# STATIONARITY AND UNIFORM IN TIME CONVERGENCE FOR THE GRAPHON PARTICLE SYSTEM

ERHAN BAYRAKTAR AND RUOYU WU

ABSTRACT. We consider the long time behavior of heterogeneously interacting diffusive particle systems and their large population limit. The interaction is of mean field type with weights characterized by an underlying graphon. The limit is given by a graphon particle system consisting of independent but heterogeneous nonlinear diffusions whose probability distributions are fully coupled. Under suitable assumptions, including a certain convexity condition, we show the exponential ergodicity for both systems, establish the uniform-in-time law of large numbers for marginal distributions as the number of particles increases, and introduce the uniform-in-time Euler approximation. The precise rate of convergence of the Euler approximation is provided.

## CONTENTS

## 1. INTRODUCTION

In this work we study the long time behavior of graphon particle systems and the finite particle approximations. The interaction is of mean-field type and characterized by a graphon

$G$, which is a symmetric measurable function from $[0,1] \times [0,1]$ to $[0,1]$ (see e.g. [27] for the theory of graphons). More precisely, denoting by $\bar{X}_u$ the state of the particle at $u \in [0,1]$,

$$\bar{X}_u(t) = \bar{X}_u(0) + \int_0^t \left( f(\bar{X}_u(s)) + \int_0^1 \int_{\mathbb{R}^d} b(\bar{X}_u(s), x) G(u,v) \, \bar{\mu}_{v,s}(dx) \, dv \right) ds$$
$$+ \sigma B_u(t), \quad u \in [0,1], \, t \geq 0, \tag{1.1}$$

where $\bar{\mu}_{v,s}$ is the probability distribution of the $\mathbb{R}^d$-valued random variable $\bar{X}_v(s)$ for each $v \in [0,1]$ and $s \geq 0$, $f$ are $b$ are suitable functions, $\sigma \in \mathbb{R}^{d \times d}$ is a constant, $\{B_u : u \in I\}$ are $d$-dimensional standard Brownian motions, and $\{\bar{X}_u(0), B_u : u \in I\}$ are mutually independent. We will also study the mean-field particle system with heterogeneous interactions given by

$$X_i^n(t) = \bar{X}_{\frac{i}{n}}(0) + \int_0^t \left( f(X_i^n(s)) + \frac{1}{n} \sum_{j=1}^n \xi_{ij}^n b(X_i^n(s), X_j^n(s)) \right) ds$$
$$+ \sigma B_{\frac{i}{n}}(t), \quad i \in \{1, \ldots, n\}, \, t \geq 0, \tag{1.2}$$

and its Euler discretization. Here $\{\xi_{ij}^n : 1 \leq i \leq j \leq n\}$ is a collection of independent $[0,1]$-valued random variables sampled from a step graphon $G_n$ that converges to the graphon $G$ in the cut metric.

   The study of mean-field *heterogeneously* interacting particle systems on random graphs converging to a graphon emerged recently ([3, 5, 15, 19, 23, 28, 30–32]). There is also a growing number of applications of graphons in game theory; see e.g. [1, 11–13, 21, 33, 37]for the study of graphon mean field games in static and dynamic settings. Among these, the only work on long time analysis is [15], which shows that the stochastic Kuramoto model defined on a sequence of graphs converging to a constant graphon behaves asymptotically as the mean-field limit (in general the manifold of McKean–Vlasov equations), up to an exponential time. More precisely, it is shown in [15] that, in the subcritical regime, the trajectory of the empirical measure of $n$ oscillators ($[0, 2\pi]$-valued diffusions), over $t \in [0, T_n]$ where $T_n = \exp(o(n))$, converges uniformly in probability to the same limit as in the Kuramoto model; while in the supercritical regime, with initial states close to a stable stationary solution, the trajectory of the empirical measure is uniformly close to the manifold of stable stationary solutions over $t \in [0, T_n]$. This is in contrast to our case where the limiting system (1.1) is heterogeneous and the stationary measure relies crucially on the underlying graphon (see Example 3.1).

   The study of classic mean-field *homogeneously* interacting particle systems and the associated limiting system given by nonlinear processes, or equivalently, McKean–Vlasov equations, dates back to works of Boltzmann, Vlasov, McKean and others; see [25, 29, 36] and references therein. Besides large population limits such as law of large numbers (LLN) and propagation of chaos (POC) on the finite time horizon, there have been an extensive collection of results on long time behaviors and Euler approximations for such systems (see e.g. [7, 8, 10, 38] and references therein) with suitable convexity assumptions. In recent ten years, there has been a growing interest in the mean-field *inhomogeneous* particle system, where the interaction between particles is governed by their own types and/or random graphs (see e.g. [2, 4, 6, 9, 16–18, 26]) and the limiting system consists of countable McKean-Vlasov processes (as opposed to uncountable heterogeneous processes in graphon particle systems like (1.1)). Among these, the paper [2] considers a collection of diffusions interacting through state-dependent fast evolving random graphs and shows a uniform-in-time averaging and LLN result. The interaction in [2], although not in the mean-field form, is close to be mean-field due to the averaging effect,

and the limiting system is given by independent and identically distributed (i.i.d.) nonlinear diffusions.

The goal of this work is to study the long time behavior, including the stationary distribution and uniform-in-time convergence, of the graphon particle system (1.1), the approximating finite particle system (1.2), and its Euler discretization, under our standing assumptions stated in Section 2, including convexity conditions (2.1) and (2.2). We are in particular interested in the following two questions:

1. Knowing that $G_n \to G$ as $n \to \infty$, does the long time behavior of $\{X_i^n : i = 1, \ldots, n\}$ approach that of $\{\bar{X}_u : u \in [0,1]\}$?
2. Given a graphon $G$ and the associated $\{\bar{X}_u : u \in [0,1]\}$, could one choose a sequence of graphons $G_n$, finite particle systems $\{X_i^n : i = 1, \ldots, n\}$, and the Euler discretizations to approximate the long time behavior of $\{\bar{X}_u : u \in [0,1]\}$ such as the stationary distribution? In order to control the approximation error, what would be the balance between the number of particles, the time to run processes, and the discretization step size?

The first question is natural and its finite time analogue has been answered in the works [3, 5, 15, 28, 32] mentioned above with (possibly) different model setups. The second question, opposite to the first one, is also (and actually more) important to us, as one may worry that the graphon particle system (1.1), consisting of uncountably many heterogeneous particles (or equivalently, their probability distributions), is not always tractable, for the either finite-time or long-time behavior (see Example 3.1). Answering the second question will suggest, for example to simulate the stationary distributions of (1.1) using the Euler discretization of (1.2), and how to choose parameters in a balanced and efficient manner. As the classic POC result was useful in finding $\varepsilon$-Nash equilibrium in the mean field game (see e.g. [14]), our uniform-in-time convergence and POC results will be helpful for the same purpose in the graphon mean field game counterpart (such as with mean-reverting drifts).

Our first main result is the exponential ergodicity of the two systems (1.1) and (1.2). For the graphon particle system (1.1), we show that $\bar{\mu}_{u,t}$ converges to a limiting distribution as $t \to \infty$ for each $u \in [0,1]$ with exponentially small errors, and the limit is invariant with respect to the system evolution (Theorem 3.1). For the finite particle system (1.2), although the random vector $(X_i^n(t) : i = 1, \ldots, n)$ is not Markovian, it is Markovian conditioned on the interaction $\{\xi_{ij}^n : 1 \le i \le j \le n\}$. Using this observation we prove the quenched (and hence annealed) exponential ergodicity and also show that the quenched limiting distribution is invariant (Theorem 4.1).

The second main result is the uniform-in-time convergence of (1.2) to (1.1) when $G_n \to G$ in the cut metric. A uniform-in-time LLN for marginal distributions is established in Theorem 5.1, which says that the empirical measure $\nu^n(t)$ of $n$ particles $\{X_i^n : i = 1, \ldots, n\}$ converges to the averaged distribution of a continuum of particles $\{\bar{X}_u : u \in [0,1]\}$. The proof relies on a truncation and approximation argument for the drift coefficients (Lemma 7.2), and certain generalization (Lemma A.1) of bounds on the Wasserstein distance between i.i.d. random variables and their common distribution established in [20]. In Theorem 5.2, we strengthen Theorem 5.1 with additional assumptions on graphons, to further obtain a uniform-in-particle convergence, LLN, and POC, all uniformly in time. Theorem 5.1 or 5.2, together with the exponential ergodicity, guarantees the interchange of limits of large $n$ and $t$ for empirical measures and the convergence of stationary measures (Corollaries 5.1, 5.2 and 5.3).

Our last main result is on the tractable computation of the graphon particle system and its stationary distributions. We study the Euler scheme associated with (1.2), and obtain a uniform-in-time bound (Theorem 6.1) for errors arising from the discretization. Under certain

conditions on the graphons, explicit rates of convergence are obtained in Corollaries 6.1 and 6.2, of the empirical measure of the Euler discretization to the stationary distribution of the graphon particle system (1.1). This in particular answers the second question above.

1.1. **Organization.** The paper is organized as follows. In Section 2 we state the space of graphons, the standing assumptions, and well-posedness of systems (1.1) and (1.2). In Section 3 we study the long time behavior of the graphon particle system (1.1). The exponential ergodicity and stationary distribution are shown in Theorem 3.1. In Section 4 we analyze the long time behavior of the finite particle system (1.2). The quenched and annealed exponential ergodicity and quenched stationary distribution are shown in Theorem 4.1. In Section 5 we study the uniform-in-time convergence of the system (1.2) to the system (1.1). LLN and POC are given in Theorems 5.1 and 5.2. We also show the interchange of limits and the convergence of limiting distributions in Corollaries 5.1, 5.2 and 5.3. In Section 6 we introduce the Euler discretization. The convergence that is uniform in time and the number of particles is shown in Theorem 6.1. The rate of convergence in given in Corollaries 6.1 and 6.2. Finally Section 7 collects the proofs of results in Sections 3–6.

We close this section by introducing some frequently used notation.

1.2. **Notation.** Given a Polish space $\mathbb{S}$, denote by $\mathcal{B}(\mathbb{S})$ the Borel $\sigma$-field. Let $\mathcal{P}(\mathbb{S})$ be the space of probability measures on $\mathbb{S}$ endowed with the topology of weak convergence. For a measurable function $f \colon \mathbb{S} \to \mathbb{R}$, let $\|f\|_\infty := \sup_{x \in \mathbb{S}} |f(x)|$. Denote by $\mathbb{C}([0,\infty) : \mathbb{S})$ (resp. $\mathbb{C}([0,T] : \mathbb{S})$ for $T \in (0,\infty)$) the space of continuous functions from $[0,\infty)$ (resp. $[0,T]$) to $\mathbb{S}$, endowed with the topology of uniform convergence on compacts (resp. uniform convergence). We will use $C$ to denote various positive constants in the paper and $C_m$ to emphasize the dependence on some parameter $m$. Their values may change from line to line. The probability law of a random variable $X$ will be denoted by $\mathcal{L}(X)$. Expectations under $\mathbb{P}$ will be denoted by $\mathbb{E}$. To simplify the notation, we will usually write $\mathbb{E}[X^k]$ as $\mathbb{E}X^k$. For vectors $x, y \in \mathbb{R}^d$, denote by $|x|$ the Euclidean norm and $x \cdot y$ the inner product. Let $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

Denote by $W_p$, $p \in \mathbb{N}$, the Wasserstein-$p$ distance (cf. [39, Chapter 6]) on $\mathcal{P}(\mathbb{R}^k)$, $k \in \mathbb{N}$:

$$W_p(m_1, m_2) := \left( \inf_\pi \int_{\mathbb{R}^k \times \mathbb{R}^k} |x - y|^p \, \pi(dx \, dy) \right)^{1/p}, \quad m_1, m_2 \in \mathcal{P}(\mathbb{R}^k),$$

where the infimum is taken over all probability measures $\pi \in \mathcal{P}(\mathbb{R}^k \times \mathbb{R}^k)$ with marginals $m_1$ and $m_2$, that is, $\pi(\cdot \times \mathbb{R}^k) = m_1(\cdot)$ and $\pi(\mathbb{R}^k \times \cdot) = m_2(\cdot)$. It is well-known that (cf. [39, Remarks 6.5 and 6.6])

$$W_p(m_1, m_2) \geq W_1(m_1, m_2)$$
$$= \sup \left\{ \int_{\mathbb{R}^k} \phi(x) \, m_1(dx) - \int_{\mathbb{R}^k} \phi(x) \, m_2(dx) \,\Big|\, \phi \colon \mathbb{R}^k \to \mathbb{R} \text{ is 1-Lipschitz} \right\}. \quad (1.3)$$

Note that

$$W_2^2(m, \tilde{m}) \geq W_2^2(m_1, \tilde{m}_1) + W_2^2(m_2, \tilde{m}_2) \tag{1.4}$$

for any $m, \tilde{m} \in \mathcal{P}(\mathbb{R}^{k_1} \times \mathbb{R}^{k_2})$ with marginals $m_1, \tilde{m}_1 \in \mathcal{P}(\mathbb{R}^{k_1})$ and $m_2, \tilde{m}_2 \in \mathcal{P}(\mathbb{R}^{k_2})$ respectively, where $k_1, k_2 \in \mathbb{N}$. In addition, if $m = m_1 \otimes m_2$ and $\tilde{m} = \tilde{m}_1 \otimes \tilde{m}_2$ are product measures, then

$$W_2^2(m, \tilde{m}) = W_2^2(m_1, \tilde{m}_1) + W_2^2(m_2, \tilde{m}_2). \tag{1.5}$$

## 2. MODEL AND ASSUMPTIONS

We follow the notation used in [27, 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 $0 \leq G \leq 1$. 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*

$$\|G\| := \|G\|_{\infty \to 1} := \sup_{\|g\|_\infty \leq 1} \|Gg\|_1 = \sup_{\|g\|_\infty \leq 1} \int_I \left| \int_I G(u, v) g(v) \, dv \right| du.$$

*From [27, 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 $G \in \mathcal{G}$ and a collection of initial distributions $\bar{\mu}(0) := (\bar{\mu}_u(0) \in \mathcal{P}(\mathbb{R}^d) : u \in I)$, recall the graphon particle system (1.1) and the finite particle system (1.2). The following assumptions will be made throughout the paper.

**Standing Assumptions:**
- The map $I \ni u \mapsto \bar{\mu}_u(0) := \mathcal{L}(\bar{X}_u(0)) \in \mathcal{P}(\mathbb{R}^d)$ is measurable, and $\sup_{u \in I} \mathbb{E}|\bar{X}_u(0)|^4 < \infty$.
- The drift functions $f$ and $b$ are Lipschiz with Lipschitz constant $K_f$ and $K_b$, respectively, namely

$$|f(x_1) - f(x_2)| \leq K_f |x_1 - x_2|, \quad \forall x_1, x_2 \in \mathbb{R}^d,$$

$$|b(x_1, y_1) - b(x_2, y_2)| \leq K_b(|x_1 - x_2| + |y_1 - y_2|), \quad \forall x_1, x_2, y_1, y_2 \in \mathbb{R}^d.$$

- Dissipativity: There exists some $c_0 \in (0, \infty)$ such that

$$(x_1 - x_2) \cdot (f(x_1) - f(x_2)) \leq -c_0 |x_1 - x_2|^2, \quad \forall x_1, x_2 \in \mathbb{R}^d \tag{2.1}$$

  and

$$\kappa := c_0 - 2K_b > 0. \tag{2.2}$$

- $G_n \in \mathcal{G}$ is a graphon and
  (i) either $\xi_{ij}^n = G_n(\frac{i}{n}, \frac{j}{n})$,
  (ii) or $\xi_{ij}^n = \xi_{ji}^n = \text{Bernoulli}(G_n(\frac{i}{n}, \frac{j}{n}))$ independently for $1 \leq i \leq j \leq n$, and independent of $\{\bar{X}_u(0), B_u : u \in I\}$.

**Remark 2.2.** *(a) The finite forth moment on $\bar{X}(0)$ is assumed to obtain Wasserstein-2 estimates in Sections 5 and 6. A weaker condition such as a finite second moment on $\bar{X}(0)$ would be sufficient to establish exponential ergodicity properties in Sections 3 and 4.*
*(b) Clearly, $f$ and $b$ have linear growth, namely there exists some $C \in (0, \infty)$ such that $|f(x)| + |b(x, y)| \leq C(1 + |x| + |y|)$ for all $x, y \in \mathbb{R}^d$.*

(c) *By the Lipschitz property of $f$, there exists some $c_0 \in [-K_f, K_f]$ such that (2.1) holds. Here the dissipativity assumption requires that $c_0 > 2K_b$. A common example of $b$ and $f$ satisfying (2.1) and (2.2) is linear (as in the study of linear quadratic graphon mean-field games in, e.g., [21]) and mean-reverting:*

$$f(x) + b(x, y) = -c_1 x + c_2 y, \quad \text{for some } c_1 > c_2 > 0.$$

*In particular, the choice of $f(x) = -(c_1 + c_2)x$ and $b(x, y) = c_2(x + y)$ satisfies (2.1) and (2.2) since $c_0 = c_1 + c_2 > 2c_2 = 2K_b$.*

The following result gives well-posedness of systems (1.1) and (1.2).

**Proposition 2.1.** *(a) There exists a unique pathwise solution to (1.1). For every $T < \infty$, the map $I \ni u \mapsto \bar{\mu}_u \in \mathcal{P}(\mathbb{C}([0, T] : \mathbb{R}^d))$ is measurable and*

$$\sup_{u \in I} \sup_{t \in [0,T]} \mathbb{E}\left[|\bar{X}_u(t)|^4\right] < \infty.$$

*(b) There exists a unique pathwise solution to (1.2). Also for every $T < \infty$,*

$$\max_{i=1,\dots,n} \sup_{t \in [0,T]} \mathbb{E}\left[|X_i^n(t)|^4\right] < \infty.$$

The proof of Proposition 2.1 is standard (see e.g. [36] and [3] for part (a), and [24, Theorems 5.2.5 and 5.2.9] for part (b)) and hence is omitted.

## 3. Exponential ergodicity of the graphon particle system

In this section we show the exponential ergodicity for the graphon particle system (1.1).

First recall the standing assumptions in Section 2. The following result guarantees that Proposition 2.1(a) holds uniformly in time.

**Proposition 3.1.** *The system (1.1) has finite fourth moments uniformly in time, namely*

$$\sup_{u \in I} \sup_{t \geq 0} \mathbb{E}\left[|\bar{X}_u(t)|^4\right] < \infty.$$

Next we introduce some notations before stating the exponential ergodicity property. For $\eta := (\eta_u : u \in I) \in [\mathcal{P}(\mathbb{R}^d)]^I$ with $\sup_{u \in I} \int_{\mathbb{R}^d} |x|^4 \, \eta_u(dx) < \infty$, consider the system $\bar{Y}^\eta = (\bar{Y}_u^\eta : u \in I)$ given by

$$\bar{Y}_u^\eta(t) = \bar{Y}_u^\eta(0) + \int_0^t \left( f(\bar{Y}_u^\eta(s)) + \int_I \int_{\mathbb{R}^d} b(\bar{Y}_u^\eta(s), x) G(u, v) \, \bar{\mu}_{v,s}^\eta(dx) \, dv \right) ds \tag{3.1}$$
$$+ \sigma B_u(t), \quad \bar{\mu}_{u,t}^\eta = \mathcal{L}(\bar{Y}_u^\eta(t)), \quad u \in I,$$

where $(\bar{Y}_u^\eta(0) : u \in I)$ are mutually independent and also independent of $\{B_u : u \in I\}$ with $\mathcal{L}(\bar{Y}_u^\eta(0)) = \eta_u$. Note that $\bar{Y}^\eta$ is well-defined and $\sup_{u \in I} \sup_{t \geq 0} \int_{\mathbb{R}^d} |x|^4 \, \bar{\mu}_{u,t}^\eta(dx) < \infty$ by Proposition 2.1(a). Denote by $P_t$ the associated Markov semigroup:

$$P_t \eta := \mathcal{L}(\bar{Y}^\eta(t)), \quad t \geq 0. \tag{3.2}$$

The following theorem shows that $\bar{\mu}_{u,t}$ (and its average) converges exponentially fast to the limiting distribution, which is also invariant with respect to $P_t$.

**Theorem 3.1.** *(a) There exists a unique collection of probability measures $(\bar{\mu}_{u,\infty} : u \in I)$ such that*

$$\sup_{u \in I} W_2(\bar{\mu}_{u,t}, \bar{\mu}_{u,\infty}) \leq \sqrt{4\kappa_1 \frac{c_0 - K_b}{\kappa}} e^{-\kappa t/2}, \quad t \geq 0, \tag{3.3}$$

*and hence*

$$W_2(\hat{\mu}(t), \hat{\mu}(\infty)) \leq \sqrt{4\kappa_1 \frac{c_0 - K_b}{\kappa}} e^{-\kappa t/2}, \quad t \geq 0, \tag{3.4}$$

*where $\kappa_1 := \sup_{u \in I} \sup_{t \geq 0} \mathbb{E}|\bar{X}_u(t)|^2$ and the averaged measures $\hat{\mu}(t)$ and $\hat{\mu}(\infty)$ are defined as*

$$\hat{\mu}(t) := \int_I \bar{\mu}_{u,t} \, du, \quad \hat{\mu}(\infty) := \int_I \bar{\mu}_{u,\infty} \, du. \tag{3.5}$$

(b) *The collection $\bar{\mu}(\infty) := (\bar{\mu}_{u,\infty} : u \in I)$ is invariant with respect to the Markov semigroup $P_t$ defined in (3.2), namely*

$$P_t \bar{\mu}(\infty) = \bar{\mu}(\infty), \quad t \geq 0.$$

(c) *There exists some $C \in (0, \infty)$ such that*

$$\sup_{t \in [0,\infty]} W_2(\bar{\mu}_{u_1,t}, \bar{\mu}_{u_2,t}) \leq \max \left\{ W_2(\bar{\mu}_{u_1,0}, \bar{\mu}_{u_2,0}), C \int_I |G(u_1, v) - G(u_2, v)| \, dv \right\}, \quad u_1, u_2 \in I.$$

Proofs of Proposition 3.1 and Theorem 3.1 are given in Section 7.1.

An immediate consequence of Theorem 3.1(c) is that the marginal distribution is (Lipschitz) continuous as long as the initial distribution and the graphon are so.

**Condition 3.1.** *There exists a finite collection of intervals $\{I_i : i = 1, \ldots, N\}$ for some $N \in \mathbb{N}$, such that $\cup_{i=1}^N I_i = I$ and for each $i \in \{1, \ldots, N\}$:*

(a) *The map $I_i \ni u \mapsto \bar{\mu}_u(0) \in \mathcal{P}(\mathbb{R}^d)$ 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 $\lambda_I$ denotes the Lebesgue measure on $I$.*

**Condition 3.2.** *There exist some $K_G \in (0, \infty)$ and a finite collection of intervals $\{I_i : i = 1, \ldots, N\}$ for some $N \in \mathbb{N}$, such that $\cup_{i=1}^N I_i = I$ and*

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

**Corollary 3.1.** (a) *Suppose Condition 3.1 holds. Then for each $i \in \{1, \ldots, N\}$, $\sup_{t \in [0,\infty]} W_2(\bar{\mu}_{u_1,t}, \bar{\mu}_{u_2,t}) \to 0$ whenever $u_1 \to u_2$ in $I_i$.*
(b) *Suppose Condition 3.2 holds. Then there exists some $C \in (0, \infty)$ such that*

$$\sup_{t \in [0,\infty]} W_2(\bar{\mu}_{u_1,t}, \bar{\mu}_{u_2,t}) \leq C |u_1 - u_2|$$

*whenever $u_1, u_2 \in I_i$ for some $i \in \{1, \ldots, N\}$.*

*Proof.* This is immediate from Theorem 3.1(c). ∎

We note that, as illustrated through the following example of Gaussian processes with linear coefficients, the graph structure plays a crucial role in the long-time behavior of the system, and hence the stationary measure, due to the heterogeneity of the system, is not necessarily tractable. This is indeed one of the main reasons we are interested in the second question in the introduction, which is answered via the uniform-in-time convergence and Euler discretization in the next few sections.

**Example 3.1.** *Suppose $d = 1$. Suppose $f$ and $b$ are linear, namely $f(x) = c_1 - c_2 x$ and $b(x, y) = c_3 + c_4 x + c_5 y$. Suppose $c_2 > 0$ and $c_2 - 2 \max\{|c_4|, |c_5|\} > 0$ so that the dissipativity assumption holds. Then (1.1) is a collection of Gaussian processes. Letting $m_u(t) := \mathbb{E}[\bar{X}_u(t)]$ and $M_u(t) := \mathbb{E}[\bar{X}_u^2(t)]$, we have*

$$m_u(t) = m_u(0) + \int_0^t \left( c_1 - c_2 m_u(s) + \int_0^1 (c_3 + c_4 m_u(s) + c_5 m_v(s)) G(u, v) \, dv \right) ds,$$

$$M_u(t) = M_u(0) + \mathbb{E}\left[ \int_0^t 2\bar{X}_u(s) \, d\bar{X}_u(s) \right] + \sigma^2 t$$

$$= M_u(0) + 2 \int_0^t \Big( c_1 m_u(s) - c_2 M_u(s)$$

$$+ \int_0^1 (c_3 m_u(s) + c_4 M_u(s) + c_5 m_u(s) m_v(s)) G(u, v) \, dv \Big) ds + \sigma^2 t.$$

*By Theorem 3.1, the limits $m_u(\infty) := \lim_{t \to \infty} m_u(t)$ and $M_u(\infty) := \lim_{t \to \infty} M_u(t)$ exist and should satisfy the following equations:*

$$c_1 - c_2 m_u(\infty) + \int_0^1 (c_3 + c_4 m_u(\infty) + c_5 m_v(\infty)) G(u, v) \, dv = 0,$$

$$c_1 m_u(\infty) - c_2 M_u(\infty) + \int_0^1 (c_3 m_u(\infty) + c_4 M_u(\infty) + c_5 m_u(\infty) m_v(\infty)) G(u, v) \, dv + \frac{1}{2} \sigma^2 = 0.$$

*Here the first equation is a Fredholm integral equation of the second kind, from which $m_u(\infty)$ could be written as a Liouville–Neumann series, and $M_u(\infty)$ could then be solved from the second equation; see e.g. [34, Chapter 13] for methods of solving such integral equations.*

*We note that even in this setup of linear systems, the long time behavior crucially depends on the graphon $G$ and the stationary distribution, such as the mean $m_u(\infty)$, is not necessarily explicit or tractable. In some special cases, one can get explicit expressions. For example, if we further assume $c_1 = c_3 = 0$, then we can get*

$$m_u(\infty) = 0, \quad M_u(\infty) = \frac{\sigma^2}{2(c_2 - c_4 \int_0^1 G(u, v) \, dv)},$$

*and the second moment of the averaged measure $\hat{\mu}(\infty)$ is*

$$\int_{\mathbb{R}} x^2 \, \hat{\mu}(\infty)(dx) = \int_0^1 M_u(\infty) \, du = \frac{\sigma^2}{2} \int_0^1 \frac{1}{c_2 - c_4 \int_0^1 G(u, v) \, dv} \, du.$$

## 4. EXPONENTIAL ERGODICITY OF THE FINITE PARTICLE SYSTEM

In this section we establish the exponential ergodicity of the joint distribution for the finite particle system (1.2).

Using the standing assumptions in Section 2, we first show that Proposition 2.1(b) holds uniformly in time, in the quenched sense by conditioning on the random interactions $\xi_{ij}^n$, and hence also in the annealed sense. Write $\xi^n := (\xi_{ij}^n)_{i,j=1}^n$ and

$$\mathbb{E}^{n,z}[\cdot] := \mathbb{E}[\cdot \mid \xi^n = (z_{ij})_{i,j=1}^n], \quad (z_{ij} = z_{ji})_{i,j=1}^n \in [0, 1]^{n^2}.$$

**Proposition 4.1.** *There exists some constant $\kappa_2 \in (0, \infty)$ such that*

$$\sup_{n \in \mathbb{N}} \max_{i=1,\ldots,n} \sup_{t \geq 0} \mathbb{E}^{n,\xi^n} \left[ |X_i^n(t)|^2 \right] \leq \kappa_2 \ a.s., \quad \sup_{n \in \mathbb{N}} \max_{i=1,\ldots,n} \sup_{t \geq 0} \mathbb{E} \left[ |X_i^n(t)|^2 \right] \leq \kappa_2.$$

Next we introduce some notations before stating the exponential ergodicity property. Define the annealed and quenched joint distributions at time $t \geq 0$ by

$$\theta^n(t) := \mathcal{L}((X_i^n(t))_{i=1}^n) \in \mathcal{P}((\mathbb{R}^d)^n)$$

and

$$\theta^{n,z}(t) := \mathcal{L}((X_i^n(t))_{i=1}^n \mid \xi^n = (z_{ij})_{i,j=1}^n), \quad z = (z_{ij} = z_{ji})_{i,j=1}^n \in [0,1]^{n^2}.$$

For $\eta \in \mathcal{P}((\mathbb{R}^d)^n)$ and $z = (z_{ij} = z_{ji})_{i,j=1}^n \in [0,1]^{n^2}$, consider the system $Y^{n,z,\eta} = (Y_i^{n,z,\eta})_{i=1}^n$ given by

$$Y_i^{n,z,\eta}(t) = Y_i^{n,z,\eta}(0) + \int_0^t \left( f(Y_i^{n,z,\eta}(s)) + \frac{1}{n} \sum_{j=1}^n z_{ij} b(Y_i^{n,z,\eta}(s), Y_j^{n,z,\eta}(s)) \right) ds \qquad (4.1)$$

$$+ \sigma B_{\frac{i}{n}}(t), \quad i \in \{1, \ldots, n\},$$

where $Y^{n,z,\eta}(0)$ is independent of $\{B_u : u \in I\}$ with $\mathcal{L}(Y^{n,z,\eta}(0)) = \eta$. Denote by $P_t^{n,z}$ the associated Markov semigroup:

$$P_t^{n,z}\eta := \mathcal{L}(Y^{n,z,\eta}(t)), \quad t \geq 0, \quad \eta \in \mathcal{P}((\mathbb{R}^d)^n). \qquad (4.2)$$

The following theorem shows that $\theta^{n,z}(t)$ (resp. $\theta^n(t)$) converges exponentially fast to the limiting distribution, which is also invariant with respect to $P_t^{n,z}$.

**Theorem 4.1.** *(a) There exists a unique collection of probability measures $\{\theta^{n,z}(\infty) : z = (z_{ij} = z_{ji})_{i,j=1}^n \in [0,1]^{n^2}\}$ such that*

$$\sup_{n \in \mathbb{N}} \frac{1}{\sqrt{n}} W_2(\theta^{n,\xi^n}(t), \theta^{n,\xi^n}(\infty)) \leq \sqrt{4\kappa_2} e^{-\kappa t}, \quad t \geq 0, \quad a.s., \qquad (4.3)$$

*and hence*

$$\sup_{n \in \mathbb{N}} \frac{1}{\sqrt{n}} W_2(\theta^n(t), \theta^n(\infty)) \leq \sqrt{4\kappa_2} e^{-\kappa t}, \quad t \geq 0, \qquad (4.4)$$

*where $\kappa_2$ is as in Proposition 4.1, and*

$$\theta^n(\infty) := \mathbb{E}[\theta^{n,\xi^n}(\infty)]. \qquad (4.5)$$

*(b) The joint distribution $\theta^{n,z}(\infty)$ is invariant with respect to the Markov semigroup $P_t^{n,z}$ defined in (4.2), namely*

$$P_t^{n,z}\theta^{n,z}(\infty) = \theta^{n,z}(\infty), \quad t \geq 0.$$

Proofs of Proposition 4.1 and Theorem 4.1 are given in Section 7.2.

## 5. UNIFORM-IN-TIME CONVERGENCE

In this section we analyze the uniform-in-time convergence of the finite particle system (1.2) to the graphon particle system (1.1).

We make the following assumption on the kernel $G_n$. Note that (5.1) is just a convenient and natural form to view $(G_n(\frac{i}{n}, \frac{j}{n}) : i, j = 1, \ldots, n)$ as a piece-wise constant graphon.

**Condition 5.1.** *$G_n$ is a step graphon, that is,*

$$G_n(u, v) = G_n\left( \frac{\lceil nu \rceil}{n}, \frac{\lceil nv \rceil}{n} \right), \quad for \ (u, v) \in I \times I. \qquad (5.1)$$

*Moreover, $G_n \to G$ in the cut metric as $n \to \infty$.*

**Remark 5.1.** *In general, if $\delta_\square(G_n, G) \to 0$ for a sequence of step graphons, then it follows from [27, Theorem 11.59] that $\|G_n - G\|_\square \to 0$, after suitable relabeling of $G_n$. Therefore we directly assume in Condition 5.1 that the convergence of $G_n$ to $G$ is in the cut metric $d_\square$, instead of assuming that $d_\square(G_n^{\varphi_n}, G) \to 0$ for some relabeling function $\varphi_n$ and $X_i^n(0) = \bar{X}_{\varphi_n^{-1}(\frac{i}{n})}(0)$.*

The following convergence on finite time intervals was shown in [3, Theorem 3.1] without the dissipativity assumption.

**Proposition 5.1.** *([3, Theorem 3.1]) Suppose Conditions 3.1 and 5.1 hold. Fix $T \in (0, \infty)$. As $n \to \infty$,*

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[ \sup_{t \in [0,T]} \left| X_i^n(t) - \bar{X}_{\frac{i}{n}}(t) \right|^2 \right] \to 0.$$

Recall $\hat{\mu}(t)$ and $\hat{\mu}(\infty)$ introduced in (3.5). Let

$$\nu^n(t) := \frac{1}{n} \sum_{i=1}^n \delta_{X_i^n(t)}, \quad \mu^n(t) := \frac{1}{n} \sum_{i=1}^n \mathcal{L}(X_i^n(t)) = \mathbb{E}\nu^n(t).$$

Although Proposition 5.1 holds for finite time horizon, it does not provide sufficient information about the convergence of stationary measures. Under the dissipativity assumption, we have the following uniform in time convergence of $X_i^n$ and LLN of $\nu^n$ and $\mu^n$, which in particular guarantees the convergence of stationary measures (see Corollary 5.3).

**Theorem 5.1.** *Suppose Conditions 3.1 and 5.1 hold.*
*(a) As $n \to \infty$,*

$$\sup_{t \geq 0} \frac{1}{n} \sum_{i=1}^n \mathbb{E}|X_i^n(t) - \bar{X}_{\frac{i}{n}}(t)|^2 \to 0.$$

*(b) (LLN) As $n \to \infty$,*

$$\sup_{t \geq 0} W_2(\mu^n(t), \hat{\mu}(t)) \to 0, \quad \sup_{t \geq 0} \mathbb{E}W_2(\nu^n(t), \hat{\mu}(t)) \to 0. \tag{5.2}$$

**Remark 5.2.** *(a) We note that Theorem 5.1 and many existing results (such as [2, Theorem 2.1(b)], [8, Theorem 3.4], [10, Theorem 3.4] and [38, Theorem 2]) work on the marginal distributions, which is sufficient for the analysis of approximating stationary measure in Corollary 6.2. The study on the trajectory level of the difference $\mathbb{E}\left[\sup_{t \geq 0} \frac{1}{n} \sum_{i=1}^n |X_i^n(t) - \bar{X}_{\frac{i}{n}}(t)|^2\right]$ is more challenging and beyond the scope of this work.*
*(b) We also note that graphs with vanishing density degrees and rescaled strength of interactions are analyzed in [3] and Proposition 5.1 is proved under certain conditions via a Girsanov's change of measure argument. It is challenging to apply such an argument to the long-time analysis. The study of the uniform-in-time convergence (and Euler discretization) for such graphs will be the future work.*

The following condition will be used for analyzing efficient Euler discretization and simulation in Section 6.

**Condition 5.2.** *$G_n$ is a graphon such that $G_n(\frac{i}{n}, \frac{j}{n}) = G(\frac{i}{n}, \frac{j}{n})$ for each $i, j \in \{1, \ldots, n\}$.*

**Remark 5.3.** *We note that Condition 5.2 is trivially satisfied if $G_n = G$. Alternatively, one may take $G_n$ to be a step graphon that is consistent with $G$:*

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

Condition 5.2 allows one to obtain POC and the rate of convergence in Theorem 5.1. Let

$$a(n) := n^{-1/d} + n^{-1/12}. \tag{5.3}$$

**Theorem 5.2.** *Suppose Conditions 3.2 and 5.2 hold. Then there exists some $C \in (0, \infty)$ such that the following hold.*

*(a) For all $n \in \mathbb{N}$,*

$$\sup_{t \geq 0} \max_{i=1,\ldots,n} \mathbb{E}|X_i^n(t) - \bar{X}_{\frac{i}{n}}(t)|^2 \leq \frac{C}{n}.$$

*(b) (LLN) For all $n \in \mathbb{N}$,*

$$\sup_{t \geq 0} W_2(\mu^n(t), \hat{\mu}(t)) \leq \frac{C}{\sqrt{n}}, \quad \sup_{t \geq 0} \mathbb{E}W_2(\nu^n(t), \hat{\mu}(t)) \leq Ca(n).$$

*(c) (POC) For all $n, k \in \mathbb{N}$ and any distinct $i_1, \ldots, i_k \in \{1, \ldots, n\}$,*

$$\sup_{t \geq 0} W_2(\mathcal{L}(X_{i_1}^n(t), \ldots, X_{i_k}^n(t)), \bar{\mu}_{\frac{i_1}{n},t} \otimes \cdots \otimes \bar{\mu}_{\frac{i_k}{n},t}) \leq \frac{C\sqrt{k}}{\sqrt{n}}.$$

Proofs of Theorems 5.1 and 5.2 are given in Section 7.3.

**Remark 5.4.** *The rate $a(n)$ is related to the upper bound of the Wasserstein distance between the empirical measure of independent random variables and their averaged distribution. It may be replaced by other function of $n$ that vanishes faster, as a result of which the constant $C$ in Theorem 5.2(b) will be larger; see Remark 7.1.*

As an immediate consequence of the exponential ergodicity of the graphon particle system (1.1) and the uniform-in-time convergence, one has the interchange of limits as $t \to \infty$ and $n \to \infty$.

**Corollary 5.1.** *Suppose Conditions 3.1 and 5.1 hold. Then*

$$\lim_{n,t \to \infty} W_2(\mu^n(t), \hat{\mu}(\infty)) = 0, \quad \lim_{n,t \to \infty} \mathbb{E}W_2(\nu^n(t), \hat{\mu}(\infty)) = 0.$$

*Proof.* This follows from (3.4) and Theorem 5.1(b). ∎

**Corollary 5.2.** *Suppose Conditions 3.2 and 5.2 hold. Then there exists $C \in (0, \infty)$ such that the following hold.*

*(a) For all $n \in \mathbb{N}$ and $t \geq 0$,*

$$W_2(\mu^n(t), \hat{\mu}(\infty)) \leq C\left(\frac{1}{\sqrt{n}} + e^{-\kappa t/2}\right), \quad \mathbb{E}W_2(\nu^n(t), \hat{\mu}(\infty)) \leq C\left(a(n) + e^{-\kappa t/2}\right).$$

*In particular,*

$$\lim_{n,t \to \infty} W_2(\mu^n(t), \hat{\mu}(\infty)) = 0, \quad \lim_{n,t \to \infty} \mathbb{E}W_2(\nu^n(t), \hat{\mu}(\infty)) = 0.$$

*(b) For all $n, k \in \mathbb{N}$, $t \geq 0$ and any distinct $i_1, \ldots, i_k \in \{1, \ldots, n\}$,*

$$W_2(\mathcal{L}(X_{i_1}^n(t), \ldots, X_{i_k}^n(t)), \bar{\mu}_{\frac{i_1}{n},\infty} \otimes \cdots \otimes \bar{\mu}_{\frac{i_k}{n},\infty}) \leq C\sqrt{k}\left(\frac{1}{\sqrt{n}} + e^{-\kappa t/2}\right).$$

*Proof.* (a) This follows from (3.4) and Theorem 5.2(b).
(b) This follows from Theorem 5.2(c), (1.5) and (3.3). ∎

From Theorem 4.1(a) we know that the limiting distribution

$$\mu^n(\infty) := \lim_{t \to \infty} \mu^n(t) = \lim_{t \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mathcal{L}(X_i^n(t))$$

is well-defined. The following corollary shows that $\mu^n(\infty)$ converges to the averaged long time distribution $\hat{\mu}(\infty)$ of the graphon particle system (1.1).

**Corollary 5.3.** *(a) Suppose Conditions* 3.1 *and* 5.1 *hold. Then*

$$\lim_{n \to \infty} W_2(\mu^n(\infty), \hat{\mu}(\infty)) = 0.$$

*(b) Suppose Conditions* 3.2 *and* 5.2 *hold. Then there exists* $C \in (0, \infty)$ *such that*

$$W_2(\mu^n(\infty), \hat{\mu}(\infty)) \leq \frac{C}{\sqrt{n}}.$$

*In addition, for all* $n, k \in \mathbb{N}$, $t \geq 0$ *and any distinct* $i_1, \ldots, i_k \in \{1, \ldots, n\}$,

$$W_2(\lim_{t \to \infty} \mathcal{L}(X_{i_1}^n(t), \ldots, X_{i_k}^n(t)), \bar{\mu}_{\frac{i_1}{n}, \infty} \otimes \cdots \otimes \bar{\mu}_{\frac{i_k}{n}, \infty}) \leq \frac{C\sqrt{k}}{\sqrt{n}}.$$

*Proof.* Write

$$W_2(\mu^n(\infty), \hat{\mu}(\infty)) \leq W_2(\mu^n(\infty), \mu^n(t)) + W_2(\mu^n(t), \hat{\mu}(\infty)).$$

Using the convexity of $W_2^2(\cdot, \cdot)$, (1.4) and (4.4), we have

$$W_2^2(\mu^n(\infty), \mu^n(t)) \leq \frac{1}{n} \sum_{i=1}^{n} W_2^2(\lim_{s \to \infty} \mathcal{L}(X_i^n(s)), \mathcal{L}(X_i^n(t))) \leq \frac{1}{n} W_2^2(\theta^n(\infty), \theta^n(t)) \leq C e^{-2\kappa t}$$

for each $t \geq 0$. Combining these with Corollary 5.1 (resp. Corollary 5.2(a)) gives part (a) (resp. the first statement in part (b)). The second statement in part (b) follows by taking $t \to \infty$ in Corollary 5.2(b). ∎

## 6. EULER DISCRETIZATION

In this section we analyze the Euler discretization of the system (1.2) with step size $h > 0$, namely, with $s_h := \lfloor \frac{s}{h} \rfloor h$,

$$X_i^{n,h}(t) = \bar{X}_{\frac{i}{n}}(0) + \int_0^t \left( f(X_i^{n,h}(s_h)) + \frac{1}{n} \sum_{j=1}^{n} \xi_{ij}^n b(X_i^{n,h}(s_h), X_j^{n,h}(s_h)) \right) ds$$

$$+ \sigma B_{\frac{i}{n}}(t), \quad i \in \{1, \ldots, n\}, \, t \geq 0. \tag{6.1}$$

The following theorem shows the convergence of the Euler scheme, uniformly in time $t$ and the number of particles $n$. The proof is given in Section 7.4.

**Theorem 6.1.** *There exist* $C, h_0 \in (0, \infty)$ *such that*

$$\sup_{n \in \mathbb{N}} \max_{i=1, \ldots, n} \sup_{t \geq 0} \mathbb{E}|X_i^{n,h}(t) - X_i^n(t)|^2 \leq Ch, \quad \forall \, h \in (0, h_0).$$

Theorem 6.1 and Corollary 5.1 guarantee that the Euler scheme (6.1) provides a good numerical approximation to the graphon particle system (1.1) uniformly in time, as shown in the following corollary. Let

$$\nu^{n,h}(t) := \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i^{n,h}(t)}, \quad \mu^{n,h}(t) := \frac{1}{n} \sum_{i=1}^{n} \mathcal{L}(X_i^{n,h}(t)) = \mathbb{E}\nu^{n,h}(t).$$

**Corollary 6.1.** *Suppose Conditions* 3.1 *and* 5.1 *hold. Then there exist* $C, h_0 \in (0, \infty)$ *such that*

$$\limsup_{n \to \infty} \sup_{t \geq 0} W_2(\mu^{n,h}(t), \hat{\mu}(t)) \leq C\sqrt{h}, \quad \limsup_{n \to \infty} \sup_{t \geq 0} \mathbb{E}W_2(\nu^{n,h}(t), \hat{\mu}(t)) \leq C\sqrt{h},$$

*and*

$$\limsup_{n,t \to \infty} W_2(\mu^{n,h}(t), \hat{\mu}(\infty)) \leq C\sqrt{h}, \quad \limsup_{n,t \to \infty} \mathbb{E}W_2(\nu^{n,h}(t), \hat{\mu}(\infty)) \leq C\sqrt{h},$$

*for all* $h \in (0, h_0)$.

*Proof.* Let $h_0 \in (0, \infty)$ be as in Theorem 6.1. Taking $\pi = \frac{1}{n} \sum_{i=1}^n \mathcal{L}(X_i^{n,h}(t), X_i^n(t))$ as the coupling of $\mu^{n,h}(t)$ and $\mu^n(t)$ gives

$$\sup_{n \in \mathbb{N}} \sup_{t \geq 0} W_2(\mu^{n,h}(t), \mu^n(t)) \leq \sup_{n \in \mathbb{N}} \sup_{t \geq 0} \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}|X_i^{n,h}(t) - X_i^n(t)|^2 \right)^{1/2} \leq C\sqrt{h}, \qquad (6.2)$$

and taking $\pi = \frac{1}{n} \sum_{i=1}^n \delta_{(X_i^{n,h}(t), X_i^n(t))}$ as the coupling of $\nu^{n,h}(t)$ and $\nu^n(t)$ gives

$$\sup_{n \in \mathbb{N}} \sup_{t \geq 0} \mathbb{E}W_2(\nu^{n,h}(t), \nu^n(t)) \leq \sup_{n \in \mathbb{N}} \sup_{t \geq 0} \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}|X_i^{n,h}(t) - X_i^n(t)|^2 \right)^{1/2} \leq C\sqrt{h}, \qquad (6.3)$$

for all $h \in (0, h_0)$. Combining these with Theorem 5.1(b) (resp. Corollary 5.1) gives the first (resp. second) statement. This completes the proof. ∎

As stated in the second question in the introduction, we are also interested in the precise rate of convergence of the Euler scheme, as an approximation to the graphon particle system (1.1) and its stationary distribution. This is answered in the following corollary.

**Corollary 6.2.** *Suppose Conditions* 3.2 *and* 5.2 *hold. Then there exist* $C, h_0 \in (0, \infty)$ *such that the following hold.*

*(a) For all* $h \in (0, h_0)$, $n \in \mathbb{N}$ *and* $t \geq 0$,

$$W_2(\mu^{n,h}(t), \hat{\mu}(t)) \leq C\left( \frac{1}{\sqrt{n}} + \sqrt{h} \right), \quad \mathbb{E}W_2(\nu^{n,h}(t), \hat{\mu}(t)) \leq C\left( a(n) + \sqrt{h} \right),$$

*and*

$$W_2(\mu^{n,h}(t), \hat{\mu}(\infty)) \leq C\left( \frac{1}{\sqrt{n}} + \sqrt{h} + e^{-\kappa t/2} \right), \mathbb{E}W_2(\nu^{n,h}(t), \hat{\mu}(\infty)) \leq C\left( a(n) + \sqrt{h} + e^{-\kappa t/2} \right).$$

*(b) For all* $n, k \in \mathbb{N}$, $t \geq 0$ *and any distinct* $i_1, \ldots, i_k \in \{1, \ldots, n\}$,

$$W_2(\mathcal{L}(X_{i_1}^{n,h}(t), \ldots, X_{i_k}^{n,h}(t)), \bar{\mu}_{\frac{i_1}{n},t} \otimes \cdots \otimes \bar{\mu}_{\frac{i_k}{n},t}) \leq C\sqrt{k} \left( \frac{1}{\sqrt{n}} + \sqrt{h} \right),$$

*and*

$$W_2(\mathcal{L}(X_{i_1}^{n,h}(t), \ldots, X_{i_k}^{n,h}(t)), \bar{\mu}_{\frac{i_1}{n},\infty} \otimes \cdots \otimes \bar{\mu}_{\frac{i_k}{n},\infty}) \leq C\sqrt{k} \left( \frac{1}{\sqrt{n}} + \sqrt{h} + e^{-\kappa t/2} \right).$$

*Proof.* (a) Combining (6.2), (6.3) and Theorem 5.2(b) (resp. Corollary 5.2(a)) gives the first (resp. last) two statements.

(b) Taking $\pi = \mathcal{L}\left(\left(X_{i_1}^{n,h}(t), \ldots, X_{i_k}^{n,h}(t)\right), \left(X_{i_1}^n(t), \ldots, X_{i_k}^n(t)\right)\right)$, we have

$$W_2(\mathcal{L}(X_{i_1}^{n,h}(t), \ldots, X_{i_k}^{n,h}(t)), \mathcal{L}(X_{i_1}^n(t), \ldots, X_{i_k}^n(t))) \leq \left(\sum_{j=1}^k \mathbb{E}|X_{i_j}^{n,h}(t) - X_{i_j}^n(t)|^2\right)^{1/2}$$
$$\leq C\sqrt{kh},$$

where the last line uses Theorem 6.1. Combining this with Theorem 5.2(c) (resp. Corollary 5.2(b)) gives the first (resp. second) statement. This completes the proof. ∎

**Remark 6.1.** *The constant $C$ here could be made explicit, but we didn't explore that direction.*

## 7. Proofs

We first present an elementary result that will be used in several later proofs.

**Lemma 7.1.** *Let $y \colon [0, \infty) \to [0, \infty)$ be a non-negative differentiable function. Suppose*

$$y(t) - y(r) \leq -a_1 \int_r^t y(s)\, ds + a_2 \int_r^t \sqrt{y(s)}\, ds + \int_r^t a_3(s)\, ds, \quad \forall\, t > r \geq 0,$$

*for some $a_1 > 0$, $a_2 \in \mathbb{R}$ and non-negative and continuous function $a_3$. Then*

$$y(t) \leq \max\left\{y(0), \left(\frac{a_2}{2a_1} + \sqrt{\frac{\sup_{0 \leq s \leq t} a_3(s)}{a_1} + \frac{a_2^2}{4a_1^2}}\right)^2\right\}, \quad \forall\, t \geq 0.$$

*In particular, $y(t) \leq \max\left\{y(0), \frac{\sup_{0 \leq s \leq t} a_3(s)}{a_1}\right\}$ if $a_2 = 0$.*

*Proof.* Fix $T \in (0, \infty)$. Since $y(t)$ is differentiable, we have

$$y'(t) \leq -a_1 y(t) + a_2\sqrt{y(t)} + a_3(t) \leq -a_1\left(\sqrt{y(t)} - \frac{a_2}{2a_1}\right)^2 + \sup_{0 \leq s \leq T} a_3(s) + \frac{a_2^2}{4a_1}, \quad t \in [0, T].$$

Noting that the right hand side above is negative when

$$\sqrt{y(t)} > \frac{a_2}{2a_1} + \sqrt{\frac{\sup_{0 \leq s \leq T} a_3(s)}{a_1} + \frac{a_2^2}{4a_1^2}},$$

we have the desired result. ∎

### 7.1. **Proofs for Section 3.**

*Proof of Proposition* 3.1. Using Itô's formula, Remark 2.2(b) and Proposition 2.1(a), we have

$$\mathbb{E}|\bar{X}_u(t)|^4 - \mathbb{E}|\bar{X}_u(0)|^4$$
$$= \mathbb{E}\int_0^t 4|\bar{X}_u(s)|^2 \bar{X}_u(s) \cdot \left(f(\bar{X}_u(s)) + \int_I \int_{\mathbb{R}^d} b(\bar{X}_u(s), x) G(u, v)\, \bar{\mu}_{v,s}(dx)\, dv\right) ds + Ct.$$

Therefore the functions

$$\alpha_u(t) := \mathbb{E}|\bar{X}_u(t)|^4, \quad \alpha(t) := \int_I \mathbb{E}|\bar{X}_v(t)|^4\, dv = \int_I \alpha_v(t)\, dv$$

are differentiable, and

$$\mathbb{E}|\bar{X}_u(t)|^4 - \mathbb{E}|\bar{X}_u(r)|^4 \tag{7.1}$$

$$= \mathbb{E} \int_r^t 4|\bar{X}_u(s)|^2 \bar{X}_u(s) \cdot \left( f(\bar{X}_u(s)) + \int_I \int_{\mathbb{R}^d} b(\bar{X}_u(s), x) G(u, v) \, \bar{\mu}_{v,s}(dx) \, dv \right) ds + C(t - r)$$

for all $t > r \geq 0$. From (2.1) we have

$$x \cdot (f(x) - f(0)) \leq -c_0 |x|^2$$

and hence for each $s \geq 0$,

$$\mathbb{E}\left[ |\bar{X}_u(s)|^2 \bar{X}_u(s) \cdot f(\bar{X}_u(s)) \right] \leq \mathbb{E}\left[ -c_0 |\bar{X}_u(s)|^4 + |f(0)||\bar{X}_u(s)|^3 \right]$$

$$\leq -c_0 \alpha_u(s) + \frac{c_0 - 2K_b}{4} \alpha_u(s) + C, \tag{7.2}$$

where the last line uses Young's inequality and (2.2). For the rest of integrand in (7.1), using the Lipschitz property of $b$ we have

$$\mathbb{E}\left[ |\bar{X}_u(s)|^2 \bar{X}_u(s) \cdot \int_I \int_{\mathbb{R}^d} b(\bar{X}_u(s), x) G(u, v) \, \bar{\mu}_{v,s}(dx) \, dv \right]$$

$$\leq \int_I \int_{\mathbb{R}^d} \mathbb{E}\left[ |\bar{X}_u(s)|^3 \left( |b(0,0)| + K_b |\bar{X}_u(s)| + K_b |x| \right) \right] \bar{\mu}_{v,s}(dx) \, dv$$

$$\leq C\alpha_u^{3/4}(s) + K_b \alpha_u(s) + K_b \alpha_u^{3/4}(s) \int_I \alpha_v^{1/4}(s) \, dv$$

$$\leq \frac{c_0 - 2K_b}{4} \alpha_u(s) + C + K_b \alpha_u(s) + \frac{3}{4} K_b \alpha_u(s) + \frac{1}{4} K_b \alpha(s),$$

where the third line uses Jensen's inequality and the last line uses Young's inequality and (2.2). Combining this with (7.1) and (7.2) gives

$$\alpha_u(t) - \alpha_u(r) \leq -(2c_0 - 3K_b) \int_r^t \alpha_u(s) \, ds + K_b \int_r^t \alpha(s) \, ds + C(t - r). \tag{7.3}$$

Integrating over $u \in I$ gives

$$\alpha(t) - \alpha(r) \leq -2(c_0 - 2K_b) \int_r^t \alpha(s) \, ds + C(t - r).$$

Since the function $\alpha(t)$ is non-negative and differentiable, using Lemma 7.1 (with $a_1 = 2(c_0 - 2K_b)$, $a_2 = 0$, $a_3 = C$) we have $\alpha(t) \leq C$. Applying this to (7.3) gives

$$\alpha_u(t) - \alpha_u(r) \leq -(2c_0 - 3K_b) \int_r^t \alpha_u(s) \, ds + C(t - r).$$

Since the function $\alpha_u(t)$ is non-negative and differentiable, using Lemma 7.1 again we have $\alpha_u(t) \leq C$, uniformly in $u \in I$. This completes the proof. ∎

*Proof of Theorem* 3.1. (a) From Proposition 3.1 we see that the quantity $\kappa_1 = \sup_{u \in I} \sup_{t \geq 0} \mathbb{E}|\bar{X}_u(t)|^2$ is finite. Let

$$A := \left\{ \eta := (\eta_u : u \in I) \in [\mathcal{P}(\mathbb{R}^d)]^I : \sup_{u \in I} \int_{\mathbb{R}^d} |x|^2 \, \eta_u(dx) \leq \kappa_1 \right\}.$$

Recall $\kappa$ in (2.2) and the process $\bar{Y}_u^\eta$ in (3.1).

We claim that

$$\sup_{u \in I} W_2^2((P_t\eta)_u, (P_t\tilde{\eta})_u) \le \frac{c_0 - K_b}{\kappa} e^{-\kappa t} \sup_{u \in I} W_2^2(\eta_u, \tilde{\eta}_u) \qquad (7.4)$$

for any $\eta, \tilde{\eta} \in A$ and $t \ge 0$. To see this, by Itô's formula, we have

$$e^{\kappa t}\mathbb{E}|\bar{Y}_u^\eta(t) - \bar{Y}_u^{\tilde{\eta}}(t)|^2 - \mathbb{E}|\bar{Y}_u^\eta(0) - \bar{Y}_u^{\tilde{\eta}}(0)|^2$$

$$= \int_0^t e^{\kappa s}\mathbb{E}\left[2\left(\bar{Y}_u^\eta(s) - \bar{Y}_u^{\tilde{\eta}}(s)\right) \cdot \left(f(\bar{Y}_u^\eta(s)) - f(\bar{Y}_u^{\tilde{\eta}}(s))\right.\right.$$

$$\left.\left. + \int_I \left(\int_{\mathbb{R}^d} b(\bar{Y}_u^\eta(s), x)\,\bar{\mu}_{v,s}^\eta(dx) - \int_{\mathbb{R}^d} b(\bar{Y}_u^{\tilde{\eta}}(s), x)\,\bar{\mu}_{v,s}^{\tilde{\eta}}(dx)\right) G(u, v)\,dv\right)\right]ds$$

$$+ \int_0^t \kappa e^{\kappa s}\mathbb{E}|\bar{Y}_u^\eta(s) - \bar{Y}_u^{\tilde{\eta}}(s)|^2\,ds.$$

Using (2.1) we have

$$\mathbb{E}\left[\left(\bar{Y}_u^\eta(s) - \bar{Y}_u^{\tilde{\eta}}(s)\right) \cdot \left(f(\bar{Y}_u^\eta(s)) - f(\bar{Y}_u^{\tilde{\eta}}(s))\right)\right] \le -c_0 \mathbb{E}|\bar{Y}_u^\eta(s) - \bar{Y}_u^{\tilde{\eta}}(s)|^2.$$

By adding and subtracting terms, we have

$$\mathbb{E}\left[\left(\bar{Y}_u^\eta(s) - \bar{Y}_u^{\tilde{\eta}}(s)\right) \cdot \int_I \left(\int_{\mathbb{R}^d} b(\bar{Y}_u^\eta(s), x)\,\bar{\mu}_{v,s}^\eta(dx) - \int_{\mathbb{R}^d} b(\bar{Y}_u^{\tilde{\eta}}(s), x)\,\bar{\mu}_{v,s}^{\tilde{\eta}}(dx)\right) G(u, v)\,dv\right]$$

$$\le \mathbb{E}\left[\left(\bar{Y}_u^\eta(s) - \bar{Y}_u^{\tilde{\eta}}(s)\right) \cdot \int_I \int_{\mathbb{R}^d} \left(b(\bar{Y}_u^\eta(s), x) - b(\bar{Y}_u^{\tilde{\eta}}(s), x)\right) \bar{\mu}_{v,s}^\eta(dx)\,G(u, v)\,dv\right]$$

$$+ \mathbb{E}\left[\left(\bar{Y}_u^\eta(s) - \bar{Y}_u^{\tilde{\eta}}(s)\right) \cdot \int_I \int_{\mathbb{R}^d} b(\bar{Y}_u^{\tilde{\eta}}(s), x)\left(\bar{\mu}_{v,s}^\eta(dx) - \bar{\mu}_{v,s}^{\tilde{\eta}}(dx)\right) G(u, v)\,dv\right]$$

$$\le K_b \mathbb{E}|\bar{Y}_u^\eta(s) - \bar{Y}_u^{\tilde{\eta}}(s)|^2 + K_b \mathbb{E}|\bar{Y}_u^\eta(s) - \bar{Y}_u^{\tilde{\eta}}(s)| \int_I W_2(\bar{\mu}_{v,s}^\eta, \bar{\mu}_{v,s}^{\tilde{\eta}})\,dv$$

$$\le \frac{3K_b}{2}\mathbb{E}|\bar{Y}_u^\eta(s) - \bar{Y}_u^{\tilde{\eta}}(s)|^2 + \frac{K_b}{2}\int_I W_2^2(\bar{\mu}_{v,s}^\eta, \bar{\mu}_{v,s}^{\tilde{\eta}})\,dv,$$

where the fourth line uses the Lipschitz property of $b$ and (1.3) and the last line uses Young's inequality and Jensen's inequality. Combining above three displays gives

$$e^{\kappa t}\mathbb{E}|\bar{Y}_u^\eta(t) - \bar{Y}_u^{\tilde{\eta}}(t)|^2 - \mathbb{E}|\bar{Y}_u^\eta(0) - \bar{Y}_u^{\tilde{\eta}}(0)|^2$$

$$\le -(2c_0 - 3K_b - \kappa)\int_0^t e^{\kappa s}\mathbb{E}|\bar{Y}_u^\eta(s) - \bar{Y}_u^{\tilde{\eta}}(s)|^2\,ds + K_b\int_0^t e^{\kappa s}\int_I W_2^2(\bar{\mu}_{v,s}^\eta, \bar{\mu}_{v,s}^{\tilde{\eta}})\,dv\,ds. \quad (7.5)$$

Since the function $t \mapsto e^{\kappa t}\mathbb{E}|\bar{Y}_u^\eta(t) - \bar{Y}_u^{\tilde{\eta}}(t)|^2$ is non-negative and differentiable, using Lemma 7.1 (with $a_1 = 2c_0 - 3K_b - \kappa$, $a_2 = 0$, $a_3(s) = K_b e^{\kappa s}\int_I W_2^2(\bar{\mu}_{v,s}^\eta, \bar{\mu}_{v,s}^{\tilde{\eta}})\,dv$) we have

$$e^{\kappa t}\mathbb{E}|\bar{Y}_u^\eta(t) - \bar{Y}_u^{\tilde{\eta}}(t)|^2 \le \max\left\{\mathbb{E}|\bar{Y}_u^\eta(0) - \bar{Y}_u^{\tilde{\eta}}(0)|^2, \frac{K_b \sup_{0 \le s \le t} e^{\kappa s}\int_I W_2^2(\bar{\mu}_{v,s}^\eta, \bar{\mu}_{v,s}^{\tilde{\eta}})\,dv}{2c_0 - 3K_b - \kappa}\right\}$$

$$\le \mathbb{E}|\bar{Y}_u^\eta(0) - \bar{Y}_u^{\tilde{\eta}}(0)|^2 + \frac{K_b}{2c_0 - 3K_b - \kappa}\sup_{0 \le s \le t} e^{\kappa s}\sup_{v \in I} W_2^2(\bar{\mu}_{v,s}^\eta, \bar{\mu}_{v,s}^{\tilde{\eta}}).$$

Taking the infimum over the joint distribution of $(\bar{Y}_u^\eta(0), \bar{Y}_u^{\tilde{\eta}}(0))$ gives

$$e^{\kappa t} W_2^2((P_t\eta)_u, (P_t\tilde{\eta})_u) \leq e^{\kappa t}\mathbb{E}|\bar{Y}_u^\eta(t) - \bar{Y}_u^{\tilde{\eta}}(t)|^2$$

$$\leq W_2^2(\eta_u, \tilde{\eta}_u) + \frac{K_b}{2c_0 - 3K_b - \kappa} \sup_{0\leq s\leq t} e^{\kappa s} \sup_{v\in I} W_2^2(\bar{\mu}_{v,s}^\eta, \bar{\mu}_{v,s}^{\tilde{\eta}}).$$

Taking the supremum over $u \in I$ and the time interval $[0,t]$ gives

$$\sup_{0\leq s\leq t} e^{\kappa s} \sup_{u\in I} W_2^2((P_s\eta)_u, (P_s\tilde{\eta})_u) \leq \sup_{u\in I} W_2^2(\eta_u, \tilde{\eta}_u) + \frac{K_b}{2c_0 - 3K_b - \kappa} \sup_{0\leq s\leq t} e^{\kappa s} \sup_{u\in I} W_2^2(\bar{\mu}_{u,s}^\eta, \bar{\mu}_{u,s}^{\tilde{\eta}}).$$

Since $\kappa = c_0 - 2K_b > 0$, by rearranging terms we have

$$\sup_{0\leq s\leq t} e^{\kappa s} \sup_{u\in I} W_2^2((P_s\eta)_u, (P_s\tilde{\eta})_u) \leq \frac{c_0 - K_b}{c_0 - 2K_b} \sup_{u\in I} W_2^2(\eta_u, \tilde{\eta}_u).$$

This gives the claim (7.4).

Note that $\bar{\mu}(t) := (\bar{\mu}_{u,t} : u \in I) \in A$ and $\bar{\mu}(t) = P_t\bar{\mu}(0)$ for each $t \geq 0$ by Propositions 2.1(a) and 3.1. It then follows from (7.4) that

$$W_2^2(\bar{\mu}_{u,t+s}, \bar{\mu}_{u,t}) = W_2^2((P_t\bar{\mu}(s))_u, (P_t\bar{\mu}(0))_u)$$

$$\leq \frac{c_0 - K_b}{\kappa} e^{-\kappa t} W_2^2(\bar{\mu}_{u,s}, \bar{\mu}_{u,0})$$

$$\leq 4\kappa_1 \frac{c_0 - K_b}{\kappa} e^{-\kappa t}. \tag{7.6}$$

This means that $\bar{\mu}_{u,t}$ is a $W_2$-Cauchy family when $t \to \infty$. So there exists a probability measure $\bar{\mu}_{u,\infty} \in \mathcal{P}(\mathbb{R}^d)$ such that

$$\lim_{t\to\infty} W_2(\bar{\mu}_{u,t}, \bar{\mu}_{u,\infty}) = 0. \tag{7.7}$$

In fact, taking $s \to \infty$ in (7.6) gives $W_2(\bar{\mu}_{u,t}, \bar{\mu}_{u,\infty}) \leq \sqrt{4\kappa_1 \frac{c_0 - K_b}{\kappa}} e^{-\kappa t/2}$, uniformly in $u \in I$. This gives (3.3). Since $W_2^2(\cdot, \cdot)$ is convex, (3.4) follows from (3.3) and (3.5).

(b) Next we argue that $\bar{\mu}(\infty)$ is invariant with respect to $P_t$. Noting that $\bar{\mu}(\infty) \in A$, we can apply (7.4) and use (7.7) to get

$$\limsup_{s\to\infty} W_2((P_t\bar{\mu}(\infty))_u, \bar{\mu}_{u,t+s}) \leq \limsup_{s\to\infty} \sqrt{\frac{c_0 - K_b}{\kappa}} e^{-\kappa t/2} W_2(\bar{\mu}_{u,\infty}, \bar{\mu}_{u,s}) = 0$$

and

$$\limsup_{s\to\infty} W_2(\bar{\mu}_{u,t+s}, \bar{\mu}_{u,s}) \leq \limsup_{s\to\infty} \sqrt{\frac{c_0 - K_b}{\kappa}} e^{-\kappa s/2} W_2(\bar{\mu}_{u,t}, \bar{\mu}_{u,0}) = 0.$$

Combining these two with (7.7) gives

$$W_2((P_t\bar{\mu}(\infty))_u, \bar{\mu}_{u,\infty}) \leq \limsup_{s\to\infty} W_2((P_t\bar{\mu}(\infty))_u, \bar{\mu}_{u,t+s})$$

$$+ \limsup_{s\to\infty} W_2(\bar{\mu}_{u,t+s}, \bar{\mu}_{u,s}) + \limsup_{s\to\infty} W_2(\bar{\mu}_{u,s}, \bar{\mu}_{u,\infty})$$

$$= 0.$$

This gives part (b).

(c) Fix $u_1, u_2 \in I$. Consider the following diffusions:

$$\tilde{X}_{u_1}(t) = \tilde{X}_{u_1}(0) + \int_0^t \left( f(\tilde{X}_{u_1}(s)) + \int_I \int_{\mathbb{R}^d} b(\tilde{X}_{u_1}(s), x) G(u_1, v) \, \bar{\mu}_{v,s}(dx) \, dv \right) ds + \sigma B(t),$$

$$\tilde{X}_{u_2}(t) = \tilde{X}_{u_2}(0) + \int_0^t \left( f(\tilde{X}_{u_2}(s)) + \int_I \int_{\mathbb{R}^d} b(\tilde{X}_{u_2}(s), x) G(u_2, v) \, \bar{\mu}_{v,s}(dx) \, dv \right) ds + \sigma B(t).$$

Here $B$ is a $d$-dimensional Brownian motion independent of $\{\tilde{X}_{u_1}(0), \tilde{X}_{u_2}(0)\}$, $\mathcal{L}(\tilde{X}_{u_1}(0)) = \bar{\mu}_{u_1,0}$, $\mathcal{L}(\tilde{X}_{u_2}(0)) = \bar{\mu}_{u_2,0}$, but $\tilde{X}_{u_1}(0)$ and $\tilde{X}_{u_2}(0)$ may not be independent. From the uniqueness property in Proposition 2.1(a) we have $\mathcal{L}(\tilde{X}_{u_1}) = \bar{\mu}_{u_1}$ and $\mathcal{L}(\tilde{X}_{u_2}) = \bar{\mu}_{u_2}$. Using Itô's formula we have

$$\mathbb{E}|\tilde{X}_{u_1}(t) - \tilde{X}_{u_2}(t)|^2 - \mathbb{E}|\tilde{X}_{u_1}(r) - \tilde{X}_{u_2}(r)|^2$$

$$= \mathbb{E} \int_r^t 2(\tilde{X}_{u_1}(s) - \tilde{X}_{u_2}(s)) \cdot \Big( f(\tilde{X}_{u_1}(s)) - f(\tilde{X}_{u_2}(s))$$

$$+ \int_I \int_{\mathbb{R}^d} \Big( b(\tilde{X}_{u_1}(s), x) G(u_1, v) - b(\tilde{X}_{u_2}(s), x) G(u_2, v) \Big) \bar{\mu}_{v,s}(dx) \, dv \Big) ds.$$

For each $s \geq 0$, from (2.1) we have

$$\mathbb{E} \left[ (\tilde{X}_{u_1}(s) - \tilde{X}_{u_2}(s)) \cdot \Big( f(\tilde{X}_{u_1}(s)) - f(\tilde{X}_{u_2}(s)) \Big) \right] \leq -c_0 \mathbb{E}|\tilde{X}_{u_1}(s) - \tilde{X}_{u_2}(s)|^2.$$

By adding and subtracting terms we get

$$\mathbb{E} \left| (\tilde{X}_{u_1}(s) - \tilde{X}_{u_2}(s)) \cdot \int_I \int_{\mathbb{R}^d} \Big( b(\tilde{X}_{u_1}(s), x) G(u_1, v) - b(\tilde{X}_{u_2}(s), x) G(u_2, v) \Big) \bar{\mu}_{v,s}(dx) \, dv \right|$$

$$\leq \mathbb{E} \left| (\tilde{X}_{u_1}(s) - \tilde{X}_{u_2}(s)) \cdot \int_I \int_{\mathbb{R}^d} \Big( b(\tilde{X}_{u_1}(s), x) - b(\tilde{X}_{u_2}(s), x) \Big) G(u_1, v) \, \bar{\mu}_{v,s}(dx) \, dv \right|$$

$$+ \mathbb{E} \left| (\tilde{X}_{u_1}(s) - \tilde{X}_{u_2}(s)) \cdot \int_I \int_{\mathbb{R}^d} b(\tilde{X}_{u_2}(s), x) \left( G(u_1, v) - G(u_2, v) \right) \bar{\mu}_{v,s}(dx) \, dv \right|.$$

For the first term on the right hand side, it follows from the Lipschitz property of $b$ that

$$\mathbb{E} \left| (\tilde{X}_{u_1}(s) - \tilde{X}_{u_2}(s)) \cdot \int_I \int_{\mathbb{R}^d} \Big( b(\tilde{X}_{u_1}(s), x) - b(\tilde{X}_{u_2}(s), x) \Big) G(u_1, v) \, \bar{\mu}_{v,s}(dx) \, dv \right|$$

$$\leq K_b \mathbb{E}|\tilde{X}_{u_1}(s) - \tilde{X}_{u_2}(s)|^2.$$

For the other term, using the Cauchy-Schwarz inequality, Young's inequality, the Lipschitz property of $b$ and Proposition 3.1 we have

$$\mathbb{E} \left| (\tilde{X}_{u_1}(s) - \tilde{X}_{u_2}(s)) \cdot \int_I \int_{\mathbb{R}^d} b(\tilde{X}_{u_2}(s), x) \left( G(u_1, v) - G(u_2, v) \right) \bar{\mu}_{v,s}(dx) \, dv \right|$$

$$\leq K_b \mathbb{E}|\tilde{X}_{u_1}(s) - \tilde{X}_{u_2}(s)|^2 + \frac{1}{4K_b} \mathbb{E} \left( \int_I \int_{\mathbb{R}^d} |b(\tilde{X}_{u_2}(s), x)| |G(u_1, v) - G(u_2, v)| \, \bar{\mu}_{v,s}(dx) \, dv \right)^2$$

$$\leq K_b \mathbb{E}|\tilde{X}_{u_1}(s) - \tilde{X}_{u_2}(s)|^2 + C \mathbb{E} \left( \int_I \left( 1 + |\tilde{X}_{u_2}(s)| \right) |G(u_1, v) - G(u_2, v)| \, dv \right)^2$$

$$\leq K_b \mathbb{E}|\tilde{X}_{u_1}(s) - \tilde{X}_{u_2}(s)|^2 + C \left( \int_I |G(u_1, v) - G(u_2, v)| \, dv \right)^2.$$

Combining above five displays gives

$$\mathbb{E}|\tilde{X}_{u_1}(t) - \tilde{X}_{u_2}(t)|^2 - \mathbb{E}|\tilde{X}_{u_1}(r) - \tilde{X}_{u_2}(r)|^2$$

$$\leq -2(c_0 - 2K_b) \int_r^t \mathbb{E}|\tilde{X}_{u_1}(s) - \tilde{X}_{u_2}(s)|^2 \, ds + C(t - r) \left( \int_I |G(u_1, v) - G(u_2, v)| \, dv \right)^2.$$

Since the function $t \mapsto \mathbb{E}|\tilde{X}_{u_1}(t) - \tilde{X}_{u_2}(t)|^2$ is non-negative and differentiable, using Lemma 7.1 (with $a_1 = 2(c_0 - 2K_b)$, $a_2 = 0$, $a_3 = C \left( \int_I |G(u_1, v) - G(u_2, v)| \, dv \right)^2$) we have

$$\mathbb{E}|\tilde{X}_{u_1}(t) - \tilde{X}_{u_2}(t)|^2 \leq \max \left\{ \mathbb{E}|\tilde{X}_{u_1}(0) - \tilde{X}_{u_2}(0)|^2, C \left( \int_I |G(u_1, v) - G(u_2, v)| \, dv \right)^2 \right\}.$$

Taking the infimum over the joint distribution of $\tilde{X}_{u_1}(0)$ and $\tilde{X}_{u_2}(0)$ gives part (c) and completes the proof. ∎

## 7.2. Proofs for Section 4.

*Proof of Proposition* 4.1. Fix $n \in \mathbb{N}$ and $z = (z_{ij} = z_{ji})_{i,j=1}^n \in [0, 1]^{n^2}$. Using the Lipschitz property of $f, b$ and a standard argument one has

$$\max_{i=1,\ldots,n} \sup_{0 \leq t \leq T} \mathbb{E}^{n,z}|X_i^n(t)|^2 < \infty \quad a.s., \quad \forall T \in (0, \infty).$$

Using this and Itô's formula, we have

$$\mathbb{E}^{n,z}|X_i^n(t)|^2 - \mathbb{E}^{n,z}|X_i^n(r)|^2$$

$$= \mathbb{E}^{n,z} \int_r^t 2X_i^n(s) \cdot \left( f(X_i^n(s)) + \frac{1}{n} \sum_{j=1}^n z_{ij} b(X_i^n(s), X_j^n(s)) \right) ds + C(t - r),$$

and hence the functions

$$\alpha_i^{n,z}(t) := \mathbb{E}^{n,z}|X_i^n(t)|^2, \quad \alpha^{n,z}(t) := \frac{1}{n} \sum_{i=1}^n \mathbb{E}^{n,z}|X_i^n(t)|^2$$

are differentiable. For each $s \geq 0$, using (2.1) and the Lipschitz property of $b$ we have

$$\mathbb{E}^{n,z} \left[ X_i^n(s) \cdot \left( f(X_i^n(s)) + \frac{1}{n} \sum_{j=1}^n z_{ij} b(X_i^n(s), X_j^n(s)) \right) \right]$$

$$\leq \mathbb{E}^{n,z} \left[ -c_0 |X_i^n(s)|^2 + C|X_i^n(s)| + \frac{1}{n} \sum_{j=1}^n |X_i^n(s)| \left( C + K_b |X_i^n(s)| + K_b |X_j^n(s)| \right) \right]$$

$$\leq -c_0 \alpha_i^{n,z}(s) + \frac{c_0 - 2K_b}{2} \alpha_i^{n,z}(s) + C + K_b \alpha_i^{n,z}(s) + \frac{K_b}{2} \alpha_i^{n,z}(s) + \frac{K_b}{2} \alpha^{n,z}(s),$$

where the last line uses Young's inequality and (2.2). Therefore

$$\alpha_i^{n,z}(t) - \alpha_i^{n,z}(r) \leq - (c_0 - K_b) \int_r^t \alpha_i^{n,z}(s) \, ds + K_b \int_r^t \alpha^{n,z}(s) \, ds + C(t - r). \tag{7.8}$$

Taking the average over $i = 1, \ldots, n$ gives

$$\alpha^{n,z}(t) - \alpha^{n,z}(r) \leq -(c_0 - 2K_b) \int_r^t \alpha^{n,z}(s) \, ds + C(t - r).$$

Since the function $\alpha^{n,z}(t)$ is non-negative and differentiable, using Lemma 7.1 (with $a_1 = c_0 - 2K_b$, $a_2 = 0$, $a_3 = C$) we have $\alpha^{n,z}(t) \leq C$. From this and (7.8) we further have

$$\alpha_i^{n,z}(t) - \alpha_i^{n,z}(r) \leq -(c_0 - K_b) \int_r^t \alpha_i^{n,z}(s)\, ds + C(t - r).$$

Since the function $\alpha_i^{n,z}(t)$ is non-negative and differentiable, using Lemma 7.1 again we have $\alpha_i^{n,z}(t) \leq C$, uniformly in $n \in \mathbb{N}$, $i = 1, \ldots, n$ and $t \geq 0$. This completes the proof. ∎

*Proof of Theorem* 4.1. (a) Recall Proposition 4.1 and $\kappa_2$ therein. Fix $n \in \mathbb{N}$ and $z = (z_{ij} = z_{ji})_{i,j=1}^n \in [0,1]^{n^2}$. Let

$$A_n := \left\{ \eta \in \mathcal{P}((\mathbb{R}^d)^n) : \max_{i=1,\ldots,n} \int_{(\mathbb{R}^d)^n} |x_i|^2\, \eta(dx) \leq \kappa_2 \right\}.$$

Recall the process $Y^{n,z,\eta}$ in (4.1).

We claim that

$$W_2^2(P_t^{n,z}\eta, P_t^{n,z}\tilde{\eta}) \leq W_2^2(\eta, \tilde{\eta}) e^{-2\kappa t} \tag{7.9}$$

for any $\eta, \tilde{\eta} \in A_n$ and $t \geq 0$. To see this, by Itô's formula, we have

$$e^{2\kappa t} \sum_{i=1}^n \mathbb{E}^{n,z} |Y_i^{n,z,\eta}(t) - Y_i^{n,z,\tilde{\eta}}(t)|^2 - \sum_{i=1}^n \mathbb{E}^{n,z} |Y_i^{n,z,\eta}(0) - Y_i^{n,z,\tilde{\eta}}(0)|^2$$

$$= \int_0^t e^{2\kappa s} \sum_{i=1}^n \mathbb{E}^{n,z} \left[ 2 \left( Y_i^{n,z,\eta}(s) - Y_i^{n,z,\tilde{\eta}}(s) \right) \cdot \left( f(Y_i^{n,z,\eta}(s)) - f(Y_i^{n,z,\tilde{\eta}}(s)) \right. \right.$$

$$\left. \left. + \frac{1}{n} \sum_{j=1}^n z_{ij} \left( b(Y_i^{n,z,\eta}(s), Y_j^{n,z,\eta}(s)) - b(Y_i^{n,z,\tilde{\eta}}(s), Y_j^{n,z,\tilde{\eta}}(s)) \right) \right) \right] ds$$

$$+ \int_0^t 2\kappa e^{2\kappa s} \sum_{i=1}^n \mathbb{E}^{n,z} |Y_i^{n,z,\eta}(s) - Y_i^{n,z,\tilde{\eta}}(s)|^2\, ds.$$

Using (2.1) we have

$$\sum_{i=1}^n \mathbb{E}^{n,z} \left[ \left( Y_i^{n,z,\eta}(s) - Y_i^{n,z,\tilde{\eta}}(s) \right) \cdot \left( f(Y_i^{n,z,\eta}(s)) - f(Y_i^{n,z,\tilde{\eta}}(s)) \right) \right]$$

$$\leq -c_0 \sum_{i=1}^n \mathbb{E}^{n,z} |Y_i^{n,z,\eta}(s) - Y_i^{n,z,\tilde{\eta}}(s)|^2.$$

Using the Cauchy-Schwarz inequality, the Lipschitz property of $b$ and Young's inequality we have

$$\sum_{i=1}^n \mathbb{E}^{n,z} \left[ \left( Y_i^{n,z,\eta}(s) - Y_i^{n,z,\tilde{\eta}}(s) \right) \cdot \frac{1}{n} \sum_{j=1}^n z_{ij} \left( b(Y_i^{n,z,\eta}(s), Y_j^{n,z,\eta}(s)) - b(Y_i^{n,z,\tilde{\eta}}(s), Y_j^{n,z,\tilde{\eta}}(s)) \right) \right]$$

$$\leq \sum_{i=1}^n \mathbb{E}^{n,z} \left[ \left| Y_i^{n,z,\eta}(s) - Y_i^{n,z,\tilde{\eta}}(s) \right| \cdot \frac{K_b}{n} \sum_{j=1}^n \left( \left| Y_i^{n,z,\eta}(s) - Y_i^{n,z,\tilde{\eta}}(s) \right| + \left| Y_j^{n,z,\eta}(s) - Y_j^{n,z,\tilde{\eta}}(s) \right| \right) \right]$$

$$\leq 2K_b \sum_{i=1}^n \mathbb{E}^{n,z} |Y_i^{n,z,\eta}(s) - Y_i^{n,z,\tilde{\eta}}(s)|^2.$$

Combining above three estimates with the definition of $\kappa$ in (2.2) gives

$$e^{2\kappa t} \sum_{i=1}^{n} \mathbb{E}^{n,z} |Y_i^{n,z,\eta}(t) - Y_i^{n,z,\tilde{\eta}}(t)|^2 \leq \sum_{i=1}^{n} \mathbb{E}^{n,z} |Y_i^{n,z,\eta}(0) - Y_i^{n,z,\tilde{\eta}}(0)|^2.$$

Therefore

$$W_2^2(P_t^{n,z}\eta, P_t^{n,z}\tilde{\eta}) \leq \sum_{i=1}^{n} \mathbb{E}^{n,z} |Y_i^{n,z,\eta}(t) - Y_i^{n,z,\tilde{\eta}}(t)|^2 \leq e^{-2\kappa t} \sum_{i=1}^{n} \mathbb{E}^{n,z} |Y_i^{n,z,\eta}(0) - Y_i^{n,z,\tilde{\eta}}(0)|^2.$$

Taking the infimum over the joint distribution of $(Y^{n,z,\eta}(0), Y^{n,z,\tilde{\eta}}(0))$ gives the claim (7.9).

Note that $\theta^{n,z}(t) = P_t \theta^{n,z}(0) \in A_n$ for each $t \geq 0$ by Proposition 4.1. Therefore

$$W_2^2(\theta^{n,z}(t), \theta^{n,z}(0)) \leq \sum_{i=1}^{n} \mathbb{E}^{n,z} |X_i^n(t) - X_i^n(0)|^2 \leq 2 \sum_{i=1}^{n} \mathbb{E}^{n,z} [|X_i^n(t)|^2 + |X_i^n(0)|^2] \leq 4n\kappa_2.$$

It then follows from (7.9) that

$$\begin{aligned} W_2^2(\theta^{n,z}(t+s), \theta^{n,z}(t)) &= W_2^2(P_t^{n,z}\theta^{n,z}(s), P_t^{n,z}\theta^{n,z}(0)) \\ &\leq W_2^2(\theta^{n,z}(s), \theta^{n,z}(0))e^{-2\kappa t} \\ &\leq 4n\kappa_2 e^{-2\kappa t}. \end{aligned} \tag{7.10}$$

This means that $\theta^{n,z}(t)$ is a $W_2$-Cauchy family when $t \to \infty$. So there exists a probability measure $\theta^{n,z}(\infty) \in \mathcal{P}((\mathbb{R}^d)^n)$ such that

$$\lim_{t\to\infty} W_2(\theta^{n,z}(t), \theta^{n,z}(\infty)) = 0. \tag{7.11}$$

In fact, taking $s \to \infty$ in (7.10) gives

$$\frac{1}{\sqrt{n}} W_2(\theta^{n,z}(t), \theta^{n,z}(\infty)) \leq \sqrt{4\kappa_2} e^{-\kappa t},$$

which gives (4.3). From this and (4.5) we have

$$\frac{1}{\sqrt{n}} W_2(\theta^n(t), \theta^n(\infty)) \leq \sup_{z\in[0,1]^{n^2}} \frac{1}{n} W_2(\theta^{n,z}(t), \theta^{n,z}(\infty)) \leq \sqrt{4\kappa_2} e^{-\kappa t}.$$

Therefore (4.4) hold.

(b) Finally we argue that $\theta^{n,z}(\infty)$ is invariant with respect to $P_t^{n,z}$. Noting that $\theta^{n,z}(\infty) \in A_n$, we can apply (7.9) and use (7.11) to get

$$\limsup_{s\to\infty} W_2(P_t^{n,z}\theta^{n,z}(\infty), \theta^{n,z}(t+s)) \leq \limsup_{s\to\infty} e^{-\kappa t} W_2(\theta^{n,z}(\infty), \theta^{n,z}(s)) = 0$$

and

$$\limsup_{s\to\infty} W_2(\theta^{n,z}(t+s), \theta^{n,z}(s)) \leq \limsup_{s\to\infty} e^{-\kappa s} W_2(\theta^{n,z}(t), \theta^{n,z}(0)) = 0.$$

Combining these two with (7.11) gives

$$\begin{aligned} W_2(P_t^{n,z}\theta^{n,z}(\infty), \theta^{n,z}(\infty)) &\leq \limsup_{s\to\infty} W_2(P_t^{n,z}\theta^{n,z}(\infty), \theta^{n,z}(t+s)) \\ &\quad + \limsup_{s\to\infty} W_2(\theta^{n,z}(t+s), \theta^{n,z}(s)) + \limsup_{s\to\infty} W_2(\theta^{n,z}(s), \theta^{n,z}(\infty)) \\ &= 0. \end{aligned}$$

This gives part (b) and completes the proof. ∎

7.3. **Proofs for Section 5.** We need the following lemma to prove Theorem 5.1.

**Lemma 7.2.** *Suppose Conditions* 3.1 *and* 5.1 *hold. For* $s \geq 0$, *write*

$$\mathcal{R}_s^n := \frac{1}{n} \sum_{i=1}^{n} \mathbb{E} \left| \frac{1}{n} \sum_{j=1}^{n} \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{j}{n}) \, \bar{\mu}_{\frac{j}{n}, s}(dx) \right.$$

$$\left. - \int_I \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v) \, \bar{\mu}_{v,s}(dx) \, dv \right|^2. \tag{7.12}$$

*Then*

$$\lim_{n \to \infty} \sup_{s \geq 0} \mathcal{R}_s^n = 0.$$

*Proof.* Fix $M > 1$ and write

$$b_M(x, y) := b(x, y) \mathbf{1}_{\{|x| \leq M, |y| \leq M\}}. \tag{7.13}$$

It then follows from [35, Corollary 2 of Theorem 3.1] that there exist some $m \in \mathbb{N}$ and polynomials

$$\tilde{b}_m(x, y) := \sum_{k=1}^{m} a_k(x) c_k(y) \mathbf{1}_{\{|x| \leq M, |y| \leq M\}}, \tag{7.14}$$

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

$$|b_M(x, y) - \tilde{b}_m(x, y)| \leq 1/M. \tag{7.15}$$

By adding and subtracting terms, we have

$$\mathcal{R}_s^n \leq \frac{5}{n} \sum_{i=1}^{n} \mathbb{E} \left| \frac{1}{n} \sum_{j=1}^{n} \int_{\mathbb{R}^d} \left( b(\bar{X}_{\frac{i}{n}}(s), x) - b_M(\bar{X}_{\frac{i}{n}}(s), x) \right) G_n(\frac{i}{n}, \frac{j}{n}) \, \bar{\mu}_{\frac{j}{n}, s}(dx) \right|^2$$

$$+ \frac{5}{n} \sum_{i=1}^{n} \mathbb{E} \left| \int_I \int_{\mathbb{R}^d} \left( b(\bar{X}_{\frac{i}{n}}(s), x) - b_M(\bar{X}_{\frac{i}{n}}(s), x) \right) G(\frac{i}{n}, v) \, \bar{\mu}_{v,s}(dx) \, dv \right|^2$$

$$+ \frac{5}{n} \sum_{i=1}^{n} \mathbb{E} \left| \frac{1}{n} \sum_{j=1}^{n} \int_{\mathbb{R}^d} \left( b_M(\bar{X}_{\frac{i}{n}}(s), x) - \tilde{b}_m(\bar{X}_{\frac{i}{n}}(s), x) \right) G_n(\frac{i}{n}, \frac{j}{n}) \, \bar{\mu}_{\frac{j}{n}, s}(dx) \right|^2$$

$$+ \frac{5}{n} \sum_{i=1}^{n} \mathbb{E} \left| \int_I \int_{\mathbb{R}^d} \left( b_M(\bar{X}_{\frac{i}{n}}(s), x) - \tilde{b}_m(\bar{X}_{\frac{i}{n}}(s), x) \right) G(\frac{i}{n}, v) \, \bar{\mu}_{v,s}(dx) \, dv \right|^2$$

$$+ \frac{5}{n} \sum_{i=1}^{n} \mathbb{E} \left| \frac{1}{n} \sum_{j=1}^{n} \int_{\mathbb{R}^d} \tilde{b}_m(\bar{X}_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{j}{n}) \, \bar{\mu}_{\frac{j}{n}, s}(dx) \right.$$

$$\left. - \int_I \int_{\mathbb{R}^d} \tilde{b}_m(\bar{X}_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v) \, \bar{\mu}_{v,s}(dx) \, dv \right|^2$$

$$=: 5 \sum_{k=1}^{5} \mathcal{R}_s^{n,k}. \tag{7.16}$$

Next we analyze each term. For $\mathcal{R}_s^{n,1}$ and $\mathcal{R}_s^{n,2}$, using (7.13), Remark 2.2(b), Proposition 3.1 and the Cauchy-Schwarz inequality we have

$$
\begin{aligned}
\mathcal{R}_s^{n,1} &\leq \frac{C}{n} \sum_{i=1}^{n} \mathbb{E}\left[ \frac{1}{n} \sum_{j=1}^{n} \int_{\mathbb{R}^d} \left(1 + |\bar{X}_{\frac{i}{n}}(s)| + |x|\right) \left(\mathbf{1}_{\{|\bar{X}_{\frac{i}{n}}(s)|>M\}} + \mathbf{1}_{\{|x|>M\}}\right) \bar{\mu}_{\frac{j}{n},s}(dx) \right]^2 \\
&\leq \frac{C}{n} \sum_{i=1}^{n} \mathbb{E}\left[ \left(1 + |\bar{X}_{\frac{i}{n}}(s)|\right)^2 \mathbf{1}_{\{|\bar{X}_{\frac{i}{n}}(s)|>M\}} \right] \\
&\leq \frac{C}{\sqrt{M}},
\end{aligned}
\tag{7.17}
$$

and

$$
\begin{aligned}
\mathcal{R}_s^{n,2} &\leq \frac{C}{n} \sum_{i=1}^{n} \mathbb{E}\left[ \int_I \int_{\mathbb{R}^d} \left(1 + |\bar{X}_{\frac{i}{n}}(s)| + |x|\right) \left(\mathbf{1}_{\{|\bar{X}_{\frac{i}{n}}(s)|>M\}} + \mathbf{1}_{\{|x|>M\}}\right) \bar{\mu}_{v,s}(dx)\, dv \right]^2 \\
&\leq \frac{C}{n} \sum_{i=1}^{n} \mathbb{E}\left[ \left(1 + |\bar{X}_{\frac{i}{n}}(s)|\right)^2 \mathbf{1}_{\{|\bar{X}_{\frac{i}{n}}(s)|>M\}} \right] + C \int_I \mathbb{E}\left[ \left(1 + |\bar{X}_v(s)|\right)^2 \mathbf{1}_{\{|\bar{X}_v(s)|>M\}} \right] dv \\
&\leq \frac{C}{\sqrt{M}}.
\end{aligned}
\tag{7.18}
$$

For $\mathcal{R}_s^{n,3}$ and $\mathcal{R}_s^{n,4}$, using (7.15) we have

$$
\mathcal{R}_s^{n,3} \leq \frac{C}{M^2}, \quad \mathcal{R}_s^{n,4} \leq \frac{C}{M^2}.
\tag{7.19}
$$

For $\mathcal{R}_s^{n,5}$, using the step graphon structure (5.1) of $G_n$ and by adding and subtracting terms, we have

$$
\begin{aligned}
\mathcal{R}_s^{n,5} &= \int_I \mathbb{E}\left| \int_I \int_{\mathbb{R}^d} \tilde{b}_m(\bar{X}_{\frac{[nu]}{n}}(s),x) G_n(u,v)\, \bar{\mu}_{\frac{[nv]}{n},s}(dx)\, dv \right. \\
&\qquad\qquad \left. - \int_I \int_{\mathbb{R}^d} \tilde{b}_m(\bar{X}_{\frac{[nu]}{n}}(s),x) G(\frac{[nu]}{n},v)\, \bar{\mu}_{v,s}(dx)\, dv \right|^2 du \\
&\leq 3 \int_I \mathbb{E}\left| \int_I \int_{\mathbb{R}^d} \tilde{b}_m(\bar{X}_{\frac{[nu]}{n}}(s),x) \left(G_n(u,v) - G(u,v)\right) \bar{\mu}_{\frac{[nv]}{n},s}(dx)\, dv \right|^2 du \\
&\quad + 3 \int_I \mathbb{E}\left| \int_I \int_{\mathbb{R}^d} \tilde{b}_m(\bar{X}_{\frac{[nu]}{n}}(s),x) \left(G(u,v) - G(\frac{[nu]}{n},v)\right) \bar{\mu}_{\frac{[nv]}{n},s}(dx)\, dv \right|^2 du \\
&\quad + 3 \int_I \mathbb{E}\left| \int_I \int_{\mathbb{R}^d} \tilde{b}_m(\bar{X}_{\frac{[nu]}{n}}(s),x) G(\frac{[nu]}{n},v) \left(\bar{\mu}_{\frac{[nv]}{n},s}(dx) - \bar{\mu}_{v,s}(dx)\right) dv \right|^2 du \\
&=: \mathcal{R}_s^{n,6} + \mathcal{R}_s^{n,7} + \mathcal{R}_s^{n,8}.
\end{aligned}
\tag{7.20}
$$

For $\mathcal{R}_s^{n,6}$, using the definition of $\tilde{b}_m$ in (7.14), Proposition 3.1 and Remark 2.1, we have

$$
\begin{aligned}
\mathcal{R}_s^{n,6} &\leq 3m \sum_{k=1}^{m} \int_I \mathbb{E}\left[ a_k^2(\bar{X}_{\frac{[nu]}{n}}(s)) \right] \\
&\qquad \cdot \left| \int_I (G_n(u,v) - G(u,v)) \left( \int_{\mathbb{R}^d} c_k(x) \mathbf{1}_{\{|x|\leq M\}} \bar{\mu}_{\frac{[nv]}{n},s}(dx) \right) dv \right|^2 du \\
&\leq C_M \|G_n - G\|,
\end{aligned}
$$

where $C_M$ depends on $M$ but not on $n$ or $s$. For $\mathcal{R}_s^{n,7}$, we have

$$\mathcal{R}_s^{n,7} \leq C_M \int_{I \times I} \left| G(u,v) - G(\frac{\lceil nu \rceil}{n}, v) \right| du \, dv.$$

For $\mathcal{R}_s^{n,8}$, using (1.3) and the Lipschitz property of $b$ (and hence $\tilde{b}_m$), we have

$$\mathcal{R}_s^{n,8} \leq C_M \int_I W_2^2(\bar{\mu}_{\frac{\lceil nv \rceil}{n},s}, \bar{\mu}_{v,s}) \, dv.$$

Combining above three estimates with (7.20) and using Remark 2.1, Condition 5.1, Condition 3.1 and Corollary 3.1(a) gives

$$\lim_{n \to \infty} \sup_{s \geq 0} \mathcal{R}_s^{n,5} = 0.$$

Combining this with (7.16)–(7.19) gives

$$\limsup_{n \to \infty} \sup_{s \geq 0} \mathcal{R}_s^n \leq \frac{C}{\sqrt{M}}.$$

Taking $\limsup_{M \to \infty}$ completes the proof. $\blacksquare$

*Proof of Theorem* 5.1. (a) Using Itô's formula, we have

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E}|X_i^n(t) - \bar{X}_{\frac{i}{n}}(t)|^2$$

$$= \frac{1}{n} \sum_{i=1}^n \mathbb{E} \int_0^t 2(X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)) \cdot \Big( f(X_i^n(s)) - f(\bar{X}_{\frac{i}{n}}(s))$$

$$+ \frac{1}{n} \sum_{j=1}^n \xi_{ij}^n b(X_i^n(s), X_j^n(s)) - \int_I \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v) \, \bar{\mu}_{v,s}(dx) \, dv \Big) \, ds.$$

This implies that the function

$$\alpha^n(t) := \frac{1}{n} \sum_{i=1}^n \mathbb{E}|X_i^n(t) - \bar{X}_{\frac{i}{n}}(t)|^2 \tag{7.21}$$

is differentiable, and

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E}|X_i^n(t) - \bar{X}_{\frac{i}{n}}(t)|^2 - \frac{1}{n} \sum_{i=1}^n \mathbb{E}|X_i^n(r) - \bar{X}_{\frac{i}{n}}(r)|^2$$

$$= \frac{1}{n} \sum_{i=1}^n \mathbb{E} \int_r^t 2(X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)) \cdot \Big( f(X_i^n(s)) - f(\bar{X}_{\frac{i}{n}}(s))$$

$$+ \frac{1}{n} \sum_{j=1}^n \xi_{ij}^n b(X_i^n(s), X_j^n(s)) - \int_I \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v) \, \bar{\mu}_{v,s}(dx) \, dv \Big) \, ds. \tag{7.22}$$

For each $s \geq 0$, using (2.1) we have

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E}\left[ (X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)) \cdot \Big( f(X_i^n(s)) - f(\bar{X}_{\frac{i}{n}}(s)) \Big) \right] \leq -c_0 \frac{1}{n} \sum_{i=1}^n \mathbb{E}\left[ |X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)|^2 \right]. \tag{7.23}$$

For the rest in the integrand of (7.22), by adding and subtracting terms, we have

$$
\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\left[ (X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)) \right.
$$

$$
\left. \cdot \left( \frac{1}{n} \sum_{j=1}^{n} \xi_{ij}^n b(X_i^n(s), X_j^n(s)) - \int_I \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v)\, \bar{\mu}_{v,s}(dx)\, dv \right) \right]
$$

$$
= \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\left[ (X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)) \cdot \left( \frac{1}{n} \sum_{j=1}^{n} \xi_{ij}^n \left( b(X_i^n(s), X_j^n(s)) - b(\bar{X}_{\frac{i}{n}}(s), \bar{X}_{\frac{j}{n}}(s)) \right) \right) \right]
$$

$$
+ \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\left[ (X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)) \right.
$$

$$
\left. \cdot \left( \frac{1}{n} \sum_{j=1}^{n} \left( \xi_{ij}^n b(\bar{X}_{\frac{i}{n}}(s), \bar{X}_{\frac{j}{n}}(s)) - \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{j}{n})\, \bar{\mu}_{\frac{j}{n},s}(dx) \right) \right) \right]
$$

$$
+ \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\left[ (X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)) \right.
$$

$$
\left. \cdot \left( \frac{1}{n} \sum_{j=1}^{n} \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{j}{n})\, \bar{\mu}_{\frac{j}{n},s}(dx) - \int_I \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v)\, \bar{\mu}_{v,s}(dx)\, dv \right) \right]
$$

$$
=: \mathcal{S}_s^{n,1} + \mathcal{S}_s^{n,2} + \mathcal{S}_s^{n,3}. \tag{7.24}
$$

For $\mathcal{S}_s^{n,1}$, using the Lipschitz property of $b$ and Young's inequality, we have

$$
\mathcal{S}_s^{n,1} \leq 2 K_b \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}|X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)|^2. \tag{7.25}
$$

For $\mathcal{S}_s^{n,2}$, using the Cauchy-Schwarz inequality we have

$$
\mathcal{S}_s^{n,2} \leq \left( \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}|X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)|^2 \right)^{1/2}
$$

$$
\cdot \left( \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\left| \frac{1}{n} \sum_{j=1}^{n} \left( \xi_{ij}^n b(\bar{X}_{\frac{i}{n}}(s), \bar{X}_{\frac{j}{n}}(s)) - \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{j}{n})\, \bar{\mu}_{\frac{j}{n},s}(dx) \right) \right|^2 \right)^{1/2}.
$$

Due to the independence of $\xi_{ij}^n$ and $\bar{X}_u$, we have

$$
\mathbb{E}\left| \frac{1}{n} \sum_{j=1}^{n} \left( \xi_{ij}^n b(\bar{X}_{\frac{i}{n}}(s), \bar{X}_{\frac{j}{n}}(s)) - \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{j}{n})\, \bar{\mu}_{\frac{j}{n},s}(dx) \right) \right|^2
$$

$$
= \frac{1}{n^2} \sum_{j=1}^{n} \mathbb{E}\left| \xi_{ij}^n b(\bar{X}_{\frac{i}{n}}(s), \bar{X}_{\frac{j}{n}}(s)) - \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{j}{n})\, \bar{\mu}_{\frac{j}{n},s}(dx) \right|^2
$$

$$
\leq \frac{C}{n}, \tag{7.26}
$$

where the last line uses Remark 2.2(b) and Proposition 3.1. Therefore

$$\mathcal{S}_s^{n,2} \leq \frac{C}{\sqrt{n}} \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}|X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)|^2 \right)^{1/2}. \tag{7.27}$$

For $\mathcal{S}_s^{n,3}$, using the Cauchy-Schwarz inequality and the definition of $\mathcal{R}_s^n$ in (7.12), we have

$$\mathcal{S}_s^{n,3} \leq (\mathcal{R}_s^n)^{1/2} \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}|X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)|^2 \right)^{1/2}. \tag{7.28}$$

Combining (7.21)–(7.25), (7.27) and (7.28), we have

$$\alpha^n(t) - \alpha^n(r) \leq -2(c_0 - 2K_b) \int_r^t \alpha^n(s)\, ds + \left( \frac{C}{\sqrt{n}} + 2 \sup_{s \geq 0} \sqrt{\mathcal{R}_s^n} \right) \int_r^t \sqrt{\alpha^n(s)}\, ds.$$

Recall that the function $\alpha^n(t)$ is differentiable, non-negative, and $\alpha^n(0) = 0$. It then follows from Lemma 7.1 (with $a_1 = 2(c_0 - 2K_b)$, $a_2 = \frac{C}{\sqrt{n}} + 2 \sup_{s \geq 0} \sqrt{\mathcal{R}_s^n}$, $a_3 = 0$) that

$$\alpha^n(t) \leq C \left( \frac{1}{n} + \sup_{s \geq 0} \mathcal{R}_s^n \right).$$

Combining this with Lemma 7.2 gives part (a).

(b) Next we prove the first convergence statement in (5.2). Write

$$\bar{\mu}^n(t) := \frac{1}{n} \sum_{i=1}^n \bar{\mu}_{\frac{i}{n},t}. \tag{7.29}$$

Using the triangle inequality we have

$$W_2(\mu^n(t), \hat{\mu}(t)) \leq W_2(\mu^n(t), \bar{\mu}^n(t)) + W_2(\bar{\mu}^n(t), \hat{\mu}(t)). \tag{7.30}$$

Taking $\pi = \frac{1}{n} \sum_{i=1}^n \mathcal{L}(X_i^n(t), \bar{X}_{\frac{i}{n}}(t))$ as the coupling of $\mu^n(t)$ and $\bar{\mu}^n(t)$ and using part (a), we have

$$\sup_{t \geq 0} W_2(\mu^n(t), \bar{\mu}^n(t)) \leq \sup_{t \geq 0} \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}|X_i^n(t) - \bar{X}_{\frac{i}{n}}(t)|^2 \right)^{1/2} \to 0$$

as $n \to \infty$. Using the convexity of $W_2^2(\cdot, \cdot)$ and Corollary 3.1(a) we have

$$\sup_{t \geq 0} W_2^2(\bar{\mu}^n(t), \hat{\mu}(t)) \leq \int_0^1 \sup_{t \geq 0} W_2^2(\bar{\mu}_{\frac{\lceil nu \rceil}{n},t}, \bar{\mu}_{u,t})\, du \to 0 \tag{7.31}$$

as $n \to \infty$. Combining these three displays gives the first convergence in part (b).

Finally, for the second convergence statement in part (b), let

$$\bar{\nu}^n(t) := \frac{1}{n} \sum_{i=1}^n \delta_{\bar{X}_{\frac{i}{n}}(t)}. \tag{7.32}$$

From the triangle inequality we have

$$\mathbb{E}W_2(\nu^n(t), \hat{\mu}(t)) \leq \mathbb{E}W_2(\nu^n(t), \bar{\nu}^n(t)) + \mathbb{E}W_2(\bar{\nu}^n(t), \bar{\mu}^n(t)) + W_2(\bar{\mu}^n(t), \hat{\mu}(t)). \tag{7.33}$$

In view of (7.31), it suffices to show

$$\sup_{t \geq 0} \mathbb{E}W_2(\nu^n(t), \bar{\nu}^n(t)) + \sup_{t \geq 0} \mathbb{E}W_2(\bar{\nu}^n(t), \bar{\mu}^n(t)) \to 0 \tag{7.34}$$

as $n \to \infty$. Taking $\pi = \frac{1}{n} \sum_{i=1}^{n} \delta_{(X_i^n(t), \bar{X}_{\frac{i}{n}}(t))}$ as the coupling of $\nu^n(t)$ and $\bar{\nu}^n(t)$ and using part (a), we have

$$\sup_{t \geq 0} \mathbb{E} W_2(\nu^n(t), \bar{\nu}^n(t)) \leq \sup_{t \geq 0} \left( \frac{1}{n} \sum_{i=1}^{n} \mathbb{E} |X_i^n(t) - \bar{X}_{\frac{i}{n}}(t)|^2 \right)^{1/2} \to 0$$

as $n \to \infty$. Applying Lemma A.1 with $Y_i = \bar{X}_{\frac{i}{n}}$, $p = 3$ and $q = 4$, we have

$$\mathbb{E} W_2(\bar{\nu}^n(t), \bar{\mu}^n(t)) \leq \left( \mathbb{E} W_p^p(\bar{\nu}^n(t), \bar{\mu}^n(t)) \right)^{1/p} \leq C \left( \int_{\mathbb{R}^d} |x|^q \, \bar{\mu}^n(t)(dx) \right)^{1/q} a(n),$$

where $a(n) = n^{-1/d} + n^{-1/12}$ is defined in (5.3). It then follows from Proposition 3.1 that

$$\sup_{t \geq 0} \mathbb{E} W_2(\bar{\nu}^n(t), \bar{\mu}^n(t)) \leq C a(n) \to 0 \tag{7.35}$$

as $n \to \infty$. Therefore (7.34) holds and hence the second convergence in part (b) holds. This completes the proof. ∎

**Remark 7.1.** *The choice of $p = 3$ above (7.35) could be replaced by any $2 < p < 4$. As a result, the constant $C$ and rate $a(n)$ will change accordingly, by Lemma A.1.*

*Proof of Theorem* 5.2. (a) Similar to the proof of Theorem 5.1(a), we apply Itô's formula and get

$$\mathbb{E} |X_i^n(t) - \bar{X}_{\frac{i}{n}}(t)|^2 = \mathbb{E} \int_0^t 2(X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)) \cdot \Big( f(X_i^n(s)) - f(\bar{X}_{\frac{i}{n}}(s))$$
$$+ \frac{1}{n} \sum_{j=1}^{n} \xi_{ij}^n b(X_i^n(s), X_j^n(s)) - \int_I \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v) \, \bar{\mu}_{v,s}(dx) \, dv \Big) ds.$$

This implies that the functions

$$\bar{\alpha}_i^n(t) := \mathbb{E} |X_i^n(t) - \bar{X}_{\frac{i}{n}}(t)|^2 \quad \text{and} \quad \bar{\alpha}^n(t) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{E} |X_i^n(t) - \bar{X}_{\frac{i}{n}}(t)|^2 \tag{7.36}$$

are differentiable, and

$$\mathbb{E} |X_i^n(t) - \bar{X}_{\frac{i}{n}}(t)|^2 - \mathbb{E} |X_i^n(r) - \bar{X}_{\frac{i}{n}}(r)|^2$$
$$= \mathbb{E} \int_r^t 2(X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)) \cdot \Big( f(X_i^n(s)) - f(\bar{X}_{\frac{i}{n}}(s))$$
$$+ \frac{1}{n} \sum_{j=1}^{n} \xi_{ij}^n b(X_i^n(s), X_j^n(s)) - \int_I \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v) \, \bar{\mu}_{v,s}(dx) \, dv \Big) ds. \tag{7.37}$$

For each $s \geq 0$, using (2.1) we have

$$\mathbb{E} \left[ (X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)) \cdot \Big( f(X_i^n(s)) - f(\bar{X}_{\frac{i}{n}}(s)) \Big) \right] \leq -c_0 \mathbb{E} |X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)|^2. \tag{7.38}$$

For the rest in the integrand, by adding and subtracting terms and using Condition 5.2, we have

$$
\mathbb{E}\left[ (X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)) \right.
$$
$$
\left. \cdot \left( \frac{1}{n}\sum_{j=1}^n \xi_{ij}^n b(X_i^n(s), X_j^n(s)) - \int_I \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v)\, \bar{\mu}_{v,s}(dx)\, dv \right) \right]
$$
$$
= \mathbb{E}\left[ (X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)) \cdot \left( \frac{1}{n}\sum_{j=1}^n \xi_{ij}^n \left( b(X_i^n(s), X_j^n(s)) - b(\bar{X}_{\frac{i}{n}}(s), \bar{X}_{\frac{j}{n}}(s)) \right) \right) \right]
$$
$$
+ \mathbb{E}\left[ (X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)) \right.
$$
$$
\left. \cdot \left( \frac{1}{n}\sum_{j=1}^n \left( \xi_{ij}^n b(\bar{X}_{\frac{i}{n}}(s), \bar{X}_{\frac{j}{n}}(s)) - \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G(\frac{i}{n}, \frac{j}{n})\, \bar{\mu}_{\frac{j}{n},s}(dx) \right) \right) \right]
$$
$$
+ \mathbb{E}\left[ (X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)) \right.
$$
$$
\left. \cdot \left( \frac{1}{n}\sum_{j=1}^n \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G(\frac{i}{n}, \frac{j}{n})\, \bar{\mu}_{\frac{j}{n},s}(dx) - \int_I \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G(\frac{i}{n}, v)\, \bar{\mu}_{v,s}(dx)\, dv \right) \right]
$$
$$
=: \bar{\mathcal{S}}_s^{n,i,1} + \bar{\mathcal{S}}_s^{n,i,2} + \bar{\mathcal{S}}_s^{n,i,3}. \tag{7.39}
$$

For $\bar{\mathcal{S}}_s^{n,i,1}$, using the Cauchy-Schwarz inequality, the Lipschitz property of $b$ and Young's inequality, we have

$$
\bar{\mathcal{S}}_s^{n,i,1} \leq \mathbb{E}\left[ \left| X_i^n(s) - \bar{X}_{\frac{i}{n}}(s) \right| \cdot \frac{K_b}{n}\sum_{j=1}^n \left( \left| X_i^n(s) - \bar{X}_{\frac{i}{n}}(s) \right| + \left| X_j^n(s) - \bar{X}_{\frac{j}{n}}(s) \right| \right) \right]
$$
$$
\leq \frac{3K_b}{2}\mathbb{E}|X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)|^2 + \frac{K_b}{2}\frac{1}{n}\sum_{j=1}^n \mathbb{E}|X_j^n(s) - \bar{X}_{\frac{j}{n}}(s)|^2. \tag{7.40}
$$

For $\bar{\mathcal{S}}_s^{n,i,2}$, using the Cauchy-Schwarz inequality and the weak LLN type estimate (7.26), we have

$$
\bar{\mathcal{S}}_s^{n,i,2} \leq \left( \mathbb{E}|X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)|^2 \right)^{1/2}
$$
$$
\cdot \left( \mathbb{E}\left| \frac{1}{n}\sum_{j=1}^n \left( \xi_{ij}^n b(\bar{X}_{\frac{i}{n}}(s), \bar{X}_{\frac{j}{n}}(s)) - \int_{\mathbb{R}^d} b(\bar{X}_{\frac{i}{n}}(s), x) G_n(\frac{i}{n}, \frac{j}{n})\, \bar{\mu}_{\frac{j}{n},s}(dx) \right) \right|^2 \right)^{1/2}
$$
$$
\leq \frac{C}{\sqrt{n}} \left( \mathbb{E}|X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)|^2 \right)^{1/2}. \tag{7.41}
$$

For $\bar{\mathcal{S}}_s^{n,i,3}$, note that

$$\mathbb{E}\left|\frac{1}{n}\sum_{j=1}^{n}\int_{\mathbb{R}^d}b(\bar{X}_{\frac{i}{n}}(s),x)G(\frac{i}{n},\frac{j}{n})\,\bar{\mu}_{\frac{j}{n},s}(dx) - \int_I\int_{\mathbb{R}^d}b(\bar{X}_{\frac{i}{n}}(s),x)G(\frac{i}{n},v)\,\bar{\mu}_{v,s}(dx)\,dv\right|^2$$

$$= \mathbb{E}\left|\int_I\int_{\mathbb{R}^d}b(\bar{X}_{\frac{i}{n}}(s),x)G(\frac{i}{n},\frac{\lceil nv\rceil}{n})\,\bar{\mu}_{\frac{\lceil nv\rceil}{n},s}(dx)\,dv - \int_I\int_{\mathbb{R}^d}b(\bar{X}_{\frac{i}{n}}(s),x)G(\frac{i}{n},v)\,\bar{\mu}_{v,s}(dx)\,dv\right|^2$$

$$\leq 2\mathbb{E}\left|\int_I\int_{\mathbb{R}^d}b(\bar{X}_{\frac{i}{n}}(s),x)\left(G(\frac{i}{n},\frac{\lceil nv\rceil}{n}) - G(\frac{i}{n},v)\right)\bar{\mu}_{\frac{\lceil nv\rceil}{n},s}(dx)\,dv\right|^2$$

$$+ 2\mathbb{E}\left|\int_I\int_{\mathbb{R}^d}b(\bar{X}_{\frac{i}{n}}(s),x)\left(\bar{\mu}_{\frac{\lceil nv\rceil}{n},s}(dx) - \bar{\mu}_{v,s}(dx)\right)G(\frac{i}{n},v)dv\right|^2$$

$$\leq \frac{C}{n^2},$$

where the last inequality uses Condition 3.2, Proposition 3.1 and Remark 2.2(b) for the first term, and the Lipschitz property of $b$, (1.3) and Corollary 3.1(b) for the second term. Therefore

$$\bar{\mathcal{S}}_s^{n,i,3} \leq \left(\mathbb{E}|X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)|^2\right)^{1/2}$$

$$\cdot\left(\mathbb{E}\left|\frac{1}{n}\sum_{j=1}^{n}\int_{\mathbb{R}^d}b(\bar{X}_{\frac{i}{n}}(s),x)G(\frac{i}{n},\frac{j}{n})\,\bar{\mu}_{\frac{j}{n},s}(dx) - \int_I\int_{\mathbb{R}^d}b(\bar{X}_{\frac{i}{n}}(s),x)G(\frac{i}{n},v)\,\bar{\mu}_{v,s}(dx)\,dv\right|^2\right)^{1/2}$$

$$\leq \frac{C}{n}\left(\mathbb{E}|X_i^n(s) - \bar{X}_{\frac{i}{n}}(s)|^2\right)^{1/2}. \tag{7.42}$$

Combining (7.36)–(7.42), we have

$$\bar{\alpha}_i^n(t) - \bar{\alpha}_i^n(r) \leq -(2c_0 - 3K_b)\int_r^t\bar{\alpha}_i^n(s)\,ds + K_b\int_r^t\bar{\alpha}^n(s)\,ds + \frac{C}{\sqrt{n}}\int_r^t\sqrt{\bar{\alpha}_i^n(s)}\,ds, \tag{7.43}$$

Taking the average over $i = 1,\ldots,n$ gives

$$\bar{\alpha}^n(t) - \bar{\alpha}^n(r) \leq -2(c_0 - 2K_b)\int_r^t\bar{\alpha}^n(s)\,ds + \frac{C}{\sqrt{n}}\int_r^t\sqrt{\bar{\alpha}^n(s)}\,ds.$$

Since the function $\bar{\alpha}^n(t)$ is non-negative and differentiable with $\bar{\alpha}^n(0) = 0$, using Lemma 7.1 (with $a_1 = 2(c_0 - 2K_b)$, $a_2 = \frac{C}{\sqrt{n}}$, $a_3 = 0$) we have $\bar{\alpha}^n(t) \leq \frac{C}{n}$. From this and (7.43) we further have

$$\bar{\alpha}_i^n(t) - \bar{\alpha}_i^n(r) \leq -(2c_0 - 3K_b)\int_r^t\bar{\alpha}_i^n(s)\,ds + \frac{C}{\sqrt{n}}\int_r^t\sqrt{\bar{\alpha}_i^n(s)}\,ds + \frac{C}{n}(t - r).$$

Since the function $\bar{\alpha}_i^n(t)$ is non-negative and differentiable with $\bar{\alpha}_i^n(0) = 0$, it follows from Lemma 7.1 again that $\bar{\alpha}_i^n(t) \leq \frac{C}{n}$, uniformly in $t \geq 0$, $n \in \mathbb{N}$ and $i = 1,\ldots,n$. This gives part (a).

(b) The proof is similar to that of Theorem 5.1(b), but we will have better estimates under Conditions 3.2 and 5.2. Recall $\bar{\mu}^n(t)$ in (7.29). Taking $\pi = \frac{1}{n}\sum_{i=1}^{n}\mathcal{L}(X_i^n(t),\bar{X}_{\frac{i}{n}}(t))$ as the coupling of $\mu^n(t)$ and $\bar{\mu}^n(t)$ and using part (a), we have

$$\sup_{t\geq 0}W_2(\mu^n(t),\bar{\mu}^n(t)) \leq \sup_{t\geq 0}\left(\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}|X_i^n(t) - \bar{X}_{\frac{i}{n}}(t)|^2\right)^{1/2} \leq \frac{C}{\sqrt{n}}.$$

Using the convexity of $W_2^2(\cdot, \cdot)$ and Corollary 3.1(b) we have

$$\sup_{t \geq 0} W_2^2(\bar{\mu}^n(t), \hat{\mu}(t)) \leq \int_0^1 \sup_{t \geq 0} W_2^2(\bar{\mu}_{\frac{\lceil nu \rceil}{n}, t}, \bar{\mu}_{u, t}) \, du \leq \frac{C}{n^2}. \qquad (7.44)$$

Combining these two estimates with (7.30) gives the first statement in part (b).

For the second statement in part (b), recall $\bar{\nu}^n(t)$ in (7.32). Taking $\pi = \frac{1}{n} \sum_{i=1}^n \delta_{(X_i^n(t), \bar{X}_{\frac{i}{n}}(t))}$ as the coupling of $\nu^n(t)$ and $\bar{\nu}^n(t)$ and using part (a), we have

$$\sup_{t \geq 0} \mathbb{E} W_2(\nu^n(t), \bar{\nu}^n(t)) \leq \sup_{t \geq 0} \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E}|X_i^n(t) - \bar{X}_{\frac{i}{n}}(t)|^2 \right)^{1/2} \leq \frac{C}{\sqrt{n}}.$$

Combining this with (7.33), (7.35) and (7.44) gives the second statement in part (b).

(c) Finally, for all $n, k \in \mathbb{N}$ and any distinct $i_1, \ldots, i_k \in \{1, \ldots, n\}$, taking $\pi = \mathcal{L}\left( \left( X_{i_1}^n(t), \ldots, X_{i_k}^n(t) \right), \left( \bar{X}_{\frac{i_1}{n}}(t), \ldots, \bar{X}_{\frac{i_k}{n}}(t) \right) \right)$ as the coupling and using part (a), we have

$$\sup_{t \geq 0} W_2(\mathcal{L}(X_{i_1}^n(t), \ldots, X_{i_k}^n(t)), \bar{\mu}_{\frac{i_1}{n}, t} \otimes \cdots \otimes \bar{\mu}_{\frac{i_k}{n}, t}) \leq \left( \sum_{j=1}^k \mathbb{E}|X_{i_j}^n(t) - \bar{X}_{\frac{i_j}{n}}(t)|^2 \right)^{1/2} \leq \frac{C\sqrt{k}}{\sqrt{n}}.$$

This gives part (c) and completes the proof.                                   ∎

### 7.4. Proofs for Section 6.

We first show the following uniform-in-time estimates.

**Lemma 7.3.** *There exist $h_0, C \in (0, \infty)$ such that*

$$\sup_{n \geq 1} \max_{i=1,\ldots,n} \mathbb{E}|X_i^{n,h}(s) - X_i^{n,h}(s_h)|^2 \leq C(s - s_h) \leq Ch, \quad \forall \, s \geq 0, \, h \in (0, h_0),$$

*and*

$$\sup_{h \in (0,h_0)} \sup_{n \geq 1} \max_{i=1,\ldots,n} \sup_{t \geq 0} \mathbb{E}|X_i^{n,h}(t)|^2 \leq C.$$

*Proof.* Before analyzing the system (6.1), consider the following equivalent discrete-time model: $Z_i^{n,h}(0) = X_i^{n,h}(0)$ and

$$Z_i^{n,h}(k+1) = Z_i^{n,h}(k) + \left( f(Z_i^{n,h}(k)) + \frac{1}{n} \sum_{j=1}^n \xi_{ij}^n b(Z_i^{n,h}(k), Z_j^{n,h}(k)) \right) h + \Delta_k B_{\frac{i}{n}}, \quad k \in \mathbb{N}_0,$$

where $\Delta_k B_{\frac{i}{n}} := B_{\frac{i}{n}}((k+1)h) - B_{\frac{i}{n}}(kh)$. Note that $Z_i^{n,h}(k) := X_i^{n,h}(kh)$.

We claim that

$$\sup_{h \in (0,h_0)} \sup_{n \geq 1} \max_{i=1,\ldots,n} \sup_{k \in \mathbb{N}_0} \mathbb{E}|Z_i^{n,h}(k)|^2 < \infty, \qquad (7.45)$$

for some $h_0 \in (0, \infty)$. To see this, write

$$|Z_i^{n,h}(k+1)|^2 - |Z_i^{n,h}(k)|^2$$

$$= 2Z_i^{n,h}(k) \cdot \left( Z_i^{n,h}(k+1) - Z_i^{n,h}(k) \right) + |Z_i^{n,h}(k+1) - Z_i^{n,h}(k)|^2$$

$$= 2Z_i^{n,h}(k) \cdot f(Z_i^{n,h}(k))h + 2Z_i^{n,h}(k) \cdot \left( \frac{1}{n} \sum_{j=1}^{n} \xi_{ij}^n b(Z_i^{n,h}(k), Z_j^{n,h}(k)) \right) h + \zeta_i^{n,h}(k) \cdot \sigma \Delta_k B_{\frac{i}{n}}$$

$$+ \left| f(Z_i^{n,h}(k)) + \frac{1}{n} \sum_{j=1}^{n} \xi_{ij}^n b(Z_i^{n,h}(k), Z_j^{n,h}(k)) \right|^2 h^2 + \left| \sigma \Delta_k B_{\frac{i}{n}} \right|^2,$$

where $\zeta_i^{n,h}(k)$ is measurable with respect to $\sigma\{B_{\frac{j}{n}}(s) : 0 \le s \le kh\}$. Let $\beta_i^{n,h}(k) :=$ $\mathbb{E}|Z_i^{n,h}(k)|^2$ and $\beta^{n,h}(k) := \frac{1}{n} \sum_{j=1}^{n} \mathbb{E}|Z_j^{n,h}(k)|^2$. Using (2.1), the Lipschitz property of $f, b$ and the Cauchy-Schwarz inequality we have

$$\beta_i^{n,h}(k+1) - \beta_i^{n,h}(k) = \mathbb{E}|Z_i^{n,h}(k+1)|^2 - \mathbb{E}|Z_i^{n,h}(k)|^2$$

$$\le \left( -2c_0 \beta_i^{n,h}(k) + C\sqrt{\beta_i^{n,h}(k)} + 2K_b \beta_i^{n,h}(k) + 2K_b \sqrt{\beta_i^{n,h}(k)\beta^{n,h}(k)} \right) h$$

$$+ C \left( 1 + \beta_i^{n,h}(k) + \beta^{n,h}(k) \right) h^2 + Ch.$$

Since $C\sqrt{\beta_i^{n,h}(k)} \le (c_0 - 2K_b)\beta_i^{n,h}(k) + \frac{C^2}{4(c_0 - 2K_b)}$ and $2\sqrt{\beta_i^{n,h}(k)\beta^{n,h}(k)} \le \beta_i^{n,h}(k) + \beta^{n,h}(k)$, we have

$$\beta_i^{n,h}(k+1) - \beta_i^{n,h}(k) \tag{7.46}$$

$$\le \left( -(c_0 - K_b)\beta_i^{n,h}(k) + K_b \beta^{n,h}(k) + C \right) h + C \left( 1 + \beta_i^{n,h}(k) + \beta^{n,h}(k) \right) h^2.$$

Taking the average over $i = 1, \ldots, n$ gives

$$\beta^{n,h}(k+1) \le (1 - h\kappa_h)\beta^{n,h}(k) + Ch,$$

where $\kappa_h := c_0 - 2K_b - Ch$. From (2.2) we can choose $h_0 > 0$ such that $\inf_{h \in (0,h_0)} \kappa_h > 0$ and $1 - h\kappa_h \in (0, 1)$ for all $h \in (0, h_0)$. Then for all $h \in (0, h_0)$,

$$\beta^{n,h}(k+1) \le (1 - h\kappa_h)^2 \beta^{n,h}(k-1) + (1 - h\kappa_h)Ch + Ch \le \cdots$$

$$\le (1 - h\kappa_h)^{k+1} \beta^{n,h}(0) + \sum_{j=0}^{k} (1 - h\kappa_h)^j Ch$$

$$\le (1 - h\kappa_h)^{k+1} C + \frac{Ch}{1 - (1 - h\kappa_h)} \le C.$$

Applying this back to (7.46) gives

$$\beta_i^{n,h}(k+1) \le (1 - h\kappa_h)\beta_i^{n,h}(k) + Ch,$$

which again gives $\beta_i^{n,h}(k+1) \le C$ and verifies (7.45).

Using (7.45) and Remark 2.2(b), we immediately have the first statement, which further implies the second statement. This completes the proof. ∎

*Proof of Theorem* 6.1. Recall $h_0$ in Lemma 7.3. Using Itô's formula, we have

$$\mathbb{E}|X_i^{n,h}(t) - X_i^n(t)|^2 = \mathbb{E}\int_0^t 2\left(X_i^{n,h}(s) - X_i^n(s)\right)\cdot\left(f(X_i^{n,h}(s_h)) - f(X_i^n(s))\right.$$

$$\left. + \frac{1}{n}\sum_{j=1}^n \xi_{ij}^n b(X_i^{n,h}(s_h), X_j^{n,h}(s_h)) - \frac{1}{n}\sum_{j=1}^n \xi_{ij}^n b(X_i^n(s), X_j^n(s))\right) ds.$$

This implies that the functions

$$\gamma_i^n(t) := \mathbb{E}|X_i^{n,h}(t) - X_i^n(t)|^2, \quad \gamma^n(t) := \frac{1}{n}\sum_{i=1}^n \mathbb{E}|X_i^{n,h}(t) - X_i^n(t)|^2$$

are differentiable.

By adding and subtracting terms, we have

$$\mathbb{E}\left[\left(X_i^{n,h}(s) - X_i^n(s)\right)\cdot\left(f(X_i^{n,h}(s_h)) - f(X_i^n(s))\right)\right]$$

$$\leq \mathbb{E}\left[\left(X_i^{n,h}(s) - X_i^n(s)\right)\cdot\left(f(X_i^{n,h}(s_h)) - f(X_i^{n,h}(s))\right)\right]$$

$$+ \mathbb{E}\left[\left(X_i^{n,h}(s) - X_i^n(s)\right)\cdot\left(f(X_i^{n,h}(s)) - f(X_i^n(s))\right)\right]$$

$$\leq C\sqrt{\gamma_i^n(s)h} - c_0\gamma_i^n(s),$$

where the last line uses the Cauchy-Schwarz inequality, the Lipschitz property of $f$ and Lemma 7.3 for the first term and (2.1) for the second term. Also by adding and subtracting terms, we have

$$\mathbb{E}\left[\left(X_i^{n,h}(s) - X_i^n(s)\right)\cdot\left(\frac{1}{n}\sum_{j=1}^n \xi_{ij}^n b(X_i^{n,h}(s_h), X_j^{n,h}(s_h)) - \frac{1}{n}\sum_{j=1}^n \xi_{ij}^n b(X_i^n(s), X_j^n(s))\right)\right]$$

$$= \mathbb{E}\left[\left(X_i^{n,h}(s) - X_i^n(s)\right)\cdot\frac{1}{n}\sum_{j=1}^n \xi_{ij}^n\left(b(X_i^{n,h}(s_h), X_j^{n,h}(s_h)) - b(X_i^{n,h}(s), X_j^{n,h}(s))\right)\right]$$

$$+ \mathbb{E}\left[\left(X_i^{n,h}(s) - X_i^n(s)\right)\cdot\frac{1}{n}\sum_{j=1}^n \xi_{ij}^n\left(b(X_i^{n,h}(s), X_j^{n,h}(s)) - b(X_i^n(s), X_j^n(s))\right)\right]$$

$$\leq C\sqrt{\gamma_i^n(s)h} + K_b\gamma_i^n(s) + K_b\sqrt{\gamma_i^n(s)\gamma^n(s)},$$

where the last line uses the Cauchy-Schwarz inequality, the Lipschitz property of $b$ and Lemma 7.3. Combining these two estimates gives

$$\gamma_i^n(t) - \gamma_i^n(r) \leq -2(c_0 - K_b)\int_r^t \gamma_i^n(s)\,ds + C\sqrt{h}\int_r^t \sqrt{\gamma_i^n(s)}\,ds + 2K_b\int_r^t \sqrt{\gamma_i^n(s)\gamma^n(s)}\,ds \quad (7.47)$$

for all $t > r \geq 0$. Taking the average over $i = 1, \ldots, n$, we get

$$\gamma^n(t) - \gamma^n(r) \leq -2(c_0 - 2K_b)\int_r^t \gamma^n(s)\,ds + C\sqrt{h}\int_r^t \sqrt{\gamma^n(s)}\,ds.$$

Since the function $\gamma^n(t)$ is non-negative and differentiable with $\gamma^n(0) = 0$, using Lemma 7.1 (with $a_1 = 2(c_0 - 2K_b), a_2 = C\sqrt{h}, a_3 = 0$) we have $\gamma^n(t) \leq Ch$. Applying this to (7.47) gives

$$\gamma_i^n(t) - \gamma_i^n(r) \leq -2(c_0 - K_b) \int_r^t \gamma_i^n(s)\, ds + C\sqrt{h} \int_r^t \sqrt{\gamma_i^n(s)}\, ds.$$

Since the function $\gamma_i^n(t)$ is non-negative and differentiable with $\gamma_i^n(0) = 0$, it follows from Lemma 7.1 again that $\gamma_i^n(t) \leq Ch$, uniformly in $h \in (0, h_0)$, $t \geq 0$, $n \in \mathbb{N}$ and $i = 1, \ldots, n$. This completes the proof. ∎

## APPENDIX A. A WASSERSTEIN DISTANCE RESULT

In this section we prove Lemma A.1 on the Wasserstein distance about the empirical measure of independent (but not necessarily identically distributed) random variables. It is a natural generalization of [20, Theorem 1] where i.i.d. samples are studied. It is also worth mentioning that for i.i.d. samples, the upper bounds are obtained in [22, Lemma 3.7 and Appendix] for complete cases with explicit constants that was not provided in [20, Theorem 1]. But the three cases in Lemma A.1 below are sufficient for our use and we provide a proof for completeness.

**Lemma A.1.** *Let $\{Y_i : i \in \mathbb{N}\}$ be independent $\mathbb{R}^d$-valued random variables. Write*

$$\bar{\mu}_i := \mathcal{L}(Y_i), \quad \nu_n := \frac{1}{n}\sum_{i=1}^n \delta_{Y_i}, \quad \bar{\nu}_n := \frac{1}{n}\sum_{i=1}^n \bar{\mu}_i.$$

*Let $p > 0$. Assume that $\sup_{i \in \mathbb{N}} \mathbb{E}|Y_i|^q < \infty$ for some $q > p$. Then there exists a constant $C$ depending only on $p, q, d$ such that, for all $n \geq 1$,*

$$\mathbb{E}W_p^p(\nu_n, \bar{\nu}_n) \leq C \left(\int_{\mathbb{R}^d} |x|^q\, \bar{\nu}_n(dx)\right)^{p/q}$$

$$\times \begin{cases} n^{-1/2} + n^{-(q-p)/q} & \text{if } p > d/2 \text{ and } q \neq 2p, \\ n^{-1/2}\log(1+n) + n^{-(q-p)/q} & \text{if } p = d/2 \text{ and } q \neq 2p, \\ n^{-p/d} + n^{-(q-p)/q} & \text{if } p \in (0, d/2) \text{ and } q \neq d/(d-p). \end{cases}$$

*Proof of Lemma* A.1. Fix $A \subset \mathbb{R}^d$. In view of the proof of [20, Theorem 1], it suffices to verify that

$$\mathbb{E}|\nu_n(A) - \bar{\nu}_n(A)| \leq \min\left\{2\bar{\nu}_n(A), \sqrt{\bar{\nu}_n(A)/n}\right\}. \tag{A.1}$$

For this, clearly we have

$$\mathbb{E}|\nu_n(A) - \bar{\nu}_n(A)| \leq \mathbb{E}\nu_n(A) + \bar{\nu}_n(A) = 2\bar{\nu}_n(A).$$

Also note that, by the independence of $\{Y_i : i \in \mathbb{N}\}$,

$$\mathbb{E}|\nu_n(A) - \bar{\nu}_n(A)|^2 = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \left(\mathbf{1}_{\{Y_i \in A\}} - \bar{\mu}_i(A)\right)\right]^2 = \frac{1}{n^2}\sum_{i=1}^n \mathbb{E}\left[\mathbf{1}_{\{Y_i \in A\}} - \bar{\mu}_i(A)\right]^2$$

$$= \frac{1}{n^2}\sum_{i=1}^n \bar{\mu}_i(A)(1 - \bar{\mu}_i(A)) \leq \frac{1}{n^2}\sum_{i=1}^n \bar{\mu}_i(A) = \frac{1}{n}\bar{\nu}_n(A).$$

This completes the proof. ∎

## References

[1] A. Aurell, R. Carmona, and M. Lauriere, *Stochastic graphon games: Ii. the linear-quadratic case*, arXiv preprint arXiv:2105.12320 (2021).

[2] J. Barré, P. Dobson, M. Ottobre, and E. Zatorska, *Fast non mean-field networks: Uniform in time averaging*, SIAM Journal on Mathematical Analysis **53** (2021), no. 1, 937–972.

[3] E. Bayraktar, S. Chakraborty, and R. Wu, *Graphon mean field systems*, arXiv preprint arXiv:2003.13180 (2020).

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

[5] G. Bet, F. Coppini, and F. R Nardi, *Weakly interacting oscillators on dense random graphs*, arXiv preprint arXiv:2006.07670 (2020).

[6] 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.

[7] F. Bolley, A. Guillin, and C. Villani, *Quantitative concentration inequalities for empirical measures on non-compact spaces*, Probability Theory and Related Fields **137** (2007), no. 3-4, 541–593.

[8] A. Budhiraja and W.-T. L. Fan, *Uniform in time interacting particle approximations for nonlinear equations of Patlak-Keller-Segel type*, Electron. J. Probab. **22** (2017), 37 pp.

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

[10] A. Budhiraja and A. Pal Majumder, *Long time results for a weakly interacting particle system in discrete time*, Stochastic Analysis and Applications **33** (2015), no. 3, 429–463.

[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] P. E Caines and M. Huang, *Graphon mean field games and the GMFG equations*, arXiv preprint arXiv:2008.10216 (2020).

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

[14] R. Carmona and F. Delarue, *Probabilistic analysis of mean-field games*, SIAM Journal on Control and Optimization **51** (2013), no. 4, 2705–2734.

[15] F. Coppini, *Long time dynamics for interacting oscillators on graphs*, arXiv preprint arXiv:1908.01520 (2019).

[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] P. Dupuis and G. Medvedev, *The large deviation principle for interacting dynamical systems on random graphs*, arXiv preprint arXiv:2007.13899 (2020).

[20] N. Fournier and A. Guillin, *On the rate of convergence in wasserstein distance of the empirical measure*, Probability Theory and Related Fields **162** (2015), no. 3-4, 707–738.

[21] S. Gao, R. F. Tchuendom, and P. E Caines, *Linear quadratic graphon field games*, arXiv preprint arXiv:2006.03964 (2020).

[22] G. Guo, J. Obłój, et al., *Computational methods for martingale optimal transport problems*, The Annals of Applied Probability **29** (2019), no. 6, 3311–3347.

[23] D. Kaliuzhnyi-Verbovetskyi and G. S Medvedev, *The mean field equation for the kuramoto model on graph sequences with non-lipschitz limit*, SIAM Journal on Mathematical Analysis **50** (2018), no. 3, 2441–2465.

[24] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics, vol. 113, Springer New York, 1991.

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

[26] D. Lacker and A. Soret, *A case study on stochastic games on large graphs in mean field and sparse regimes*, arXiv preprint arXiv:2005.14102 (2020).

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

[28] E. Luçon, *Quenched asymptotics for interacting diffusions on inhomogeneous random graphs*, Stochastic Processes and their Applications (2020).

[29] 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.

[30] G. S Medvedev, *The nonlinear heat equation on dense graphs and graph limits*, SIAM Journal on Mathematical Analysis **46** (2014), no. 4, 2743–2766.

[31] G. S Medvedev, *The nonlinear heat equation on w-random graphs*, Archive for Rational Mechanics and Analysis **212** (2014), no. 3, 781–803.

[32] R. Oliveira and G. Reis, *Interacting diffusions on random graphs with diverging average degrees: Hydrodynamics and large deviations*, Journal of Statistical Physics **176** (2019), 1057–1087.

[33] 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).

[34] A. D Polyanin and A. V Manzhirov, *Handbook of integral equations*, CRC press, 2008.

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

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

[37] D. Vasal, R. K Mishra, and S. Vishwanath, *Sequential decomposition of graphon mean field games*, arXiv preprint arXiv:2001.05633 (2020).

[38] A Y. Veretennikov, *On ergodic measures for mckean-vlasov stochastic equations*, Monte carlo and quasi-monte carlo methods 2004, 2006, pp. 471–486.

[39] C. Villani, *Optimal transport: old and new*, Vol. 338, Springer Science & Business Media, 2008.

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

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