&\qquad \cdot \Bigg(\xi_{ik}^n b(X_{\frac{i}{n}}(s), X_{\frac{k}{n}}(s)) - \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{k}{n})\,\mu_{\frac{k}{n},s}(dx)\Bigg)\Bigg] \\
&= \frac{1}{n}\sum_{i=1}^n \frac{1}{n^2}\sum_{j=1}^n\sum_{k\in\{i,j\}} \mathbb{E}\Bigg[\Bigg(\xi_{ij}^n b(X_{\frac{i}{n}}(s), X_{\frac{j}{n}}(s)) - \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{j}{n})\,\mu_{\frac{j}{n},s}(dx)\Bigg) \\
&\qquad \cdot \Bigg(\xi_{ik}^n b(X_{\frac{i}{n}}(s), X_{\frac{k}{n}}(s)) - \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{k}{n})\,\mu_{\frac{k}{n},s}(dx)\Bigg)\Bigg] \\
&\le \frac{\kappa}{n},
\end{aligned}$$

where the second equality follows from the observation that the expectation is zero whenever $k \notin \{i,j\}$ by Condition 3.1 and the independence of $\{X_{\frac{i}{n}}\}$ and $\{\xi_{ij}^n\}$, and the inequality uses the boundedness of b, ξ_{ij}^n, G_n. This gives (6.4).

The analysis of $\mathcal{T}_s^{n,3}$ is similar to that of $\mathcal{J}_s^{n,3}$ in the proof of Theorem 2.1(c) but is more involved. From Condition 2.1 it follows that (5.10) holds. Fix $M \in (0,\infty)$ and let b_M and $\tilde{b}_m$ be defined as in (5.11) and (5.12), such that (5.13) holds. By adding and subtracting terms and using the boundedness of b, we have

$$\begin{aligned}
\mathcal{T}_s^{n,3} &\le \frac{\kappa}{n}\sum_{i=1}^n \mathbb{E}\left|\frac{1}{n}\sum_{j=1}^n \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{j}{n})\,\mu_{\frac{j}{n},s}(dx) - \int_I\int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v)\,\mu_{v,s}(dx)\,dv\right| \\
&\le \frac{\kappa}{n}\sum_{i=1}^n \mathbb{E}\left|\frac{1}{n}\sum_{j=1}^n \int_{\mathbb{R}^d} [b(X_{\frac{i}{n}}(s), x) - b_M(X_{\frac{i}{n}}(s), x)] G_n(\frac{i}{n}, \frac{j}{n})\,\mu_{\frac{j}{n},s}(dx)\right| \\
&\quad + \frac{\kappa}{n}\sum_{i=1}^n \mathbb{E}\left|\int_I\int_{\mathbb{R}^d} [b(X_{\frac{i}{n}}(s), x) - b_M(X_{\frac{i}{n}}(s), x)] G(\frac{i}{n}, v)\,\mu_{v,s}(dx)\,dv\right| \\
&\quad + \frac{\kappa}{n}\sum_{i=1}^n \mathbb{E}\left|\frac{1}{n}\sum_{j=1}^n \int_{\mathbb{R}^d} [b_M(X_{\frac{i}{n}}(s), x) - \tilde{b}_m(X_{\frac{i}{n}}(s), x)] G_n(\frac{i}{n}, \frac{j}{n})\,\mu_{\frac{j}{n},s}(dx)\right| \\
&\quad + \frac{\kappa}{n}\sum_{i=1}^n \mathbb{E}\left|\int_I\int_{\mathbb{R}^d} [b_M(X_{\frac{i}{n}}(s), x) - \tilde{b}_m(X_{\frac{i}{n}}(s), x)] G(\frac{i}{n}, v)\,\mu_{v,s}(dx)\,dv\right| \\
&\quad + \frac{\kappa}{n}\sum_{i=1}^n \mathbb{E}\Bigg|\frac{1}{n}\sum_{j=1}^n \int_{\mathbb{R}^d} \tilde{b}_m(X_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{j}{n})\,\mu_{\frac{j}{n},s}(dx) \\
&\qquad - \int_I\int_{\mathbb{R}^d} \tilde{b}_m(X_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v)\,\mu_{v,s}(dx)\,dv\Bigg| \\
&=: \kappa\sum_{k=1}^5 \mathcal{T}_s^{n,3,k}.
\end{aligned} \tag{6.6}$$

Next we analyze each term. For $\mathcal{T}_s^{n,3,1}$ and $\mathcal{T}_s^{n,3,2}$, using (5.10) and (5.11) we have

$$\mathcal{T}_s^{n,3,1} \le \kappa \frac{1}{n} \sum_{i=1}^n \mathbb{E}\left[\frac{1}{n}\sum_{j=1}^n \int_{\mathbb{R}^d} \left[\mathbf{1}_{\{|X_{\frac{i}{n}}(s)|>M\}} + \mathbf{1}_{\{|x|>M\}}\right] \mu_{\frac{j}{n},s}(dx)\right] \le \frac{\kappa}{M}, \tag{6.7}$$

$$\mathcal{T}_s^{n,3,2} \le \kappa \frac{1}{n} \sum_{i=1}^n \mathbb{E} \int_I \int_{\mathbb{R}^d} \left[\mathbf{1}_{\{|X_{\frac{i}{n}}(s)|>M\}} + \mathbf{1}_{\{|x|>M\}}\right] \mu_{v,s}(dx)\, dv \le \frac{\kappa}{M}. \tag{6.8}$$

For $\mathcal{T}_s^{n,3,3}$ and $\mathcal{T}_s^{n,3,4}$, using (5.13) we have

$$\mathcal{T}_s^{n,3,3} \le \frac{\kappa}{M}, \quad \mathcal{T}_s^{n,3,4} \le \frac{\kappa}{M}. \tag{6.9}$$

For $\mathcal{T}_s^{n,3,5}$, using the step graphon structure (3.2) of G_n we have

$$\begin{aligned}
\mathcal{T}_s^{n,3,5} &= \int_I \mathbb{E}\left|\int_I \int_{\mathbb{R}^d} \tilde{b}_m(X_{\frac{\lceil nu\rceil}{n}}(s), x) G_n(u,v)\, \mu_{\frac{\lceil nv\rceil}{n},s}(dx)\, dv\right. \\
&\qquad \left. - \int_I \int_{\mathbb{R}^d} \tilde{b}_m(X_{\frac{\lceil nu\rceil}{n}}(s), x) G(\frac{\lceil nu\rceil}{n}, v)\, \mu_{v,s}(dx)\, dv\right| du \\
&\le \int_I \mathbb{E}\left|\int_I \int_{\mathbb{R}^d} \tilde{b}_m(X_{\frac{\lceil nu\rceil}{n}}(s), x) \left[G_n(u,v) - G(u,v)\right] \mu_{\frac{\lceil nv\rceil}{n},s}(dx)\, dv\right| du \\
&\quad + \int_I \mathbb{E}\left|\int_I \int_{\mathbb{R}^d} \tilde{b}_m(X_{\frac{\lceil nu\rceil}{n}}(s), x) G(u,v)\, \mu_{\frac{\lceil nv\rceil}{n},s}(dx)\, dv\right. \\
&\qquad \left. - \int_I \int_{\mathbb{R}^d} \tilde{b}_m(X_{\frac{\lceil nu\rceil}{n}}(s), x) G(\frac{\lceil nu\rceil}{n}, v)\, \mu_{v,s}(dx)\, dv\right| du \\
&=: \mathcal{T}_s^{n,3,6} + \mathcal{T}_s^{n,3,7}.
\end{aligned} \tag{6.10}$$

For $\mathcal{T}_s^{n,3,6}$, using the definition of $\tilde{b}_m$ in (5.12) we have

$$\begin{aligned}
\mathcal{T}_s^{n,3,6} &\le \kappa \sum_{k=1}^m \int_I \left| \int_I \left[G_n(u,v) - G(u,v)\right] \left[\int_{\mathbb{R}^d} c_k(x) \mathbf{1}_{\{\|x\|\le M\}} \mu_{\frac{\lceil nv\rceil}{n},s}(dx)\right] dv \right| du \\
&\le \kappa(M) \|G_n - G\|,
\end{aligned}$$

where $\kappa(M)$ depends on M but not on n. It then follows from Remark 2.1 that

$$\lim_{n\to\infty} \int_0^T \mathcal{T}_s^{n,3,6}\, ds = 0. \tag{6.11}$$

For $\mathcal{T}_s^{n,3,7}$, using Condition 2.2 and Theorem 2.1(a) we see that the integrand goes to 0 for almost every pair $(u,v) \in I \times I$ as $n \to \infty$, and hence

$$\lim_{n\to\infty} \int_0^T \mathcal{T}_s^{n,3,7}\, ds = 0. \tag{6.12}$$

Combining (6.6)–(6.12) gives

$$\limsup_{n\to\infty} \int_0^T \mathcal{T}_s^{n,3}\, ds \le \frac{\kappa}{M}.$$

Further taking $\limsup_{M\to\infty}$ gives (6.5). This completes the proof of Lemma 6.1. ∎

6.3. **Proof of Theorem 3.2.** Fix $t \in [0,T]$ and $i \in \{1,\dots,n\}$. Similar to the proof of Theorem 3.1, we have

$$\begin{aligned}
&\mathbb{E}\|X_i^n - X_{\frac{i}{n}}\|_{*,t}^2 \\
&\le \kappa \mathbb{E}\int_0^t \left| \frac{1}{n}\sum_{j=1}^n \xi_{ij}^n b(X_i^n(s), X_j^n(s)) - \int_I \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v)\, \mu_{v,s}(dx)\, dv \right|^2 ds \\
&\quad + \kappa \mathbb{E}\int_0^t \left| \frac{1}{n}\sum_{j=1}^n \xi_{ij}^n \sigma(X_i^n(s), X_j^n(s)) - \int_I \int_{\mathbb{R}^d} \sigma(X_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v)\, \mu_{v,s}(dx)\, dv \right|^2 ds. \qquad (6.13)
\end{aligned}$$

We will analyze the first integrand above for fixed $s \in [0,t]$, and the analysis for σ is similar. By adding and subtracting terms, we have

$$\begin{aligned}
&\mathbb{E}\left| \frac{1}{n}\sum_{j=1}^n \xi_{ij}^n b(X_i^n(s), X_j^n(s)) - \int_I \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v)\, \mu_{v,s}(dx)\, dv \right|^2 \\
&\le 3\mathbb{E}\left| \frac{1}{n}\sum_{j=1}^n \xi_{ij}^n \left(b(X_i^n(s), X_j^n(s)) - b(X_{\frac{i}{n}}(s), X_{\frac{j}{n}}(s)) \right) \right|^2 \\
&\quad + 3\mathbb{E}\left| \frac{1}{n}\sum_{j=1}^n \left(\xi_{ij}^n b(X_{\frac{i}{n}}(s), X_{\frac{j}{n}}(s)) - \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G(\frac{i}{n}, \frac{j}{n})\, \mu_{\frac{j}{n},s}(dx) \right) \right|^2 \\
&\quad + 3\mathbb{E}\left| \frac{1}{n}\sum_{j=1}^n \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G(\frac{i}{n}, \frac{j}{n})\, \mu_{\frac{j}{n},s}(dx) - \int_I \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v)\, \mu_{v,s}(dx)\, dv \right|^2 \\
&=: 3\left(\tilde{\mathcal{T}}_s^{n,1} + \tilde{\mathcal{T}}_s^{n,2} + \tilde{\mathcal{T}}_s^{n,3} \right). \qquad (6.14)
\end{aligned}$$

For $\tilde{\mathcal{T}}_s^{n,1}$, using the Lipschitz property of b we have

$$\tilde{\mathcal{T}}_s^{n,1} \le \kappa \mathbb{E}\left[\frac{1}{n}\sum_{j=1}^n \left(|X_i^n(s) - X_{\frac{i}{n}}(s)|^2 + |X_j^n(s) - X_{\frac{j}{n}}(s)|^2 \right) \right] \le 2\kappa \max_{i=1,\dots,n} \mathbb{E}|X_i^n(s) - X_{\frac{i}{n}}(s)|^2. \qquad (6.15)$$

For $\tilde{\mathcal{T}}_s^{n,2}$, using Condition 3.2, the independence of $\{X_{\frac{i}{n}}\}$ and $\{\xi_{ij}^n\}$, the boundedness of b, and a weak LLN type argument, we have

$$\tilde{\mathcal{T}}_s^{n,2} \le \frac{\kappa}{n}. \qquad (6.16)$$

For $\tilde{\mathcal{T}}_s^{n,3}$, we have

$$\begin{aligned}
\tilde{\mathcal{T}}_s^{n,3} &= \mathbb{E}\left|\int_I\int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)G(\frac{i}{n},\frac{\lceil nv\rceil}{n})\,\mu_{\frac{\lceil nv\rceil}{n},s}(dx)\,dv - \int_I\int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)G(\frac{i}{n},v)\,\mu_{v,s}(dx)\,dv\right|^2 \\
&\le 2\mathbb{E}\left|\int_I\int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)\left[G(\frac{i}{n},\frac{\lceil nv\rceil}{n}) - G(\frac{i}{n},v)\right]\mu_{\frac{\lceil nv\rceil}{n},s}(dx)\,dv\right|^2 \\
&\quad + 2\mathbb{E}\left|\int_I\left[\int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)\,\mu_{\frac{\lceil nv\rceil}{n},s}(dx)\,dv - \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)\,\mu_{v,s}(dx)\right]G(\frac{i}{n},v)\,dv\right|^2 \\
&\le \frac{\kappa}{n^2},
\end{aligned} \tag{6.17}$$

where the last inequality uses Condition 2.3, Theorem 2.1(b) and Remark 2.3. Combining (6.13)–(6.17) with Gronwall's inequality gives (3.5) and completes the proof of Theorem 3.2. ■

7. Proofs for Section 4

In this section we prove Theorems 4.1 and 4.2. The following moment estimate will be needed.

Lemma 7.1. *Suppose either Condition* 4.1 *or Condition* 4.2 *holds. Suppose* $\liminf_{n\to\infty} n\beta_n > 0$. *Then*

$$\sup_{n\in\mathbb{N}}\max_{i=1,\dots,n}\mathbb{E}\|X_i^n - X_{\frac{i}{n}}\|_{*,T}^k < \infty, \quad \forall k\in\mathbb{N}.$$

Proof of Lemma 7.1. Fix $k\in\mathbb{N}$ and $i\in\{1,\dots,n\}$. Since b and σ are bounded, we have

$$\begin{aligned}
\mathbb{E}\|X_i^n - X_{\frac{i}{n}}\|_{*,T}^k &\le \kappa\mathbb{E}\left|\frac{1}{n\beta_n}\sum_{j=1}^n \xi_{ij}^n\right|^k + \kappa \\
&\le \kappa\frac{\sum_{j=1}^n \mathbb{E}[(\xi_{ij}^n)^k] + \left(\sum_{j=1}^n E[(\xi_{ij}^n)^2]\right)^{k/2}}{(n\beta_n)^k} + \kappa \le \kappa,
\end{aligned}$$

where the second inequality follows from the Rosenthal's inequality [28, Theorem 3] and the last inequality uses the assumption $\liminf_{n\to\infty} n\beta_n > 0$. ■

7.1. Proofs of Theorems 4.1 and 4.2.

For the first assertion in Theorem 4.1, fix $t\in[0,T]$.

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n \mathbb{E}\|X_i^n - X_{\frac{i}{n}}\|_{*,t}^2 \\
&\le \kappa\int_0^t\left[\frac{1}{n}\sum_{i=1}^n\mathbb{E}\left|\frac{1}{n\beta_n}\sum_{j=1}^n \xi_{ij}^n b(X_i^n(s),X_j^n(s)) - \int_I\int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)G(\frac{i}{n},v)\,\mu_{v,s}(dx)\,dv\right|^2\right]ds \\
&\quad + \kappa\int_0^t\left[\frac{1}{n}\sum_{i=1}^n\mathbb{E}\left|X_i^n(s) - X_{\frac{i}{n}}(s)\right|^2\right]ds.
\end{aligned} \tag{7.1}$$

We will analyze the first integrand above for fixed $s \in [0,t]$. By adding and subtracting terms, we have

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left|\frac{1}{n\beta_n}\sum_{j=1}^n \xi_{ij}^n b(X_i^n(s),X_j^n(s)) - \int_I\int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)G(\frac{i}{n},v)\,\mu_{v,s}(dx)\,dv\right|^2 \\
&\le \frac{4}{n}\sum_{i=1}^n \mathbb{E}\left|\frac{1}{n}\sum_{j=1}^n \frac{\xi_{ij}^n}{\beta_n}\Big(b(X_i^n(s),X_j^n(s)) - b(X_{\frac{i}{n}}(s),X_j^n(s))\Big)\right|^2 \\
&\quad + \frac{4}{n}\sum_{i=1}^n \mathbb{E}\left|\frac{1}{n}\sum_{j=1}^n \frac{\xi_{ij}^n}{\beta_n}\Big(b(X_{\frac{i}{n}}(s),X_j^n(s)) - b(X_{\frac{i}{n}}(s),X_{\frac{j}{n}}(s))\Big)\right|^2 \\
&\quad + \frac{4}{n}\sum_{i=1}^n \mathbb{E}\left|\frac{1}{n}\sum_{j=1}^n \left(\frac{\xi_{ij}^n}{\beta_n} b(X_{\frac{i}{n}}(s),X_{\frac{j}{n}}(s)) - \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)G_n(\frac{i}{n},\frac{j}{n})\,\mu_{\frac{j}{n},s}(dx)\right)\right|^2 \\
&\quad + \frac{4}{n}\sum_{i=1}^n \mathbb{E}\left|\frac{1}{n}\sum_{j=1}^n \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)G_n(\frac{i}{n},\frac{j}{n})\,\mu_{\frac{j}{n},s}(dx) - \int_I\int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)G(\frac{i}{n},v)\,\mu_{v,s}(dx)\,dv\right|^2 \\
&=: 4\left(\mathcal{R}_s^{n,1} + \mathcal{R}_s^{n,2} + \mathcal{R}_s^{n,3} + \mathcal{R}_s^{n,4}\right). \qquad (7.2)
\end{aligned}
$$

For each term $\mathcal{R}_s^{n,k}$, $k = 1,2,3$, we have the following key estimates, whose proof will be given in Section 7.2.

Lemma 7.2. *Suppose either Condition* 4.1 *or Condition* 4.2 *holds. Suppose Condition* 2.2 *holds and* $\liminf_{n\to\infty} n\beta_n > 0$. *Then there exists some* $\kappa, \kappa(q) \in (0,\infty)$ *for each* $q \in (1,\infty)$ *such that the following holds for each* $s \in [0,T]$*:*

$$\mathcal{R}_s^{n,1} \le \frac{\kappa}{n\beta_n} + \frac{\kappa}{n}\sum_{i=1}^n \mathbb{E}|X_i^n(s) - X_{\frac{i}{n}}(s)|^2, \qquad (7.3)$$

$$\mathcal{R}_s^{n,2} \le \frac{\kappa}{n}\sum_{j=1}^n \mathbb{E}|X_j^n(s) - X_{\frac{j}{n}}(s)|^2 + \frac{\kappa(q)}{(n\beta_n)^{1/q}}, \qquad (7.4)$$

$$\mathcal{R}_s^{n,3} \le \frac{\kappa}{n\beta_n}. \qquad (7.5)$$

Completing the proof of Theorem 4.1: For $\mathcal{R}_s^{n,4}$, similar to the proof of (6.5) in Lemma 6.1, using Conditions 2.1, 2.2 and 4.1 we could get

$$\lim_{n\to\infty}\int_0^T \mathcal{R}_s^{n,4}\,ds = 0.$$

Combining this, (7.1)–(7.5) with Gronwall's inequality gives (4.2). The proof of the second assertion in Theorem 4.1 is the same as that of the second assertion in Theorem 3.1, and hence is omitted.

Completing the proof of Theorem 4.2: In view of (7.1)–(7.5), in order to show (4.3), it suffices to argue

$$\mathcal{R}_s^{n,4} \le \frac{\kappa}{n}.$$

For this, using Condition 4.2 we have

$$
\begin{aligned}
\mathcal{R}^{n,4}_s &= \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left|\frac{1}{n}\sum_{j=1}^n \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)G(\frac{i}{n},\frac{j}{n})\,\mu_{\frac{j}{n},s}(dx)\right. \\
&\qquad \left. - \int_I\int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)G(\frac{i}{n},v)\,\mu_{v,s}(dx)\,dv\right|^2 \\
&= \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left|\int_I\int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)G(\frac{i}{n},\frac{\lceil nv\rceil}{n})\,\mu_{\frac{\lceil nv\rceil}{n},s}(dx)\,dv\right. \\
&\qquad \left. - \int_I\int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)G(\frac{i}{n},v)\,\mu_{v,s}(dx)\,dv\right|^2 \\
&\le 2\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left|\int_I\int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)\left[G(\frac{i}{n},\frac{\lceil nv\rceil}{n}) - G(\frac{i}{n},v)\right]\mu_{\frac{\lceil nv\rceil}{n},s}(dx)\,dv\right|^2 \\
&\quad + 2\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left|\int_I\left[\int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)\,\mu_{\frac{\lceil nv\rceil}{n},s}(dx)\,dv - \int_{\mathbb{R}^d} b(X_{\frac{i}{n}}(s),x)\,\mu_{v,s}(dx)\right]G(\frac{i}{n},v)\,dv\right|^2 \\
&\le \frac{\kappa}{n^2},
\end{aligned}
$$

where the last inequality uses Condition 2.3, Theorem 2.1(b) and Remark 2.3. This completes the proof of Theorem 4.2. ■

7.2. **Proof of Lemma 7.2.** Fix $s \in [0,T]$. For $\mathcal{R}^{n,1}_s$, using the Lipschitz property of b we have

$$
\begin{aligned}
\mathcal{R}^{n,1}_s &\le \kappa\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left(\frac{1}{n}\sum_{j=1}^n \frac{\xi^n_{ij}}{\beta_n}|X^n_i(s) - X_{\frac{i}{n}}(s)|\right)^2 \\
&\le \kappa\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left(\frac{1}{n}\sum_{j=1}^n \frac{\xi^n_{ij} - \beta_n G_n(\frac{i}{n},\frac{j}{n})}{\beta_n}|X^n_i(s) - X_{\frac{i}{n}}(s)|\right)^2 \\
&\quad + \kappa\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left(\frac{1}{n}\sum_{j=1}^n G_n(\frac{i}{n},\frac{j}{n})|X^n_i(s) - X_{\frac{i}{n}}(s)|\right)^2.
\end{aligned}
$$

Since $\{\xi_{ij}^n - \beta_n G_n(\frac{i}{n}, \frac{j}{n})\}$ are centered and independent, we have

$$
\begin{aligned}
&\mathbb{E}\left(\frac{1}{n}\sum_{j=1}^n \frac{\xi_{ij}^n - \beta_n G_n(\frac{i}{n}, \frac{j}{n})}{\beta_n}\right)^4 \\
&= \frac{1}{(n\beta_n)^4}\sum_{j=1}^n \mathbb{E}\left(\xi_{ij}^n - \beta_n G_n(\frac{i}{n}, \frac{j}{n})\right)^4 \\
&\quad + \frac{3}{(n\beta_n)^4}\sum_{j=1}^n\sum_{k\neq j}^n \mathbb{E}\left(\xi_{ij}^n - \beta_n G_n(\frac{i}{n}, \frac{j}{n})\right)^2 \mathbb{E}\left(\xi_{ik}^n - \beta_n G_n(\frac{i}{n}, \frac{k}{n})\right)^2 \\
&\le \frac{1}{(n\beta_n)^3} + \frac{3}{(n\beta_n)^2}.
\end{aligned}
$$

From these two estimates, Cauchy–Schwarz inequality and Lemma 7.1 we have (7.3).

For $\mathcal{R}_s^{n,3}$, using Condition 4.1, the independence of $\{X_{\frac{i}{n}}\}$ and $\{\xi_{ij}^n\}$, the boundedness of b, and a weak LLN type argument, we have (7.5).

The analysis of $\mathcal{R}_s^{n,2}$ is based on a collection of change of measure arguments. First note that

$$
\begin{aligned}
\mathcal{R}_s^{n,2} &\le \kappa \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left(\frac{1}{n}\sum_{j=1}^n \frac{\xi_{ij}^n}{\beta_n}|X_j^n(s) - X_{\frac{j}{n}}(s)|\right)^2 \\
&= \kappa \frac{1}{n}\sum_{i=1}^n \frac{1}{(n\beta_n)^2}\left(\sum_{j=1}^n \mathbb{E}\left[\xi_{ij}^n |X_j^n(s) - X_{\frac{j}{n}}(s)|^2\right]\right. \\
&\qquad \left. + \sum_{j=1}^n\sum_{k\neq j}^n \mathbb{E}\left[\xi_{ij}^n\xi_{ik}^n |X_j^n(s) - X_{\frac{j}{n}}(s)||X_k^n(s) - X_{\frac{k}{n}}(s)|\right]\right). \qquad (7.6)
\end{aligned}
$$

Fix $i, j \in \{1, \dots, n\}$. Consider auxiliary processes given by

$$
\begin{aligned}
\tilde{X}_{\frac{i}{n}}(t) &= X_{\frac{i}{n}}(0) + \int_0^t \int_I \int_{\mathbb{R}^d} b(\tilde{X}_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v)\, \mu_{v,s}(dx)\, dv\, ds + \int_0^t \sigma(\tilde{X}_{\frac{i}{n}}(s))\, dB_{\frac{i}{n}}(s) \\
&\quad + \int_0^t \frac{\sigma(\tilde{X}_{\frac{i}{n}}(s))}{\sigma(X_i^n(s))} \frac{1}{n\beta_n}\left(\xi_{ij}^n - 1\right) b(X_i^n(s), X_j^n(s))\, ds, \\
\tilde{X}_{\frac{j}{n}}(t) &= X_{\frac{j}{n}}(0) + \int_0^t \int_I \int_{\mathbb{R}^d} b(\tilde{X}_{\frac{j}{n}}(s), x) G(\frac{j}{n}, v)\, \mu_{v,s}(dx)\, dv\, ds + \int_0^t \sigma(\tilde{X}_{\frac{j}{n}}(s))\, dB_{\frac{j}{n}}(s) \\
&\quad + \int_0^t \frac{\sigma(\tilde{X}_{\frac{j}{n}}(s))}{\sigma(X_j^n(s))} \frac{1}{n\beta_n}\left(\xi_{ji}^n - 1\right) b(X_j^n(s), X_i^n(s))\, ds.
\end{aligned}
$$

Note that the existence and uniqueness of such processes are guaranteed by the bounded and Lipschitz properties of b, σ and σ^{-1}. Also using these properties and Gronwall's inequality we can show that

$$
\mathbb{E}\|X_{\frac{j}{n}} - \tilde{X}_{\frac{j}{n}}\|_{*,T}^m \le \frac{\kappa(m)}{(n\beta_n)^m}, \quad m \ge 0. \qquad (7.7)
$$

Define $Q^{i,j,n}$ by

$$\frac{dQ^{i,j,n}}{d\mathbb{P}} = \mathcal{E}_T\left(\int_0^{\cdot} \frac{1}{\sigma(X_i^n(s))}\frac{1}{n\beta_n}\left(1-\xi_{ij}^n\right) b(X_i^n(s), X_j^n(s))\, dB_{\frac{i}{n}} + \int_0^{\cdot} \frac{1}{\sigma(X_j^n(s))}\frac{1}{n\beta_n}\left(1-\xi_{ji}^n\right) b(X_j^n(s), X_i^n(s))\, dB_{\frac{j}{n}}\right),$$

where

$$\mathcal{E}_t(M) := \exp\left\{M_t - \frac{1}{2}[M]_t\right\}$$

is the Doleans exponential for a semi-martingale M_t. Since b and σ^{-1} are bounded, we have

$$\mathbb{P}\left(\left(X_i^n, X_j^n, X_{\frac{i}{n}}, X_{\frac{j}{n}}\right) \in \cdot \,|\, \xi_{ij}^n = 1\right) = Q^{i,j,n}\left(\left(X_i^n, X_j^n, \tilde{X}_{\frac{i}{n}}, \tilde{X}_{\frac{j}{n}}\right) \in \cdot\right) \tag{7.8}$$

by the Girsanov's theorem, and

$$\mathbb{E}\left[\left(\frac{dQ^{i,j,n}}{d\mathbb{P}}\right)^m\right] \le \exp\left\{\frac{m|m-1|\kappa}{(n\beta_n)^2}\right\}, \quad m \ge 0.$$

From this it then follows that

$$\begin{aligned}
\mathbb{E}\left[\left|\frac{dQ^{i,j,n}}{d\mathbb{P}} - 1\right|^m\right] &\le \sqrt{\mathbb{E}\left[\left|\frac{dQ^{i,j,n}}{d\mathbb{P}} - 1\right|^2\right]\mathbb{E}\left[\left|\frac{dQ^{i,j,n}}{d\mathbb{P}} - 1\right|^{2m-2}\right]} \\
&\le \sqrt{\mathbb{E}\left[\left(\frac{dQ^{i,j,n}}{d\mathbb{P}}\right)^2 - 1\right]\kappa(m)\mathbb{E}\left[\left|\frac{dQ^{i,j,n}}{d\mathbb{P}}\right|^{2m-2} + 1\right]} \\
&\le \sqrt{\left[1 + \frac{\kappa(m)}{(n\beta_n)^2} - 1\right]\kappa(m)} \le \frac{\kappa(m)}{n\beta_n}, \quad m \ge 1,
\end{aligned} \tag{7.9}$$

where the third inequality uses the assumption that $\liminf_{n\to\infty} n\beta_n > 0$. From (7.8) we have

$$\begin{aligned}
\mathbb{E}\left[\xi_{ij}^n (X_j^n(s) - X_{\frac{j}{n}}(s))^2\right] &= \mathbb{E}\left[(X_j^n(s) - X_{\frac{j}{n}}(s))^2 \,|\, \xi_{ij}^n = 1\right]\beta_n G_n(\frac{i}{n}, \frac{j}{n}) \\
&= \mathbb{E}_{Q^{i,j,n}}\left[(X_j^n(s) - \tilde{X}_{\frac{j}{n}}(s))^2\right]\beta_n G_n(\frac{i}{n}, \frac{j}{n}).
\end{aligned}$$

Note that

$$\begin{aligned}
\mathbb{E}_{Q^{i,j,n}}\left[(X_j^n(s) - \tilde{X}_{\frac{j}{n}}(s))^2\right] &= \mathbb{E}\left[(X_j^n(s) - \tilde{X}_{\frac{j}{n}}(s))^2 \frac{dQ^{i,j,n}}{d\mathbb{P}}\right] \\
&\le 2\mathbb{E}\left[(X_j^n(s) - X_{\frac{j}{n}}(s))^2 \frac{dQ^{i,j,n}}{d\mathbb{P}}\right] + 2\mathbb{E}\left[(X_{\frac{j}{n}}(s) - \tilde{X}_{\frac{j}{n}}(s))^2 \frac{dQ^{i,j,n}}{d\mathbb{P}}\right].
\end{aligned}$$

For the first term, using Holder's inequality, Lemma 7.1 and (7.9) we have

$$\begin{aligned}
&\mathbb{E}\left[(X_j^n(s) - X_{\frac{j}{n}}(s))^2 \frac{dQ^{i,j,n}}{d\mathbb{P}}\right] - \mathbb{E}\left[(X_j^n(s) - X_{\frac{j}{n}}(s))^2\right] \\
&\le \left(\mathbb{E}\left[(X_j^n(s) - X_{\frac{j}{n}}(s))^{2p}\right]\right)^{1/p}\left(\mathbb{E}\left[\left|\frac{dQ^{i,j,n}}{d\mathbb{P}} - 1\right|^q\right]\right)^{1/q} \\
&\le \frac{\kappa(q)}{(n\beta_n)^{1/q}}, \quad \forall q > 1.
\end{aligned}$$

For the second term, using Holder's inequality and (7.7) we have

$$\mathbb{E}\left[\left(X_{\frac{j}{n}}(s)-\tilde{X}_{\frac{j}{n}}(s)\right)^2\frac{dQ^{i,j,n}}{d\mathbb{P}}\right]\le\sqrt{\mathbb{E}\left[\left(X_{\frac{j}{n}}(s)-\tilde{X}_{\frac{j}{n}}(s)\right)^4\right]\mathbb{E}\left[\left(\frac{dQ^{i,j,n}}{d\mathbb{P}}\right)^2\right]}\le\frac{\kappa}{(n\beta_n)^2}.$$

Combining these three estimates gives

$$\mathbb{E}\left[\xi^n_{ij}(X^n_j(s)-X_{\frac{j}{n}}(s))^2\right]\le\left(2\mathbb{E}\left[(X^n_j(s)-X_{\frac{j}{n}}(s))^2\right]+\frac{\kappa(q)}{(n\beta_n)^{1/q}}\right)\beta_nG_n(\frac{i}{n},\frac{j}{n}).\tag{7.10}$$

Now fix $i,j,k\in\{1,\dots,n\}$ with $j\neq k$. The following argument is similar to the above change of measure, but we provide the proof for completeness. Consider auxiliary processes given by

$$\begin{aligned}\tilde{X}_{\frac{i}{n}}(t)&=X_{\frac{i}{n}}(0)+\int_0^t\int_I\int_{\mathbb{R}^d}b(\tilde{X}_{\frac{i}{n}}(s),x)G(\frac{i}{n},v)\,\mu_{v,s}(dx)\,dv\,ds+\int_0^t\sigma(\tilde{X}_{\frac{i}{n}}(s))\,dB_{\frac{i}{n}}(s)\\&\quad+\sum_{l=j,k}\int_0^t\frac{\sigma(\tilde{X}_{\frac{i}{n}}(s))}{\sigma(X^n_i(s))}\frac{1}{n\beta_n}\,(\xi^n_{il}-1)\,b(X^n_i(s),X^n_l(s))\,ds,\\\tilde{X}_{\frac{j}{n}}(t)&=X_{\frac{j}{n}}(0)+\int_0^t\int_I\int_{\mathbb{R}^d}b(\tilde{X}_{\frac{j}{n}}(s),x)G(\frac{j}{n},v)\,\mu_{v,s}(dx)\,dv\,ds+\int_0^t\sigma(\tilde{X}_{\frac{j}{n}}(s))\,dB_{\frac{j}{n}}(s)\\&\quad+\int_0^t\frac{\sigma(\tilde{X}_{\frac{j}{n}}(s))}{\sigma(X^n_j(s))}\frac{1}{n\beta_n}\left(\xi^n_{ji}-1\right)b(X^n_j(s),X^n_i(s))\,ds\\\tilde{X}_{\frac{k}{n}}(t)&=X_{\frac{k}{n}}(0)+\int_0^t\int_I\int_{\mathbb{R}^d}b(\tilde{X}_{\frac{k}{n}}(s),x)G(\frac{k}{n},v)\,\mu_{v,s}(dx)\,dv\,ds+\int_0^t\sigma(\tilde{X}_{\frac{k}{n}}(s))\,dB_{\frac{k}{n}}(s)\\&\quad+\int_0^t\frac{\sigma(\tilde{X}_{\frac{k}{n}}(s))}{\sigma(X^n_k(s))}\frac{1}{n\beta_n}\,(\xi^n_{ki}-1)\,b(X^n_k(s),X^n_i(s))\,ds.\end{aligned}$$

Note that the existence and uniqueness of such processes are again guaranteed by the bounded and Lipschitz properties of b, σ and σ^{-1}. Also using these properties and Gronwall's inequality we can show that

$$\mathbb{E}\|X_{\frac{j}{n}}-\tilde{X}_{\frac{j}{n}}\|^m_{*,T}+\mathbb{E}\|X_{\frac{k}{n}}-\tilde{X}_{\frac{k}{n}}\|^m_{*,T}\le\frac{\kappa(m)}{(n\beta_n)^m},\quad m\ge0.\tag{7.11}$$

Define $Q^{i,j,k,n}$ by

$$\begin{aligned}\frac{dQ^{i,j,k,n}}{d\mathbb{P}}=\mathcal{E}_T\Bigg(&\sum_{l=j,k}\int_0^\cdot\frac{1}{\sigma(X^n_i(s))}\frac{1}{n\beta_n}\,(1-\xi^n_{il})\,b(X^n_i(s),X^n_l(s))\,dB_{\frac{i}{n}}\\&+\sum_{l=j,k}\int_0^\cdot\frac{1}{\sigma(X^n_l(s))}\frac{1}{n\beta_n}\,(1-\xi^n_{li})\,b(X^n_l(s),X^n_i(s))\,dB_{\frac{l}{n}}\Bigg).\end{aligned}$$

Since b and σ^{-1} are bounded, we have

$$\begin{aligned}&\mathbb{P}\left(\left(X^n_i,X^n_j,X^n_k,X_{\frac{i}{n}},X_{\frac{j}{n}},X_{\frac{k}{n}}\right)\in\cdot\,|\,\xi^n_{ij}=1,\xi^n_{ik}=1\right)\\&=Q^{i,j,k,n}\left(\left(X^n_i,X^n_j,X^n_k,\tilde{X}_{\frac{i}{n}},\tilde{X}_{\frac{j}{n}},\tilde{X}_{\frac{k}{n}}\right)\in\cdot\right)\end{aligned}\tag{7.12}$$

by the Girsanov's theorem, and

$$\mathbb{E}\left[\left(\frac{dQ^{i,j,k,n}}{d\mathbb{P}}\right)^m\right] \leq \exp\left\{\frac{m|m-1|\kappa}{(n\beta_n)^2}\right\}, m \geq 0.$$

Using this and the assumption that $\liminf_{n\to\infty} n\beta_n > 0$, we have

$$\mathbb{E}\left[\left|\frac{dQ^{i,j,n}}{d\mathbb{P}}-1\right|^m\right] \leq \sqrt{\mathbb{E}\left[\left|\frac{dQ^{i,j,n}}{d\mathbb{P}}-1\right|^2\right]\mathbb{E}\left[\left|\frac{dQ^{i,j,n}}{d\mathbb{P}}-1\right|^{2m-2}\right]} \leq \frac{\kappa(m)}{n\beta_n}, m \geq 1. \tag{7.13}$$

From (7.12) we have

$$\begin{aligned}
&\mathbb{E}\left[\xi^n_{ij}\xi^n_{ik}|X^n_j(s)-X_{\frac{j}{n}}(s)||X^n_k(s)-X_{\frac{k}{n}}(s)|\right] \\
&= \mathbb{E}\left[|X^n_j(s)-X_{\frac{j}{n}}(s)||X^n_k(s)-X_{\frac{k}{n}}(s)|\,|\,\xi^n_{ij}=1,\xi^n_{ik}=1\right]\beta_n^2 G_n(\frac{i}{n},\frac{j}{n})G_n(\frac{i}{n},\frac{k}{n}) \\
&= \mathbb{E}_{Q^{i,j,k,n}}\left[|X^n_j(s)-\tilde{X}_{\frac{j}{n}}(s)||X^n_k(s)-\tilde{X}_{\frac{k}{n}}(s)|\right]\beta_n^2 G_n(\frac{i}{n},\frac{j}{n})G_n(\frac{i}{n},\frac{k}{n}).
\end{aligned}$$

Note that

$$\begin{aligned}
&\mathbb{E}_{Q^{i,j,k,n}}\left[|X^n_j(s)-\tilde{X}_{\frac{j}{n}}(s)||X^n_k(s)-\tilde{X}_{\frac{k}{n}}(s)|\right] \\
&\leq \frac{1}{2}\mathbb{E}_{Q^{i,j,k,n}}\left[|X^n_j(s)-\tilde{X}_{\frac{j}{n}}(s)|^2\right]+\frac{1}{2}\mathbb{E}_{Q^{i,j,k,n}}\left[|X^n_k(s)-\tilde{X}_{\frac{k}{n}}(s)|^2\right] \\
&= \frac{1}{2}\mathbb{E}\left[(X^n_j(s)-\tilde{X}_{\frac{j}{n}}(s))^2\frac{dQ^{i,j,k,n}}{d\mathbb{P}}\right]+\frac{1}{2}\mathbb{E}\left[(X^n_k(s)-\tilde{X}_{\frac{k}{n}}(s))^2\frac{dQ^{i,j,k,n}}{d\mathbb{P}}\right] \\
&\leq \mathbb{E}\left[(X^n_j(s)-X_{\frac{j}{n}}(s))^2\frac{dQ^{i,j,k,n}}{d\mathbb{P}}\right]+\mathbb{E}\left[(X_{\frac{j}{n}}(s)-\tilde{X}_{\frac{j}{n}}(s))^2\frac{dQ^{i,j,k,n}}{d\mathbb{P}}\right] \\
&\quad+\mathbb{E}\left[(X^n_k(s)-X_{\frac{k}{n}}(s))^2\frac{dQ^{i,j,k,n}}{d\mathbb{P}}\right]+\mathbb{E}\left[(X_{\frac{k}{n}}(s)-\tilde{X}_{\frac{k}{n}}(s))^2\frac{dQ^{i,j,k,n}}{d\mathbb{P}}\right].
\end{aligned}$$

Similar to the derivation of (7.10), using Holder's inequality, Lemma 7.1, (7.13), and (7.11) we have

$$\begin{aligned}
&\mathbb{E}\left[\xi^n_{ij}\xi^n_{ik}|X^n_j(s)-X_{\frac{j}{n}}(s)||X^n_k(s)-X_{\frac{k}{n}}(s)|\right] \qquad (7.14)\\
&\leq \left(\mathbb{E}\left[(X^n_j(s)-X_{\frac{j}{n}}(s))^2\right]+\mathbb{E}\left[(X^n_k(s)-X_{\frac{k}{n}}(s))^2\right]+\frac{\kappa_q}{(n\beta_n)^{1/q}}\right)\beta_n^2 G_n(\frac{i}{n},\frac{j}{n})G_n(\frac{i}{n},\frac{k}{n}).
\end{aligned}$$

Applying (7.10) and (7.14) to (7.6) gives

$$\mathcal{R}^{n,2}_s \leq \frac{\kappa}{n}\sum_{j=1}^n \mathbb{E}|X^n_j(s)-X_{\frac{j}{n}}(s)|^2+\frac{\kappa(q)}{(n\beta_n)^{1/q}}.$$

This completes the proof of Lemma 7.2. ■

References

[1] J. Baladron, D. Fasoli, O. Faugeras, and J. Touboul, *Mean-field description and propagation of chaos in networks of Hodgkin–Huxley and FitzHugh–Nagumo neurons*, The Journal of Mathematical Neuroscience **2** (2012), no. 1, 10.

[2] A. Basak, S. Bhamidi, S. Chakraborty, and A. Nobel, *Large subgraphs in pseudo-random graphs*, arXiv preprint arXiv:1610.03762 (2016).

[3] E. Bayraktar and R. Wu, *Mean field interaction on random graphs with dynamically changing multi-color edges*, arXiv preprint arXiv:1912.01785 (2019).

[4] M. J Beeson, *Foundations of constructive mathematics: Metamathematical studies*, Vol. 6, Springer Science & Business Media, 2012.

[5] S. Bhamidi, A. Budhiraja, and R. Wu, *Weakly interacting particle systems on inhomogeneous random graphs*, Stochastic Processes and their Applications **129** (2019), no. 6, 2174–2206.

[6] S. Bhamidi, S. Chakraborty, S. Cranmer, and B. Desmarais, *Weighted exponential random graph models: Scope and large network limits*, Journal of Statistical Physics **173** (2018), no. 3-4, 704–735.

[7] B. Bollobás, C. Borgs, J. Chayes, O. Riordan, et al., *Percolation on dense graph sequences*, The Annals of Probability **38** (2010), no. 1, 150–183.

[8] A. Budhiraja, P. Dupuis, M. Fischer, and K. Ramanan, *Limits of relative entropies associated with weakly interacting particle systems*, Electronic Journal of Probability **20** (2015), no. 80, 1–22.

[9] A. Budhiraja and R. Wu, *Some fluctuation results for weakly interacting multi-type particle systems*, Stochastic Processes and their Applications **126** (2016), no. 8, 2253–2296.

[10] A. Budhiraja, D. Mukherjee, and R. Wu, *Supermarket model on graphs*, Ann. Appl. Probab. **29** (201906), no. 3, 1740–1777.

[11] P. E Caines and M. Huang, *Graphon mean field games and the GMFG equations*, 2018 IEEE Conference on Decision and Control (CDC), 2018, pp. 4129–4134.

[12] R. Carmona, D. Cooney, C. Graves, and M. Lauriere, *Stochastic graphon games: I. the static case*, arXiv preprint arXiv:1911.10664 (2019).

[13] S. Chatterjee, P. Diaconis, et al., *Estimating and understanding exponential random graph models*, The Annals of Statistics **41** (2013), no. 5, 2428–2461.

[14] F. Collet, *Macroscopic limit of a bipartite Curie–Weiss model: A dynamical approach*, Journal of Statistical Physics **157** (2014), no. 6, 1301–1319.

[15] P. Contucci, I. Gallo, and G. Menconi, *Phase transitions in social sciences: Two-population mean field theory*, International Journal of Modern Physics B **22** (2008), no. 14, 2199–2212.

[16] F. Coppini, H. Dietert, and G. Giacomin, *A law of large numbers and large deviations for interacting diffusions on Erdős–Rényi graphs*, Stochastics and Dynamics **0** (2019), no. 0, 2050010, available at `https://doi.org/10.1142/S0219493720500100`.

[17] F. Delarue, *Mean field games: A toy model on an erdös-renyi graph.*, ESAIM: Proceedings and Surveys **60** (2017), 1–26.

[18] S. Delattre, G. Giacomin, and E. Luçon, *A note on dynamical models on random graphs and Fokker–Planck equations*, Journal of Statistical Physics **165** (2016), no. 4, 785–798.

[19] R. M Dudley, *Real analysis and probability*, CRC Press, 2018.

[20] J. Hadamard, *Lectures on Cauchy's problem in linear partial differential equations*, Courier Corporation, 2003.

[21] V. N. Kolokoltsov, *Nonlinear Markov Processes and Kinetic Equations*, Cambridge Tracts in Mathematics, vol. 182, Cambridge University Press, 2010.

[22] M. Krivelevich and B. Sudakov, *Pseudo-random graphs*, More sets, graphs and numbers, 2006, pp. 199–262.

[23] L. Lovász, *Large networks and graph limits*, Vol. 60, American Mathematical Soc., 2012.

[24] H. P. McKean, *Propagation of chaos for a class of non-linear parabolic equations*, Stochastic differential equations (Lecture Series in Differential Equations, Session 7, Catholic University, 1967), 1967, pp. 41–57.

[25] S. Nadtochiy and M. Shkolnikov, *Mean field systems on networks, with singular interaction through hitting times*, arXiv preprint arXiv:1807.02015 (2018).

[26] R. I. Oliveira and G. H. Reis, *Interacting diffusions on random graphs with diverging average degrees: Hydrodynamics and large deviations*, Journal of Statistical Physics (2019Jul).

[27] F. Parise and A. E Ozdaglar, *Graphon games: A statistical framework for network games and interventions*, Available at SSRN: https://ssrn.com/abstract=3437293 (2019).

[28] H. P Rosenthal, *On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables*, Israel Journal of Mathematics **8** (1970), no. 3, 273–303.

[29] M. H. Schultz, *L^∞-multivariate approximation theory*, SIAM Journal on Numerical Analysis **6** (1969), no. 2, 161–183.

[30] A-S. Sznitman, *Topics in propagation of chaos*, Ecole d'Eté de Probabilités de Saint-Flour XIX—1989, 1991, pp. 165–251.

Department of Mathematics, University of Michigan, 530 Church Street, Ann Arbor, MI 48109

Department of Mathematics and Computer Science, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands

Department of Mathematics, Iowa State University, 411 Morrill Road, Ames, IA 50011
Email address: erhan@umich.edu, s.chakraborty1@tue.nl, ruoyu@iastate.edu