CONVERGENCE TOWARD EQUILIBRIUM OF THE FIRST-ORDER CONSENSUS MODEL WITH RANDOM BATCH INTERACTIONS

SEUNG-YEAL HA, SHI JIN, DOHEON KIM, AND DONGNAM KO

ABSTRACT. We study convergence analysis of the first-order stochastic consensus model with random batch interactions. The proposed model can be obtained via random batch method (RBM) from the first-order nonlinear consensus model. This model has two competing mechanisms, namely intrinsic free flow and nonlinear consensus interaction terms. From the competition between the two mechanisms, the original (full batch) model can admit relative equilibria and relaxation of the dynamics to the relative equilibrium in a large coupling regime. In authors' earlier work, we have studied the RBM approximation and its uniform error analysis. In this paper, we present two convergence analysis of RBM solutions toward the relative equilibrium. More precisely, we show that the variances of displacement processes between the full batch and random batch solutions tend to zero exponentially fast, as time goes to infinity. Second, we also show that, almost surely, the diameter process of displacement tends to zero exponentially fast.

1. INTRODUCTION

Emergent behaviors of interacting particle systems are ubiquitous in human society and nature, e.g., flocking of birds [12, 13, 42], swarming of fish [4, 40], synchronization of pacemaker cells and fireflies [7, 35, 36], coordinated control of vehicular traffic [2], consensus phenomenon [38], and so on. Thus, it is a very challenging and important task to figure out how coherent motions can emerge from a set of simple rules in complex systems [1, 19, 39, 41]. Recently, due to practical needs in engineering applications, mathematical modeling and simulations of such collective models have received lots of attention [5, 8, 21, 29, 34, 42] from the communities of applied mathematics and control theory.

In this paper, we deal with the first-order nonlinear consensus model which was studied in author's recent work [28], and study convergence analysis for the model with the random batch interactions. To be more specific, let $q_i = q_i(t) \in \mathbb{R}$ be an observable of the i-th particle in the whole ensemble which can exhibit consensus dynamics, and we set $Q = (q_1, \cdots, q_N) \in \mathbb{R}^N$ to be an ensemble state vector. In what follows, we also use the same notation Q to denote a set $\{q_i\}$ as long as there is no confusion. In [28], we assume that the temporal dynamics of q_i is governed by the Cauchy problem to the following first-order

Date: March 19, 2021.

2010 *Mathematics Subject Classification.* 37M99, 37N30, 65P99.

Key words and phrases. consensus, interacting particle system, random batch.

Acknowledgment. The work of S.-Y. Ha was supported by National Research Foundation of Korea (NRF-2020R1A2C3A01003881), the work of S. Jin was supported by National Natural Science Foundation of China grant 12031013, and Science and Technology Commission of Shanghai Municipality grant No. 20JC1414100, the work of D. Kim was supported by a KIAS Individual Grant (MG073901) at Korea Institute for Advanced Study, and the work of D. Ko was supported by the Catholic University of Korea, Research Fund, 2021.

nonlinear consensus model [28]:

$$\begin{cases} \dfrac{dq_i}{dt} = \nu_i + \dfrac{\kappa}{N-1} \displaystyle\sum_{j \neq i} \Gamma(q_j - q_i), \quad t > 0, \\ q_i(0) = q_i^{in}, \quad i = 1, \cdots, N. \end{cases} \tag{1.1}$$

Here ν_i and κ are natural (generalized) velocity of the i-th particle and the nonnegative coupling strength, respectively. Throughout the paper, without loss of generality, we assume that the total sum of ν_i and initial data are zero (see Section 2.1):

$$\sum_{i=1}^{N} \nu_i = 0 \quad \text{and} \quad \sum_{i=1}^{N} q_i^{in} = 0. \tag{1.2}$$

The interaction kernel Γ in (1.1) is also assumed as a function of relative states and satisfies the bi-Lipschitz regularity and anti-symmetric property on some symmetric interval around zero: there exist positive constants r_0, L_1 and L_2 such that

$$\begin{aligned} &\Gamma(-q) = -\Gamma(q) \quad \text{and} \\ &L_1|p-q|^2 \leq (\Gamma(p) - \Gamma(q)) \cdot (p-q) \leq L_2|p-q|^2, \quad \forall\, p, q \in [-r_0, r_0]. \end{aligned} \tag{1.3}$$

The anti-symmetric convolution-type interaction function Γ can be found in many physics-based consensus models with translation invariance, e.g., synchronization [15, 10, 11, 20, 30, 43], flocking [12, 17], and consensus [38], though the symmetry may break down from the nonsymmetric nature [27, 33] or geometric constraints [6, 9, 22, 23, 32].

In system (1.1)–(1.2), it is known from [28] that if the coupling strength κ is sufficiently large compared to $\|\nu\|_2 = \Big(\sum_{i=1}^{N} |\nu_i|^2 \Big)^{\frac{1}{2}}$, it admits a unique equilibrium $Q^\infty := (q_1^\infty, \cdots, q_N^\infty)$ satisfying the following relations (See Proposition 2.1 for details):

$$\sum_{i=1}^{N} q_i^\infty = 0 \quad \text{and} \quad \max_{1 \leq i,j \leq N} |q_i^\infty - q_j^\infty| < r_0.$$

Although the dynamic feature of the state vector Q may be simple, numerical simulation of (1.1) is not an easy problem when N is large. Note that q_i interacts with all the other particles $q_1, \cdots, q_N$ in the all-to-all interacting network. This implies that the computational complexity for the whole interactions is of $\mathcal{O}(N^2)$ per time step, which quickly becomes unfeasible, as the number of particles N increases. This curse of dimensionality poses main computational challenge for numerical integration of interacting particle systems, when we deal with collective dynamics models. Thus, the designing of a suitable *"approximate model"* preserving key properties of the original system, yet computationally affordable, becomes an important problem in the era of big data. Many attempts have been done to approximate the dynamics with less complexity, to name a few, the fast multipole method [37] is a classical method, and recently, the random batch method (RBM) [25] is proposed and applied to many interacting particle models. We refer to [24] for related issues.

In this paper, we address the following two questions:

- (Q1): Are there suitable RBM approximations of system (1.1) which can preserve emergent behaviors of the full batch model?
- (Q2): If there exists such an RBM approximation in (Q1), are there suitable frameworks which guarantee convergence analysis in suitable sense?

In what follows, we discuss the aforementioned questions in depth and present two main results in relation with the second question (Q2).

The first question (Q1) is concerned with a complexity reduction problem. The RBM approximates the total interaction on each particle by randomly sampling a small batch of particles of size P (we choose P a priori with $1 < P \ll N$). In particular, we consider a random partition of the whole population $\{1, \cdots, N\}$ into small batches at each time interval (If P does not divide N, then one batch may have a smaller number of particles), and we assume that each particle only interacts with the ones in the same batch for a small duration of time, and then the batches are randomly rearranged for the next time interval and so on. Therefore, the RBM system can be regarded as a kind of (randomly) switching networked system (see [25, 28] for justification). In this way, the number of interactions we take reduces to the order of $\mathcal{O}(NP)$ which has linear growth in N. Below, we briefly explain RBM approximation of (1.1) in more detail.

For simplicity, we consider the case $\frac{N}{P} \in \mathbb{N}$. Let τ be a positive small time-step and $t_m := m\tau$ for $m = 0, 1, 2, \cdots$ be the instant of time at which we change batches. For each time interval $[t_m, t_{m+1})$, we consider a random partition of $\{1, 2, \ldots, N\}$:

$$\{1, \ldots, N\} = C_1^m \cup C_2^m \cup \cdots \cup C_{N/P}^m, \quad |C_i^m| = P, \quad i = 1, \cdots, N/P.$$

We set $[i]_m \subset \{1, \ldots, N\}$ to be the batch containing the i-th particle in the time interval $[t_m, t_{m+1})$. Then, the naive and honest RBM approximated system for (1.1) will be

$$\frac{dq_i^R}{dt} = \nu_i + \frac{\kappa}{P-1} \sum_{j \in [i]_m} \Gamma(q_j^R - q_i^R), \quad t \in [t_m, t_{m+1}). \tag{1.4}$$

However, as noticed in [28], system (1.4) may not admit an equilibrium in general and may have unbounded solutions. Even if the equilibrium Q^∞ of the original system exists, the RBM solution can fluctuate around Q^∞ continuously. Therefore, one cannot expect a convergence result to the equilibrium (though the error between the approximated and original solutions is small for $\tau \ll 1$), and the simulations in [16, 26, 31] average the data in time to approximate the equilibrium profile. More precisely, we set displacements:

$$\hat{q}_i^R := q_i^R(t) - q_i^\infty, \quad 1 \le i \le N, \quad \widehat{Q}^R := (\hat{q}_1^R, \cdots, \hat{q}_N^R).$$

Then, for system (1.4), the boundedness of the displacement process $\widehat{Q}^R(t)$ cannot be verified as it is, even if it is initially small enough (see [28] or Lemma 4.1 for a detailed statement).

Therefore, we need to leverage the effect of the free flow registered by natural velocities ν_i's accordingly when we take small number of interactions in each batch. It seems that there is no systematic way to decompose ν_i's in general, because ν_i's are just super parameters to be specified a priori. Fortunately, they can be related to the equilibrium solution if it exists. In [28], the following RBM approximation was proposed in a perturbed regime of an equilibrium $Q^\infty = (q_1^\infty, \cdots, q_N^\infty)$:

$$\begin{cases} \dfrac{dq_i^R}{dt} = \dfrac{\kappa}{P-1} \displaystyle\sum_{j \in [i]_m} \Big(\Gamma(q_j^R - q_i^R) - \Gamma(q_j^\infty - q_i^\infty) \Big), \quad t \in [t_m, t_{m+1}), \\ q_i^R(0) = q_i^{in}, \quad i = 1, \cdots, N, \ m = 0, 1, 2, \cdots. \end{cases} \tag{1.5}$$

When $P = N$, one can easily check that both (1.4) and (1.5) become the original full batch model (1.1). A significant difference of (1.5) from (1.4) is the boundedness of the maximal displacement from the background equilibrium profile:

$$\|\widehat{Q}^R(t)\|_\infty := \max_{1\le i\le N} |\hat{q}_i^R(t)| < \infty.$$

That is why we restrict our discussion to system (1.5) throughout this paper.

Before we discuss our main results, we briefly introduce the error estimate of (1.5) in authors' recent work [28]. Let Q and Q^R be solutions to (1.1) and (1.5), respectively. Then, the uniform error analysis yields

$$\sup_{0\le t<\infty} \mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^N |q_i(t) - q^R(t)|^2\Big) \lesssim \Big[\tau\left(\frac{1}{P-1} - \frac{1}{N-1}\right) + \tau^2\Big]. \tag{1.6}$$

If we combine the decay estimate of the original system (see Remark 2.2 for details) and the error estimate (1.6) by a triangle inequality, one may conclude that solution to the system stay near the equilibrium exponentially fast, however, up to the error in (1.6). Therefore, this error analysis is not sufficient to infer the convergence speed of $Q^R(t)$ to the background equilibrium profile Q^∞.

Now, we discuss our two main results on the emergent dynamics of (1.5) as follows. To estimate the error from the RBM approximation, we introduce two functionals:

$$\mathcal{M}_2(\widehat{Q}^R) := \frac{1}{N}\sum_{j=1}^N |\hat{q}_j^R|^2, \quad \mathcal{D}(\widehat{Q}^R) := \max_{1\le i,j\le N} |\hat{q}_i^R - \hat{q}_j^R|.$$

Our first main result is concerned with the exponential decay of the second moment of $\hat{q}_i^R$ using the stochastic properties suggested in [18, 25]. More precisely, we obtain the following decay estimate (see Theorem 3.1): there exists a positive constant $\Lambda_1 = \Lambda_1(N, P, \tau, \kappa, L_1)$ satisfying

$$\mathbb{E}\Big(\mathcal{M}_2(\widehat{Q}^R(t))\Big) \le e^{-\Lambda_1 t}\mathbb{E}\Big(\mathcal{M}_2(\widehat{Q}^R(0))\Big), \quad t \ge 0.$$

Next, our second main result deals with almost sure (a.s.) convergence of $\widehat{Q}^R$ by improving techniques on switching network topology in [14]: there exists a positive constant $\Lambda_2 = \Lambda_2(N, P, \tau, \kappa, L_1, L_2)$ such that

$$\mathcal{D}(\widehat{Q}^R(t)) \le \mathcal{D}(\widehat{Q}^R(0)) C e^{-\Lambda_2 t}, \quad t \ge 0,$$

where the detailed estimation of Λ_2 is described later in Theorem 3.2.

To the authors' knowledge, this is the first result on convergence of the RBM toward the global equilibrium.

The rest of this paper is organized as follows. In Section 2, we briefly discuss the properties of the consensus model (1.1) and its RBM approximated system (1.5). Then, in Section 3, we state our framework and main results on the convergence analysis of solutions to the RBM approximation toward the equilibrium. In Section 4 and Section 5, we provide detailed proofs for these main results. In Section 6, we present numerical simulations of linear consensus model to compare with theoretical results. Finally, Section 7 is devoted to a brief summary of our results and discussions on some remaining issues for a future work.

Notation: For $\nu = (\nu_1, \cdots, \nu_N)$ and $Q = (q_1, \cdots, q_N)$, we set

$$\|\nu\|_2 := \Big(\sum_{i=1}^N |\nu_i|^2 \Big)^{\frac{1}{2}}, \quad \|Q\|_2 := \Big(\sum_{i=1}^N |q_i|^2 \Big)^{\frac{1}{2}}, \quad \|\nu\|_\infty := \max_{1\le i\le N} |\nu_i|,$$
$$\|Q\|_\infty := \max_{1\le i\le N} |q_i|, \quad \mathcal{D}(\nu) := \max_{i,j} |\nu_i - \nu_j|, \quad \mathcal{D}(Q) := \max_{i,j} |q_i - q_j|.$$

2. Preliminaries

In this section, we briefly review a reformulation of (1.1) into another consensus model with zero free flow, existence of relative equilibrium and the RBM approximation of the reformulated system.

2.1. A nonlinear consensus model with zero free flow.

In this subsection, we consider system (1.1) in a perturbed regime around an equilibrium. Note that the coupling function Γ is anti-symmetric and locally bounded near zero, hence

$$\Gamma(0) = 0$$

and system (1.1) becomes

$$\begin{cases} \dfrac{dq_i}{dt} = \nu_i + \dfrac{\kappa}{N-1} \displaystyle\sum_{j=1}^N \Gamma(q_j - q_i), \quad t > 0, \\ q_i(0) = q_i^{in}, \quad i = 1, \cdots, N. \end{cases} \tag{2.1}$$

For a solution $Q = \{q_i\}$ to (2.1), we set

$$S(t) := \sum_{i=1}^N q_i(t) - t \sum_{i=1}^N \nu_i, \quad t \ge 0. \tag{2.2}$$

In the sequel, we will see that the functional S is constant.

Lemma 2.1. [28] *Let $Q = \{q_i\}$ be a solution to* (2.1). *Then, the following structural assertions hold.*

(1) (Translation-invariance): system (2.1) *is invariant under the translation $\tilde{q}_i := q_i + \alpha,\ \alpha \in \mathbb{R},\ i = 1, \cdots, N$:*

$$\frac{d\tilde{q}_i}{dt} = \nu_i + \frac{\kappa}{N-1} \sum_{j=1}^N \Gamma(\tilde{q}_j - \tilde{q}_i).$$

(2) (Conservation law): the quantity $S(t)$ in (2.2) *is conserved along* (2.1)*:*

$$S(t) = S(0), \quad t > 0.$$

Remark 2.1. *Due to the translation invariance, without loss of generality, we may assume that*

$$\sum_{j=1}^N q_j^{in} = 0 \quad \textit{and} \quad \sum_{j=1}^N \nu_j = 0,$$

otherwise, we may consider a reference frame moving with velocity $\frac{1}{N}\sum_{j=1}^N \nu_j$, *with the origin* $\frac{1}{N}\sum_{j=1}^N q_j^{in}$ *at* $t = 0$.

Next, we assume the existence of an equilibrium Q^∞ and derive the RBM–approximated model (1.5) with local conservation. First, we consider an equilibrium system:

$$\nu_i + \frac{\kappa}{N-1}\sum_{j=1}^N \Gamma(\phi_j - \phi_i) = 0, \quad i = 1, \cdots, N \tag{2.3}$$

subject to zero sum constraints:

$$\sum_{j=1}^N \phi_j = 0 \quad \text{and} \quad \sum_{j=1}^N \nu_j = 0. \tag{2.4}$$

In general, due to translational invariance, without any constraints, system (2.3) may not admit a solution. Thus, we need to consider a relative equilibrium. This is why we impose the constraints (2.4) to guarantee a unique equilibrium. Note that the interaction terms in the L.H.S. of (2.3) satisfy

$$\Big|\frac{\kappa}{N-1}\sum_{j=1}^N \Gamma(\phi_j - \phi_i)\Big| \le \frac{\kappa N \Gamma^\infty}{N-1}, \quad i = 1, \cdots, N, \qquad \text{where} \quad \Gamma^\infty := \max_{-r_0 \le q \le r_0} |\Gamma(q)|.$$

If the coupling strength κ is sufficiently small such that

$$\kappa < \frac{\min_i |\nu_i|}{\Gamma^\infty}\Big(\frac{N-1}{N}\Big),$$

then equilibrium system (2.3) does not have a solution. The existence issue of equilibrium will be discussed in the following subsection. For the time being, we assume that equilibrium $Q^\infty := \{q_i^\infty\}$ exists, and one has

$$\sum_{j=1}^N q_j^\infty = \sum_{j=1}^N q_j^{in} = 0, \quad \nu_i = \frac{\kappa}{N-1}\sum_{j=1}^N \Big[-\Gamma(q_j^\infty - q_i^\infty)\Big] \quad i = 1, \cdots, N. \tag{2.5}$$

Now, we combine (2.1) and $(2.5)_2$ to get a nonlinear consensus model with zero free flow and its effective RBM approximation:

$$\begin{aligned} \frac{dq_i}{dt} &= \frac{\kappa}{N-1}\sum_{j=1}^N \Big(\Gamma(q_j - q_i) - \Gamma(q_j^\infty - q_i^\infty)\Big), \ t > 0, \\ \frac{dq_i^R}{dt} &= \frac{\kappa}{P-1}\sum_{j \in [i]_m} \Big(\Gamma(q_j^R - q_i^R) - \Gamma(q_j^\infty - q_i^\infty)\Big), \ t \in [t_m, t_{m+1}), \quad m = 0, 1, \cdots. \end{aligned} \tag{2.6}$$

Although we need to know the equilibrium Q^∞ exactly, the RBM approximation $(2.6)_2$ preserves good properties of the original system $(2.6)_1$. One of such properties is the anti-symmetry of interaction term $(\Gamma(q_j^R(t) - q_i^R(t)) - \Gamma(q_j^\infty - q_i^\infty))$ under the exchange symmetry $i \longleftrightarrow j$. This induces a local conservation law in each batch to be crucially used in our main theorems presented in Section 3.

Lemma 2.2. *Let $Q^R = \{q_i^R\}$ be a solution to* (1.5). *Then, for any $k = 1, \cdots, N$ and $m = 0, 1, 2, \cdots$, we have*

$$\frac{d}{dt}\Big(\frac{1}{P}\sum_{i\in[k]_m} q_i^R(t)\Big) = 0, \quad a.e.\ t \in [t_m, t_{m+1}).$$

Proof. We sum up $(2.6)_2$ over all i in the batch $[k]_m$ to get

$$\frac{d}{dt}\Big(\frac{1}{P}\sum_{i\in[k]_m} q_i^R(t)\Big) = \frac{\kappa}{P(P-1)}\sum_{i,j\in[k]_m}\Big(\Gamma(q_j^R(t) - q_i^R(t)) - \Gamma(q_j^\infty - q_i^\infty)\Big) = 0,$$

where the second equality is due to the anti-symmetry of the summands in the interchange $i \leftrightarrow j$. □

2.2. **Existence of equilibrium.** In this subsection, we recall the existence of an equilibrium for (1.1) using the contraction mapping principle, and discuss the RBM approximation from the view point of switching network topology. Next, under suitable assumptions, we show that the existence of such Q^∞ can be guaranteed, if coupling strength is large enough compared to the relative size of natural velocities.

Proposition 2.1. [28] *Suppose that condition* (1.3) *holds, and*

$$\sum_{j=1}^N \nu_j = 0, \qquad \kappa > \frac{\sqrt{2}}{r_0 L_1}\left(\frac{N-1}{N}\right)\|\nu\|_2.$$

Then, system (2.5) *has a unique solution $Q^\infty = (q_1^\infty, \ldots, q_N^\infty)$ in $[-r_0, r_0]^N$ such that*

$$\sum_{j=1}^N q_j^\infty = 0, \qquad \|Q^\infty\|_2 < \frac{r_0}{\sqrt{2}}, \qquad \mathcal{D}(Q^\infty) < r_0.$$

Proof. A proof can be found in Corollary 4.1 in [28]. □

Without the zero sum condition (2.5), there are infinitely many equilibria in addition to Q^∞ from translation invariance. The conservation law in Lemma 2.1 guarantees that Q^∞ is the only equilibrium the solution may converges to. The following remark shows that the solution exponentially converges.

Remark 2.2. [18] *The equilibrium $Q^\infty = (q_1^\infty, \cdots, q_N^\infty)$ obtained in Proposition 2.1 emerges from the dynamic solution Q to* (1.1) *as an asymptotic state, and it satisfies*

$$\left(\frac{1}{N}\sum_{i=1}^N |q_i(t) - q_i^\infty|^2\right) \le \exp\left(-\frac{2\kappa L_1 N}{N-1}t\right)\left(\frac{1}{N}\sum_{i=1}^N |q_i(0) - q_i^\infty|^2\right). \tag{2.7}$$

The proof is also stated later in Lemma 4.2 (if we put $P = N$).

2.3. **The RBM approximation.** Consider the Cauchy problem to the RBM approximated system:

$$\begin{cases} \dfrac{dq_i^R}{dt} = \dfrac{\kappa}{P-1}\displaystyle\sum_{j\in[i]_m}\Big(\Gamma(q_j^R - q_i^R) - \Gamma(q_j^\infty - q_i^\infty)\Big), \quad t \in [t_m, t_{m+1}), \\ q_i^R(0) = q_i^{in}, \quad i = 1, \cdots, N,\ m = 0, 1, 2, \cdots. \end{cases} \tag{2.8}$$

For $P = N$, note that one can easily check that system (2.8) reduces to the full batch model (1.1).

Our interest lies in the emergent behavior of the collective behavior of system (2.8). The error analysis in [28] suggests that the solutions $Q(t)$ and $Q^R(t)$ to (1.1) and (2.8), respectively, differ by the following relation:

$$\sup_{0\le t<\infty} \mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^N |q_i^R(t) - q_i(t)|^2\Big) \lesssim \Big[\tau\left(\frac{1}{P-1} - \frac{1}{N-1}\right) + \tau^2\Big]. \tag{2.9}$$

If we combine the decay of system (2.7) toward the equilibrium Q^∞ and the error estimate (2.9) by a triangle inequality, one may conclude that the system goes near the equilibrium exponentially fast, however, up to the error in (2.9). This error analysis is not enough to estimate the convergence speed of $Q^R(t)$ to the equilibrium.

As we described in (1.4), the RBM approximates the original system (1.3) by using coarser interacting network, ignoring the interactions between different batches. At $t = t_m$, the connectivity matrix $\chi(t_m) = (\chi_{ij}(t_m))_{i,j}$ from the random batches can be defined by

$$\chi_{ij}(t_m) := \begin{cases} 1 & \text{if} \quad j \in [i]_m, \\ 0 & \text{if} \quad j \notin [i]_m. \end{cases} \tag{2.10}$$

Since this is used for $t \in [t_m, t_{m+1})$, we set $\chi_{ij} = \chi_{ij}(t)$ to be the time-dependent piecewise-constant network connectivity:

$$\chi_{ij}(t) = \chi_{ij}(t_m) \quad \text{for} \quad t \in [t_m, t_{m+1}).$$

Then, the RBM approximation (2.8) can be rewritten as

$$\begin{cases} \dfrac{dq_i^R}{dt} = \dfrac{\kappa}{P-1}\displaystyle\sum_{j=1}^N \chi_{ij}\Big(\Gamma(q_j^R - q_i^R) - \Gamma(q_j^\infty - q_i^\infty)\Big), \quad t > 0, \\ q_i^R(0) = q_i^{in}, \quad i = 1, \cdots, N. \end{cases} \tag{2.11}$$

In this formulation, it is more clear that formally system (2.11) converges to system (1.3): the interaction network coefficient $\chi_{ij}(t)$ is averaged over time and becomes a uniform constant:

$$\frac{1}{M}\sum_{m=0}^{M-1} \frac{\chi_{ij}(t_m)}{P-1} \to \frac{1}{N-1} \quad \text{as} \;\; M \to \infty.$$

3. A framework and description of main results

In this section, we present a framework for convergence and state main results on the almost sure convergence and mean convergence toward to equilibrium.

3.1. **A framework.** Let $Q^\infty = \{q_i^\infty\}$ and $Q^R = \{q_i^R\}$ be the equilibrium and RBM solution, respectively. Then, we set the deviations of Q^R from Q^∞ by $\widehat{Q}^R = \{\hat{q}_i^R\}$:

$$\hat{q}_i(t) := q_i^R(t) - q_i^\infty \quad i = 1, \cdots, N. \tag{3.1}$$

Then, the fluctuation $\hat{Q}^R$ satisfies

$$\frac{d\hat{q}_i^R(t)}{dt} = \frac{\kappa}{P-1}\sum_{j=1}^N \chi_{ij}(t)(\Gamma(q_j^R(t) - q_i^R(t)) - \Gamma(q_j^\infty - q_i^\infty)), \quad t > 0, \tag{3.2}$$

or equivalently,

$$\frac{d\hat{q}_i^R(t)}{dt} = \frac{\kappa}{P-1}\sum_{j=1}^N \chi_{ij}(t)\gamma_{ij}(t)(\hat{q}_j^R(t) - \hat{q}_i^R(t)), \quad t > 0, \tag{3.3}$$

where $\chi_{ij}(t)$ is defined in (2.10) and we also define the nonlinear coefficient:

$$\gamma_{ij}(t) := \begin{cases} \dfrac{\Gamma(q_j^R(t) - q_i^R(t)) - \Gamma(q_j^\infty - q_i^\infty)}{(q_j^R(t) - q_i^R(t)) - (q_j^\infty - q_i^\infty)}, & \text{if } \ (q_j^R(t) - q_i^R(t)) \neq (q_j^\infty - q_i^\infty), \\ L_1, & \text{otherwise.} \end{cases} \tag{3.4}$$

Next, we list conditions imposed on natural velocities, coupling strength, initial data and coupling function:

- $(\mathcal{A}_1)$: intrinsic velocities and initial states have mean zero:
$$\sum_{j=1}^N \nu_j = 0, \qquad \sum_{j=1}^N q_j^{in} = 0.$$
- $(\mathcal{A}_2)$: Γ is bi-Lipschitz on $[-r_0, r_0]$: there exists positive constants L_1 and L_2 such that
$$L_1|p-q|^2 \le (\Gamma(p) - \Gamma(q))\cdot(p-q) \le L_2|p-q|^2, \quad \forall\ p, q \in [-r_0, r_0].$$
- $(\mathcal{A}_3)$: κ is sufficiently large such that
$$\kappa > \frac{\sqrt{2}}{r_0 L_1}\left(\frac{N-1}{N}\right)\|\nu\|_2.$$

Note that by the bi-Lipshitz condition $(\mathcal{A}_2)$, the linearized coefficient $\gamma_{ij}(t)$ in (3.4) is bounded:

$$\gamma_{ij}(t) = \gamma_{ji}(t) \in [L_1, L_2] \quad \text{if } \ |q_i^R - q_j^R| \le r_0 \quad \text{for all } \ i, j = 1, \dots, N. \tag{3.5}$$

3.2. **Two main results.** Next, we present the statements of two main results. Our first main result is concerned with the decay of the averaged variances of fluctuations $\widehat{Q}^R$.

Theorem 3.1. *Suppose assumptions $(\mathcal{A}_1) - (\mathcal{A}_3)$ hold, and let Q^R and Q^∞ be a solution of (1.5) and (2.3), respectively. Then, the averaged ℓ^2-distance from Q^R to Q^∞ decays to zero exponentially fast: there exists a positive constant Λ_1 such that*

$$\mathbb{E}\Big(\frac{1}{N}\sum_{j=1}^N |q_j^R(t) - q_j^\infty|^2\Big) \le e^{-\Lambda_1 t}\mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^N |q_i^R(0) - q_i^\infty|^2\Big), \quad t \ge 0,$$

where $\Lambda_1 = \Lambda_1(N, P, \tau, \kappa, L_1)$ *is explicitly given as follows.*

$$\Lambda_1 := -\left(\frac{P-1}{N-1}\right)\left(1 + \frac{2\kappa L_1 P}{P-1}\tau\right)^{-1}\frac{2\kappa L_1 P}{P-1}.$$

Proof. Since the proof is rather lengthy, we postpone its detailed proof in Section 4. The proof is motivated by the argument in [18], which treats the asymptotic flocking behavior of the RBM Cucker-Smale system by estimating the decay of $|\hat{q}_j(t) - \hat{q}_i(t)|^2$ when $[i]_m = [j]_m$. In [18], the model could be seen as a homogeneous linear system with a time-varying network topology. In order to apply the same argument in [18] to the RBM approximation of the model (1.1), we had to find an appropriate RBM formulation for (1.1), so that it resembles a homogeneous linear system. We dealt with this problem in Section 2.1 by adopting the RBM model introduced in [28] using a priori knowledge on the equilibria of system. □

Remark 3.1. *Although the decay exponent* Λ_1 *is proportional to* $(P-1)/(N-1)$*, this does not imply that the convergence of the RBM is slow as a numerical scheme. For a given tolerance level* $\varepsilon > 0$*, we may choose* τ *and* P *to make the error bound of* (2.9) *(which measures distance between the original and the RBM approximated solutions) smaller than* $\varepsilon/2$*. Since the original consensus model follows Remark 2.2, its* ℓ^2*-distance toward the equilibrium becomes less than* $\varepsilon/2$ *exponentially fast. Hence, the RBM approximated solution approaches to the equilibrium in an* ε *bound with the same decay rate as in the original consensus model. Theorem 3.1 explains the behavior below the* ε *bound; even after it becomes close to the equilibrium, it exponentially converges to the equilibrium, with a possible loss of decay rate.*

Our second main result deals with the almost-sure convergence of Q^R toward Q^∞ using proper stopping times.

Theorem 3.2. *Let* Q^R *and* Q^∞ *be solutions of* (1.5) *and* (2.3)*, respectively. Then, the diameter process* $\mathcal{D}(\widehat{Q})$ *tends to zero exponentially fast almost surely. More precisely, for* Λ_2 *satisfying*

$$0 < \Lambda_2 < \frac{\kappa L_1 \exp(-\kappa(L_1/(P-1) + L_2)\tau\mathbb{E}(b_1))}{(P-1)\mathbb{E}(b_1)},$$

there exists a positive constant $C > 0$ *such that*

$$\mathcal{D}(\widehat{Q}^R(t)) \le Ce^{-\Lambda_2 t}, \quad a.s. \quad t \ge 0,$$

where C *is a constant depending on the sample path, and* $\mathbb{E}(b_1)$ *appearing in the upper bound of* Λ_2 *is a constant depending only on* P *and* N *which can be estimated by*

$$1 \le \mathbb{E}(b_1) \le (N+1)\frac{N-P}{P-1} + 1 = \mathcal{O}(N^2).$$

Proof. We leave its detailed proof to Section 5 which adopts a strategy in [14], but its statement and proof are more elegant than that of a result in [14]. The random network structure in RBM, which is a special case of the randomly switching graphs in [14], allows us to devise a shorter and simpler way to derive a desired asymptotic behavior. Unlike [14], we proved that the emergent behavior emerges exponentially fast. □

Remark 3.2. *A practical use of RBM requires that the batch size is small* ($P \ll N$) *and the time step is short* ($\tau \ll 1$)*. The decay exponents* Λ_1 *and* Λ_2 *can be roughly estimated as follows:*

$$\Lambda_1 = \mathcal{O}\left(\frac{\kappa P}{N(1+\kappa\tau)}\right) \quad and \quad \Lambda_2 = \mathcal{O}\left(\frac{\kappa P}{N^2}\exp\left(-\frac{\kappa P}{N^2}\tau\right)\right).$$

In contrast, the original system has a decay rate of $\mathcal{O}(\kappa)$ independent of N from Remark 2.2.

4. Zero convergence of the variance of fluctuations

In this section, we present a proof of Theorem 3.1. First, we briefly delineate our proof strategy in several steps:

- Step A: recall that $\widehat{Q}^R = \{\hat{q}_i^R\}$ with $\hat{q}_i^R = q_i^R - q_i^\infty$ for $i = 1, \cdots, N$,
$$\mathcal{D}(\widehat{Q}^R) := \max_{i,j} |\hat{q}_i^R - \hat{q}_j^R| \quad \text{and} \quad \mathcal{D}(Q^\infty) := \max_{i,j} |q_i^\infty - q_j^\infty|.$$
Then, we show an existence of positively invariant set (see Lemma 4.1):
$$\mathcal{D}(\widehat{Q}^R(0)) + \mathcal{D}(Q^\infty) < r_0 \quad \Longrightarrow \quad \sup_{0\le t<\infty} \mathcal{D}(\widehat{Q}^R(t)) + \mathcal{D}(Q^\infty) < r_0.$$
- Step B: we derive a dissipative estimate (see Lemma 4.2):
$$\begin{aligned}&\frac{d}{dt}\Big(\frac{1}{P^2}\sum_{i\in[k]_m}\sum_{j\in[l]_m}|\hat{q}_i^R - \hat{q}_j^R|^2\Big)\\ &\qquad = -\frac{\kappa L_1 P}{P-1}\left(\frac{1}{P^2}\sum_{i,j\in[k]_m}|\hat{q}_i^R - \hat{q}_j^R|^2 + \frac{1}{P^2}\sum_{i,j\in[l]_m}|\hat{q}_i^R - \hat{q}_j^R|^2\right) \le 0.\end{aligned}$$
- Step C: for a sequence of sampling times $\{t_m = m\tau\}$, we derive a recursive relation (see Lemma 4.3):
$$\begin{aligned}&\mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^N |q_i^R(t_{m+1}) - q_i^\infty|^2\Big)\\ &\qquad \le \left[\frac{N-P}{N-1} + \frac{P-1}{N-1}\exp\Big(-\frac{2\kappa L_1 P}{P-1}\tau\Big)\right]\mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^N |q_j^R(t_m) - q_i^\infty|^2\Big).\end{aligned}$$
- Final step: We use the recursive relation in Step C and non-increasing property of $\mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^N |q_i^R(t_m) - q_i^\infty|^2\Big)$ to finish the proof of Theorem 3.1.

Detailed proofs for the above four steps will be given in the following subsection.

4.1. Elementary estimates.

From (3.5), again note that the nonlinear coefficient γ_{ij} in (3.4) is symmetric and bounded when $|q_i^R - q_j^R| \le r_0$ for all i, j. Hence, we first need to show the boundedness of Q^R in the whole time interval. The boundedness property is proved in [28] as in the following lemma.

Lemma 4.1. *Suppose that the condition $(\mathcal{A}_2)$ holds and initial data satisfy*
$$\mathcal{D}(\widehat{Q}^R(0)) + \mathcal{D}(Q^\infty) < r_0, \tag{4.1}$$
and let $Q^R(t)$ be a solution to system (1.5). Then, one has
$$\sup_{0\le t<\infty} \mathcal{D}(\widehat{Q}^R(t)) + \mathcal{D}(Q^\infty) < r_0.$$

Proof. The proof can be followed from the same argument in [28]. For reader's convenience, we briefly sketch a proof below.

• Step A: Suppose that the boundedness condition holds for a time interval $[0, T)$:

$$\mathcal{D}(\widehat{Q}^R(t)) + \mathcal{D}(Q^\infty) < r_0, \quad t \in [0, T). \tag{4.2}$$

Then, we claim:

$$\mathcal{D}(\widehat{Q}^R(t)) \le \mathcal{D}(\widehat{Q}^R(0)), \quad t \in [0, T].$$

First, we choose time-dependent indices i_t and j_t such that

$$\hat{q}_{i_t}(t) = \max_{1 \le i \le N} \hat{q}_i(t), \quad \hat{q}_{j_t}(t) = \min_{1 \le i \le N} \hat{q}_i(t). \tag{4.3}$$

◇ Case A (Non-increasing property of $\hat{q}_{i_t}$): It follows from (3.2) that

$$\frac{d\hat{q}^R_{i_t}}{dt} = \frac{\kappa}{P-1} \sum_{k \in [i_t]_m} \Big(\Gamma(q^R_k - q^R_{i_t}) - \Gamma(q^\infty_k - q^\infty_{i_t}) \Big) \le \frac{L_1 \kappa}{P-1} \sum_{k \in [i_t]_m} (\hat{q}^R_k - \hat{q}^R_{i_t}) \le 0,$$

where we use (4.3) to see the following relation:

$$(q^R_k - q^R_{i_t}) - (q^\infty_k - q^\infty_{i_t}) = \hat{q}^R_k - \hat{q}^R_{i_t} \le 0, \quad k \in \{1, \cdots, N\}.$$

◇ Case B (Non-decreasing property of $\hat{q}_{j_t}$): By the same argument as in Case A, one has

$$\frac{d\hat{q}^R_{j_t}}{dt} \ge 0.$$

Finally, we combine estimates for Case A and Case B to get

$$\mathcal{D}(\widehat{Q}^R(t)) \le \mathcal{D}(\widehat{Q}^R(0)).$$

• Step B: By the continuous induction argument used in [28], we show that the a priori condition (4.2) holds for the whole time $t \ge 0$, since (4.2) is true at $t = 0$ from (4.1). For the details, we refer to [28]. □

Note that for a time interval $[t_m, t_{m+1})$, each connected component has all-to-all network structure, independent of other batches. Next, we estimate the following quantity:

$$\frac{1}{P^2} \sum_{i \in [k]_m} \sum_{j \in [l]_m} |\hat{q}^R_i - \hat{q}^R_j|^2.$$

Lemma 4.2. *Let Q^R and Q^∞ be a solution of* (1.5) *and* (2.3)*, respectively, and let $k, l = 1, \ldots, N$. Then for $t \in [t_m, t_{m+1})$, we have*

$$\begin{aligned} &\frac{d}{dt}\Big(\frac{1}{P^2} \sum_{i \in [k]_m} \sum_{j \in [l]_m} |\hat{q}^R_i - \hat{q}^R_j|^2 \Big) \\ &\quad = -\frac{\kappa L_1 P}{P-1} \left(\frac{1}{P^2} \sum_{i,j \in [k]_m} |\hat{q}^R_i - \hat{q}^R_j|^2 + \frac{1}{P^2} \sum_{i,j \in [l]_m} |\hat{q}^R_i - \hat{q}^R_j|^2 \right) \le 0. \end{aligned} \tag{4.4}$$

Proof. Note that

$$
\begin{aligned}
&\frac{d}{dt}\Big(\frac{1}{P^2}\sum_{i\in[k]_m}\sum_{j\in[l]_m}|\hat{q}_i^R-\hat{q}_j^R|^2\Big)\\
&=\frac{2}{P^2}\sum_{i\in[k]_m}\sum_{j\in[l]_m}(\hat{q}_i^R-\hat{q}_j^R)\cdot\Big(\frac{d\hat{q}_i^R}{dt}-\frac{d\hat{q}_j^R}{dt}\Big)=\frac{2}{P^2}\sum_{i\in[k]_m}\sum_{j\in[l]_m}\Big(\hat{q}_i^R\cdot\frac{d\hat{q}_i^R}{dt}+\hat{q}_j^R\cdot\frac{d\hat{q}_j^R}{dt}\Big)\\
&=\frac{d}{dt}\Big(\frac{1}{P}\sum_{i\in[k]_m}|\hat{q}_i^R|^2\Big)+\frac{d}{dt}\Big(\frac{1}{P}\sum_{j\in[l]_m}|\hat{q}_j^R|^2\Big),
\end{aligned}
\tag{4.5}
$$

where we used Lemma 2.2:

$$
\frac{d}{dt}\Big(\frac{1}{P}\sum_{i\in[k]_m}\hat{q}_i^R\Big)=\frac{1}{P}\sum_{i\in[k]_m}\frac{d\hat{q}_i^R}{dt}=0.
$$

Now we estimate the last terms in (4.5) as follows. It follows from (1.5) that

$$
\begin{aligned}
\frac{d}{dt}\Big(\frac{1}{P}\sum_{i\in[k]_m}|\hat{q}_i^R|^2\Big)&=\frac{1}{P}\sum_{i\in[k]_m}2\hat{q}_i^R\cdot\frac{d\hat{q}_i^R}{dt}\\
&=\frac{1}{P}\sum_{i\in[k]_m}2\hat{q}_i^R\cdot\frac{\kappa}{P-1}\sum_{j\in[k]_m}\Big(\Gamma(q_j^R-q_i^R)-\Gamma(q_j^\infty-q_i^\infty)\Big)\\
&=\frac{\kappa}{P(P-1)}\sum_{i,j\in[k]_m}2\hat{q}_i^R\cdot\Big(\Gamma(q_j^R-q_i^R)-\Gamma(q_j^\infty-q_i^\infty)\Big)\\
&=\frac{\kappa}{P(P-1)}\sum_{i,j\in[k]_m}\underbrace{(\hat{q}_i^R-\hat{q}_j^R)\cdot\Big(\Gamma(q_j^R-q_i^R)-\Gamma(q_j^\infty-q_i^\infty)\Big)}_{=:\Delta_{ij}}.
\end{aligned}
\tag{4.6}
$$

The term Δ_{ij} can be estimated using the assumption $(\mathcal{A}_2)$:

$$
\begin{aligned}
\Delta_{ij}&=(\hat{q}_i^R-\hat{q}_j^R)\cdot(\Gamma(q_j^R-q_i^R)-\Gamma(q_j^\infty-q_i^\infty))\\
&=-((q_j^R-q_i^R)-(q_j^\infty-q_i^\infty))\cdot(\Gamma(q_j^R-q_i^R)-\Gamma(q_j^\infty-q_i^\infty))\\
&\le -L_1|\hat{q}_i^R-\hat{q}_j^R|^2.
\end{aligned}
\tag{4.7}
$$

Finally, we combine (4.5), (4.6) and (4.7) to get (4.4). □

Remark 4.1. *Note that* (4.4) *yields dissipative estimates: for* $t\in[t_m,t_{m+1})$,

$$
\begin{aligned}
&\frac{d}{dt}\Big(\frac{1}{P^2}\sum_{i\in[k]_m}\sum_{j\in[l]_m}|\hat{q}_i^R-\hat{q}_j^R|^2\Big)\le 0,\\
&\frac{d}{dt}\Big(\frac{1}{P^2}\sum_{i,j\in[k]_m}|\hat{q}_i^R-\hat{q}_j^R|^2\Big)\le -\frac{2\kappa L_1 P}{P-1}\Big(\frac{1}{P^2}\sum_{i,j\in[k]_m}|\hat{q}_i^R-\hat{q}_j^R|^2\Big).
\end{aligned}
\tag{4.8}
$$

Consider the random second moment process:

$$
\mathcal{M}_2:=\frac{1}{N^2}\sum_{i,j=1}^N|\hat{q}_j^R(t)-\hat{q}_i^R(t)|^2,\quad t\ge 0.
\tag{4.9}
$$

Then, it follows from (2.5), Lemma 2.2 and (3.1) that $\sum_{i=1}^N \hat{q}_i^R = 0$. Thus, the above second moment (4.9) is indeed the quantity proportional to the second moment of $q_i^R - q_i^\infty$:

$$\mathcal{M}_2 = \frac{1}{N^2}\sum_{i,j=1}^N |\hat{q}_j^R(t) - \hat{q}_i^R(t)|^2 = \frac{2}{N}\sum_{i=1}^N |\hat{q}_i^R|^2 - 2\Big|\frac{1}{N}\sum_{i=1}^N \hat{q}_i^R\Big|^2 = \frac{2}{N}\sum_{i=1}^N |\hat{q}_i^R|^2.$$

Remark 4.2. [18] *If the random vector $\widehat{Q}(t)$ for $t \in [t_{m-1}, t_m)$ was independent of the batch $[i]_m$, then we would have obtained the following simple estimate on $\mathbb{E}(\mathcal{M}_2)$ by summing* (4.8) *over k and then taking expectation on it:*

$$\frac{d}{dt}\mathbb{E}\Big(\sum_{1\le i,j\le N} |\hat{q}_i^R - \hat{q}_j^R|^2\Big) \le -\frac{2\kappa L_1 P}{P-1}\mathbb{E}\Big(\sum_{1\le i,j\le N} |\hat{q}_i^R - \hat{q}_j^R|^2\Big).$$

However, the independence is only guaranteed at $t = t_m$.

In the following lemma, we show the decay of the expectation of the random process $\mathcal{M}_2$.

Lemma 4.3. *For a positive $\tau > 0$, let $\{t_m = m\tau\}$ be a sequence of discrete times. Then, for $t \in [t_m, t_{m+1}]$, one has*

$$\mathbb{E}(\mathcal{M}_2(t)) \le \left[\frac{N-P}{N-1} + \frac{P-1}{N-1}\exp\Big(-\frac{2\kappa L_1 P}{P-1}\tau\Big)\right]^m \mathbb{E}(\mathcal{M}_2(0)). \tag{4.10}$$

Proof. We split its proof into two steps.

• Step A (Derivation of a recursive relation): we claim

$$\mathbb{E}(\mathcal{M}_2(t)) \le \left[\frac{N-P}{N-1} + \frac{P-1}{N-1}\exp\Big(-\frac{2\kappa L_1 P}{P-1}(t-t_m)\Big)\right] \mathbb{E}(\mathcal{M}_2(t_m)). \tag{4.11}$$

Proof of (4.11): for $t \in [t_m, t_{m+1})$, we use (4.8) to obtain

$$\begin{aligned}
\mathcal{M}_2(t) &= \frac{1}{N^2}\sum_{i,j=1}^N |\hat{q}_j^R(t) - \hat{q}_i^R(t)|^2 \\
&= \frac{1}{N^2}\sum_{\substack{1\le i,j\le N\\ [i]_m\neq[j]_m}} |\hat{q}_j^R(t) - \hat{q}_i^R(t)|^2 + \frac{1}{N^2}\sum_{\substack{1\le i,j\le N\\ [i]_m=[j]_m}} |\hat{q}_j^R(t) - \hat{q}_i^R(t)|^2 \\
&\le \frac{1}{N^2}\sum_{\substack{1\le i,j\le N\\ [i]_m\neq[j]_m}} |\hat{q}_j^R(t_m) - \hat{q}_i^R(t_m)|^2 \\
&\quad + \frac{1}{N^2}\sum_{\substack{1\le i,j\le N\\ [i]_m=[j]_m}} |\hat{q}_j^R(t_m) - \hat{q}_i^R(t_m)|^2 \exp\Big[-\frac{2\kappa L_1 P}{P-1}(t-t_m)\Big].
\end{aligned} \tag{4.12}$$

Now, we take the expectation to both sides of (4.12) to get

$$
\begin{aligned}
\mathbb{E}(\mathcal{M}_2(t)) &\le \mathbb{E}\Big(\frac{1}{N^2}\sum_{\substack{1\le i,j\le N\\ [i]_m\neq[j]_m}} |\hat{q}_j^R(t_m)-\hat{q}_i^R(t_m)|^2\Big)\\
&\quad + \mathbb{E}\Big(\frac{1}{N^2}\sum_{\substack{1\le i,j\le N\\ [i]_m=[j]_m}} |\hat{q}_j^R(t_m)-\hat{q}_i^R(t_m)|^2\Big)\exp\Big[-\frac{2\kappa L_1 P}{P-1}(t-t_m)\Big]\\
&=: \mathcal{I}_{11}+\mathcal{I}_{12}\exp\Big[-\frac{2\kappa L_1 P}{P-1}(t-t_m)\Big].
\end{aligned}
\tag{4.13}
$$

• (Estimate of $\mathcal{I}_{12}$): At each time discretization $t\in[t_m,t_{m+1})$, the random connectivity process χ_{ij} in (2.10)

$$
\chi_{ij}(t) := \begin{cases} 1 & \text{if} \quad j\in[i]_m,\\ 0 & \text{if} \quad j\notin[i]_m,\end{cases}
$$

has the expectation of $(P-1)/(N-1)$ when $i\neq j$:

$$
\mathbb{E}(\chi_{ij}(t)) = \mathbb{P}\{[i]_m=[j]_m \mid i\neq j\} = \frac{P-1}{N-1}. \tag{4.14}
$$

At each t_m, the random variables $\chi_{ij}(t_m)$ and $|\hat{q}_j^R(t_m)-\hat{q}_i^R(t_m)|^2$ are independent. Hence, we use (4.14) and properties of expectation to find

$$
\begin{aligned}
&\mathbb{E}\Big(\frac{1}{N^2}\sum_{\substack{1\le i,j\le N\\ [i]_m=[j]_m}} |\hat{q}_j^R(t_m)-\hat{q}_i^R(t_m)|^2\Big)\\
&\quad = \mathbb{E}\Big(\frac{1}{N^2}\sum_{1\le i,j\le N} \chi_{ij}(t_m)\cdot|\hat{q}_j^R(t_m)-\hat{q}_i^R(t_m)|^2\Big)\\
&\quad = \frac{1}{N^2}\sum_{1\le i,j\le N} \mathbb{E}\Big(\chi_{ij}(t_m)\cdot|\hat{q}_j^R(t_m)-\hat{q}_i^R(t_m)|^2\Big)\\
&\quad = \frac{1}{N^2}\sum_{1\le i,j\le N} \mathbb{E}\Big(\chi_{ij}(t_m)\Big)\mathbb{E}\Big(|\hat{q}_j^R(t_m)-\hat{q}_i^R(t_m)|^2\Big)\\
&\quad = \frac{P-1}{N-1}\frac{1}{N^2}\sum_{1\le i,j\le N} \mathbb{E}\Big(|\hat{q}_j^R(t_m)-\hat{q}_i^R(t_m)|^2\Big)\\
&\quad = \frac{P-1}{N-1}\mathbb{E}(\mathcal{M}_2(t_m)).
\end{aligned}
\tag{4.15}
$$

• (Estimate of $\mathcal{I}_{11}$): Similar to the estimate for $\mathcal{I}_{12}$, one has

$$
\mathbb{E}\Big(\frac{1}{N^2}\sum_{\substack{1\le i,j\le N\\ [i]_m\neq[j]_m}} |\hat{q}_j^R(t_m)-\hat{q}_i^R(t_m)|^2\Big) = \frac{N-P}{N-1}\mathbb{E}(\mathcal{M}_2(t_m)). \tag{4.16}
$$

In (4.13), for $t\in[t_m,t_{m+1}]$, we combine (4.15) and (4.16) to get (4.11):

$$
\mathbb{E}(\mathcal{M}_2(t)) \le \Big[\frac{N-P}{N-1}+\frac{P-1}{N-1}\exp\Big(-\frac{2\kappa L_1 P}{P-1}(t-t_m)\Big)\Big]\mathbb{E}(\mathcal{M}_2(t_m)). \tag{4.17}
$$

• Step B (Iteration of a recursive relation): in (4.17), we use $t_m - t_{m-1} = \tau$ to get

$$\mathbb{E}[\mathcal{M}_2(t)] \le \mathbb{E}[\mathcal{M}_2(t_m)] \le \left[\frac{N-P}{N-1} + \frac{P-1}{N-1}\exp\Big(-\frac{2\kappa L_1 P}{P-1}\tau\Big)\right]\mathbb{E}[\mathcal{M}_2(t_{m-1})]. \tag{4.18}$$

We iterate the recursive relation (4.18) from t_m to $t_0 = 0$ backward to obtain the desired estimate (4.10). □

Next, we provide an elementary inequality from [18].

Lemma 4.4. [18] *Let $0 \le a \le 1$ and $b > 0$. Then, for $x \in [0, b]$,*

$$(1-a) + ae^{-x} \le \exp\left(-\frac{a}{1+b}x\right).$$

Next, we provide a proof of Theorem 3.1.

4.2. **Proof of Theorem 3.1.** It follows from (4.17) and Lemma 4.4 that for $t \in [t_m, t_{m+1}]$, we have

$$\begin{aligned}\mathbb{E}(\mathcal{M}_2(t)) &= \left[\frac{N-P}{N-1} + \frac{P-1}{N-1}\exp\Big(-\frac{2\kappa L_1 P}{P-1}(t-t_m)\Big)\right]\mathbb{E}(\mathcal{M}_2(t_m))\\ &\le \exp\left[-\left(\frac{P-1}{N-1}\right)\left(1+\frac{2\kappa L_1 P}{P-1}\tau\right)^{-1}\frac{2\kappa L_1 P}{P-1}(t-t_m)\right]\mathbb{E}(\mathcal{M}_2(t_m)).\end{aligned} \tag{4.19}$$

Next, as in Lemma 4.3, we iterate the above relation (4.19) from t_m to $t_0 = 0$ to derive

$$\mathbb{E}(\mathcal{M}_2(t)) \le \exp\left[-\left(\frac{P-1}{N-1}\right)\left(1+\frac{2\kappa L_1 P}{P-1}\tau\right)^{-1}\frac{2\kappa L_1 P}{P-1}t\right]\mathbb{E}(\mathcal{M}_2(0)).$$

5. Almost-sure zero convergence of the diameter process

In this section, we deduce a proof of Theorem 3.2. Note that taking expectations is crucial in (4.17) to integrate the decay on each batch to get an exponential decay of variation of $\widehat{Q}^R$. To prove the convergence toward zero on the dynamics of the diameter process of $\widehat{Q}^R$, we use several properties of the connection topology.

First, we rewrite (3.3)

$$\frac{d\hat{q}_i^R}{dt} = \frac{\kappa}{P-1}\sum_{j=1}^N \chi_{ij}\gamma_{ij}(\hat{q}_j^R - \hat{q}_i^R), \quad t > 0, \tag{5.1}$$

in a matrix form:

$$\frac{d\widehat{Q}^R}{dt} = -\frac{\kappa L(t)}{P-1}\widehat{Q}^R. \tag{5.2}$$

Here the matrix $L(t) = (L_{ij}(t))$ is given by

$$L_{ij}(t) := \begin{cases} -\chi_{ij}(t)\gamma_{ij}(t), & i \ne j,\\ \displaystyle\sum_{k\ne i}\chi_{ik}(t)\gamma_{ik}(t), & i = j. \end{cases} \tag{5.3}$$

Let $\Phi(t,\tau)$ be the state transition matrix associated with (5.2) on the time interval $[\tau, t]$:

$$\widehat{Q}^R(t) = \Phi(t,\tau)\widehat{Q}^R(\tau). \tag{5.4}$$

First, we briefly discuss basic ideas for a proof of Theorem 3.2 in two steps:

- Step A (Derivation of recursive relation for $\mathcal{D}(\widehat{Q}^R(t))$): recall that

$$\widehat{Q}^R = \{q_i^R - q_i^\infty\}, \quad \mathcal{D}(\widehat{Q}^R) := \max_{1\le i,j\le N} |\hat{q}_i^R - \hat{q}_j^R|.$$

  We choose a random sequence:

$$t_n^* := t_{b_n} = b_n \tau \tag{5.5}$$

  which is a subsequence of $\{t_n\}_{n=0}^\infty$, where the random integer sequence $\{b_n\}_{n=0}^\infty$ will be determined later. If the network connectivity is mixed enough (in the sense of the ergodicity in Definition 5.1 and Definition 5.2 below), we show that

$$\mathcal{D}(\widehat{Q}^R(t_{n+1}^*)) \le (1-\varepsilon)\mathcal{D}(\widehat{Q}^R(t_n^*)), \quad n = 1, 2, \cdots, \tag{5.6}$$

  for some constant $0 < \varepsilon < 1$.

- Step B (Derivation of decay estimate for $\mathcal{D}(\widehat{Q}^R(t))$): by iterating the recursive relation (5.6), one obtains

$$\mathcal{D}(\widehat{Q}^R(t_n^*)) \le (1-\varepsilon)^n \mathcal{D}(\widehat{Q}^R(0)) \le e^{-\varepsilon n}\mathcal{D}(\widehat{Q}^R(0)).$$

  This implies the decay of $\mathcal{D}(\widehat{Q}^R(t))$ as $t \to \infty$.

In what follows, we perform the above two steps one by one. The estimate (5.6) will be verified in the following two subsections.

5.1. **Scrambling matrix and ergodicity coefficient.** First, we begin with the several notions on stochastic matrix, scrambling matrix and ergodicity coefficient as follows.

Definition 5.1. *Let $A = (a_{ij})$ be an $N \times N$ square matrix with nonnegative entries.*

(1) A is stochastic, if the row sums are equal to unity:

$$\sum_{j=1}^N a_{ij} = 1, \quad \textit{for all } i = 1, \cdots, N.$$

(2) A is scrambling, if for any pair of rows (i, j), there is a column k such that

$$a_{ik} > 0 \quad \textit{and} \quad a_{jk} > 0. \tag{5.7}$$

(3) The ergodicity coefficient of A is

$$\mu(A) := \min_{i,j} \sum_{k=1}^N \min\{a_{ik}, a_{jk}\}.$$

Remark 5.1. *1. By the nonnegative of entries in a scrambling matrix, it is easy to see that a scrambling matrix has a positive ergodicity coefficient.*
2. For two nonnegative matrices $A = (a_{ij})$ and $B = (b_{ij})$,

$$a_{ij} \ge b_{ij} \textit{ for any } i, j \quad \Longrightarrow \quad \mu(A) \ge \mu(B).$$

Lemma 5.1. [3] *Suppose that a nonnegative $N \times N$ matrix $A = (a_{ij})$ is stochastic, and let $Z = (z_1, \dots, z_N)^\top$ and $W = (w_1, \dots, w_N)^\top$ be vectors such that*

$$W = AZ.$$

Then, we have

$$\max_{1 \le i,j \le N} |w_i - w_j| \le (1 - \mu(A)) \max_{1 \le i,j \le N} |z_i - z_j|.$$

Next, we estimates on the time sequence t_n^* appearing in (5.6) in Step A. From the arguments in Definition 5.1 and Lemma 5.1, we would like to show that the state transition matrix $\Phi(t_{n+1}^*, t_n^*)$ is a scrambling matrix for $|t_{n+1}^* - t_n^*| \gg 1$. To achieve this, we define the sequence $\{b_n\}_{n=0}^\infty$ appearing in (5.5) inductively.

Definition 5.2. *We define the random integer sequence $\{b_n\}_{n=0}^\infty$ as follows.*

$$b_{n+1} := \begin{cases} 0, & \text{if } n = -1, \\ \min\left\{ b \ge b_n + 1 \ : \ \displaystyle\sum_{k=b_n}^{b-1} \chi(t_k) \text{ is scrambling} \right\} & \text{if } n \ge 0, \end{cases}$$

where $\chi = (\chi_{ij})$ is the adjacency matrix defined in (2.10)*.*

Remark 5.2. *An intuitive explanation for the sequence (b_n) appearing in Definition 5.2 is as follows. Let $\mathcal{G}(t)$ $(t \ge 0)$ be the graph of the network generated by $\chi_{ij}(t)$, where the set of vertices is $\{1, 2, \dots, N\}$, and the edge between i and j exists if and only if $\chi_{ij}(t) = 1$.*

Note that

$$\sum_{k=b_n}^{b_{n+1}-1} \chi(t_k) \text{ is scrambling}$$

$$\iff \text{the adjacency matrix of the graph } \bigcup_{k=b_n}^{b_{n+1}-1} \mathcal{G}(t_k) \text{ is scrambling}$$

$$\iff \text{the adjacency matrix of the graph } \bigcup_{t \in [t_{b_n}, t_{b_{n+1}})} \mathcal{G}(t) \text{ is scrambling.}$$

Therefore, one has

$$b_{n+1} = \min\left\{ b \ge b_n + 1 \ : \ \text{the adjacency matrix of } \bigcup_{t \in [t_{b_n}, t_b)} \mathcal{G}(t) \text{ is scrambling} \right\}.$$

In the following proposition, we show that t_n^* is a well-defined stopping time.

Proposition 5.1. *For any $n \in \mathbb{N}$, there exists $m \ge n$ such that the matrix $\sum_{k=n}^m \chi(t_k)$ is scrambling almost surely. In particular, $t_n^* < \infty$ for any n.*

Proof. Instead of checking (5.7) for $\sum_{k=n}^m \chi(t_k)$, we look for a much stronger condition that every vertex is directly connected to the vertex labeled by 1 at some time $t_k \in \{t_n, \dots, t_m\}$.

Note that

$$\begin{aligned}
&\mathbb{P}\left(\sum_{k=n}^{m}\chi(t_k) \text{ is not scrambling}\right)\\
&\quad \le \mathbb{P}\Big(\exists \text{ a vertex } j \text{ with } j\ge 2 \text{ such that } \chi_{1j}(t_k)=0 \text{ for all } k\in[n,m]\Big)\\
&\quad \le \sum_{j=2}^{N}\prod_{k=n}^{m}\mathbb{P}\left(\chi_{1j}(t_k)=0\right)=\sum_{j=2}^{N}\prod_{k=n}^{m}\left(\frac{N-P}{N-1}\right)=(N-1)\left(\frac{N-P}{N-1}\right)^{m-n+1},
\end{aligned}$$

i.e.,

$$\mathbb{P}\left(\sum_{k=n}^{m}\chi(t_k) \text{ is not scrambling}\right)\le (N-1)\left(\frac{N-P}{N-1}\right)^{m-n+1}. \tag{5.8}$$

Note that the R.H.S. of the above relation tends to 0 as $m\to\infty$. Therefore, for any $n\in\mathbb{N}$, the probability $\mathbb{P}\{\sum_{k=n}^{\infty}\chi(t_k)$ is not scrambling$\}$ is zero, which says that $m<\infty$ almost surely. □

In the next lemma, we estimate successive gaps in the length of $t_n^*=b_n\tau$. First, we set

$$a_n := \frac{b_n}{n},\quad n=1,2,\cdots.$$

Lemma 5.2. *Let $t_n^*=b_n\tau$ be the sequence of times as in Definition 5.2. Then, one has the following assertions.*

$$(i)\ a_n\ge 1,\quad \mathbb{E}(a_n)=\mathbb{E}(b_1),\quad n\in\mathbb{N},\quad \lim_{n\to\infty}a_n=\mathbb{E}(b_1)\quad a.s.$$

$$(ii)\ \mathbb{E}(b_1)\le (N-1)\frac{N-P}{P-1}+1.$$

Proof. (i) By definition 5.2 and Remark 5.2, one has

$$b_n\ge n\quad \text{i.e.,}\quad a_n=\frac{b_n}{n}\ge 1.$$

Note that the choices of batches at $t=t_m$ are independent of system dynamics in the time interval $[t_m,t_{m+1})$, the successive gaps $\{b_n-b_{n-1}\}_{n\ge1}$ are independent and identically distributed random variables. On the other hand, one has

$$a_n=\frac{b_n}{n}=\frac{1}{n}\sum_{m=1}^{n}(b_m-b_{m-1}). \tag{5.9}$$

We take an expectation to both sides of (5.9) to find

$$\mathbb{E}(a_n)=\frac{1}{n}\sum_{m=1}^{n}\mathbb{E}(b_m-b_{m-1}). \tag{5.10}$$

Since b_m-b_{m-1} is i.i.d. and $b_0=0$, we have

$$\mathbb{E}(b_m-b_{m-1})=\mathbb{E}(b_1-b_0)=\mathbb{E}(b_1). \tag{5.11}$$

By combining (5.10) and (5.11), one has

$$\mathbb{E}(a_n)=\mathbb{E}(b_1),\quad \text{for}\ \ n\ge 1.$$

Now, we apply the strong law of large numbers to $\{b_m - b_{m-1}\}$ to get

$$\exists \lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n}(b_m - b_{m-1}) = \mathbb{E}(b_1 - b_0) = \mathbb{E}(b_1), \quad \text{a.s.} \tag{5.12}$$

Finally, the last relation in (i) follows from (5.9) and (5.12).

(ii) We use definition of $\mathbb{E}(b_1)$ and

$$n = \sum_{m=1}^{\infty} \mathbf{1}_{\{m\le n\}}, \quad \mathbb{P}(b_1 \ge 1) = 1$$

to get

$$\begin{aligned}
\mathbb{E}(b_1) &= \sum_{n=1}^{\infty} n\mathbb{P}\big(b_1 = n\big) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \mathbf{1}_{\{m\le n\}}\mathbb{P}\big(b_1 = n\big) \\
&= \sum_{m=1}^{\infty} \mathbb{P}\big(b_1 \ge m\big) = 1 + \sum_{m=2}^{\infty} \mathbb{P}\big(b_1 \ge m\big) \\
&= 1 + \sum_{m=2}^{\infty} \mathbb{P}\left(\sum_{k=0}^{b-1}\chi(t_k) \text{ is not a scrambling matrix for } b < m\right) \qquad (5.13)\\
&= 1 + \sum_{m=2}^{\infty} \mathbb{P}\left(\sum_{k=0}^{m-2}\chi(t_k) \text{ is not a scrambling matrix}\right) \\
&= 1 + \sum_{m=0}^{\infty} \mathbb{P}\left(\sum_{k=0}^{m}\chi(t_k) \text{ is not a scrambling matrix}\right).
\end{aligned}$$

We combine (5.8) and (5.13) to obtain

$$\mathbb{E}(b_1) \le 1 + \sum_{m=0}^{\infty}(N-1)\left(\frac{N-P}{N-1}\right)^{m+1} = 1 + (N-1)\frac{N-P}{P-1}.$$

□

5.2. **Estimate on the ergodic coefficient.** Recall that we have chosen a random sequence $\{b_n\}$ such that

$$\sum_{k=b_n}^{b_{n+1}-1}\chi(t_k) \text{ is a scrambling matrix for all } n.$$

Next, we show that the ergodicity coefficient of $\Phi(t^*_{n+1}, t^*_n) = \Phi(\tau b_{n+1}, \tau b_n)$ is always positive.

Lemma 5.3. [Lemma 4.1 in [14]] *Suppose the coupling function Γ in (3.3)–(3.4) satisfies the bi-Lipschitz condition $(\mathcal{A}_2)$. Then, system (3.3) enjoys the following properties.*

(1) The ergodicity coefficient $\mu\left(\Phi(t^*_{n+1}, t^*_n)\right)$ has a positive lower bound for any n:

$$\mu\left(\Phi(t^*_{n+1}, t^*_n)\right) \ge \frac{\tau\kappa L_1}{P-1}\exp\Big[-\kappa\Big(\frac{L_1}{P-1} + L_2\Big)(t^*_{n+1} - t^*_n)\Big].$$

(2) The state transition matrix $\Phi(t^*_{n+1}, t^*_n)$ is stochastic for any n.

Proof. (i) The second assertion follows the same idea of Lemma 4.1 in [14], where the Peano-Baker series are used. However, for reader's convenience, we briefly sketch its proof below. Note that the coefficient matrix $L(t)$ in (5.2)–(5.3) can be estimated element-wise:

$$-\frac{\kappa L(t)}{P-1} \geq \frac{\kappa L_1}{P-1}\chi(t) - \kappa L_3 I, \quad L_3 = \frac{L_1}{P-1} + L_2.$$

Let $\Psi(t_m, t_n)$ be the state transition matrix corresponding to the following dynamical system:

$$\frac{d\widehat{Q}^R}{dt} = \left(\kappa L_3 I - \frac{\kappa L(t)}{P-1}\right)\widehat{Q}^R.$$

Then, one has

$$\Phi(t_m, t_n) = \exp\Big(- \kappa L_3(t_m - t_n)\Big)\Psi(t_m, t_n), \tag{5.14}$$

and the transition matrix $\Psi(t_{n+1}, t_n)$ can be estimated by the adjacency matrix $\chi(t_n)$ using the Peano-Baker series component wise:

$$\begin{aligned}
\Psi(t_{n+1}, t_n) &= I + \sum_{\ell=1}^{\infty} \int_{t_n}^{t_{n+1}} \int_{t_n}^{\tau_1} \cdots \int_{t_n}^{\tau_{\ell-1}} \left[\left(\kappa L_3 I - \frac{\kappa L(\tau_1)}{P-1}\right) \cdots \left(\kappa L_3 I - \frac{\kappa L(\tau_\ell)}{P-1}\right)\right] d\tau_\ell \dots d\tau_1 \\
&\geq I + \sum_{\ell=1}^{\infty} \int_{t_n}^{t_{n+1}} \int_{t_n}^{\tau_1} \cdots \int_{t_n}^{\tau_{\ell-1}} \left[\left(\frac{\kappa L_1 \chi(\tau_1)}{P-1}\right) \cdots \left(\frac{\kappa L_1 \chi(\tau_\ell)}{P-1}\right)\right] d\tau_\ell \dots d\tau_1 \\
&= I + \sum_{\ell=1}^{\infty} \int_{t_n}^{t_{n+1}} \int_{t_n}^{\tau_1} \cdots \int_{t_n}^{\tau_{\ell-1}} \left(\frac{\kappa L_1 \chi(t_n)}{P-1}\right)^{\ell} d\tau_\ell \dots d\tau_1 \\
&= I + \sum_{\ell=1}^{\infty} \frac{1}{\ell!}\tau^\ell \left(\frac{\kappa L_1 \chi(t_n)}{P-1}\right)^{\ell} \\
&\geq I + \frac{\tau\kappa L_1}{P-1}\chi(t_n).
\end{aligned}$$

On the other hand, since the state-transition matrix satisfies a semi-group property, we have

$$\begin{aligned}
\Psi(t_{n+1}^*, t_n^*) &= \Psi(t_{b_{n+1}}, t_{b_{n+1}-1})\Psi(t_{b_{n+1}-1}, t_{b_{n+1}-2}) \dots \Psi(t_{b_n+1}, t_{b_n}) \\
&\geq \left(I + \frac{\tau\kappa L_1}{P-1}\chi(t_{b_{n+1}-1})\right) \cdots \left(I + \frac{\tau\kappa L_1}{P-1}\chi(t_{b_n})\right) \\
&\geq I + \frac{\tau\kappa L_1}{P-1} \sum_{k=b_n}^{b_{n+1}-1} \chi(t_k).
\end{aligned} \tag{5.15}$$

Next, we combine (5.14) and (5.15) to obtain the following relation element-wise.

$$\Phi(t_{n+1}^*, t_n^*) \geq \exp(-\kappa L_3(t_{n+1}^* - t_n^*)) \left(I + \frac{\tau\kappa L_1}{P-1} \sum_{k=b_n}^{b_{n+1}-1} \chi(t_k) \right). \tag{5.16}$$

Since $\sum_{k=b_n}^{b_{n+1}-1} \chi(t_k)$ is scrambling and consists of integer elements, we have

$$\mu\Big(\sum_{k=b_n}^{b_{n+1}-1} \chi(t_k)\Big) \geq 1. \tag{5.17}$$

Now, we use Remark 5.2(2), (5.16) and (5.17) to get

$$\mu\left(\Phi(t^*_{n+1}, t^*_n)\right) \geq \frac{\tau\kappa L_1}{P-1}\exp(-\kappa L_3(t^*_{n+1} - t^*_n)) > 0.$$

(ii) It follows from (5.16) that each element of $\Phi(t,s)$ are nonnegative. Moreover, note that constant state $((1,\dots,1)^T$ is clearly a solution of (5.1). Thus, we substitute this constant state into (5.4) to get

$$(1,\dots,1)^T = \Phi(t,s)(1,\dots,1)^T.$$

This establishes the stochasticity of $\Phi(t,s)$.

□

5.3. **Proof of Theorem 3.2.** We are now ready to provide a proof of our second main result by combining Lemma 5.1, Lemma 5.2 and Lemma 5.3 altogether. We split its proof into three steps.

• Step A (Derivation of recursive relation): it follows from Lemma 5.1 and Lemma 5.3 that

$$\begin{aligned}
\mathcal{D}(\widehat{Q}^R(t^*_n)) &\leq \mathcal{D}(\widehat{Q}^R(t^*_{n-1}))\left[1 - \mu\left(\Phi(t^*_n, t^*_{n+1})\right)\right] \\
&\leq \mathcal{D}(\widehat{Q}^R(t^*_{n-1}))\left(1 - \frac{\tau\kappa L_1}{P-1}e^{-\kappa L_3(t^*_n - t^*_{n-1})}\right) \\
&\leq \mathcal{D}(\widehat{Q}^R(t^*_{n-1}))\exp\left(-\frac{\tau\kappa L_1}{P-1}e^{-\kappa L_3(t^*_n - t^*_{n-1})}\right).
\end{aligned} \tag{5.18}$$

• Step B (Estimate of $\mathcal{D}(\widehat{Q}^R)$ at discrete time t^*_n: we iterate the recursive relation (5.18) on n and use Jensen's inequality to get

$$\begin{aligned}
\mathcal{D}(\widehat{Q}^R(t^*_n)) &\leq \|Q(0)\|\exp\left[-\frac{\tau\kappa L_1}{P-1}\sum_{m=1}^{n} e^{-\kappa L_3(t^*_m - t^*_{m-1})}\right] \\
&\leq \mathcal{D}(\widehat{Q}^R(0))\exp\left[-\frac{\tau\kappa L_1}{P-1}n\exp\left(-\frac{\kappa L_3}{n}\sum_{m=1}^{n}(t^*_m - t^*_{m-1})\right)\right] \\
&= \mathcal{D}(\widehat{Q}^R(0))\exp\left[-\frac{n\tau\kappa L_1}{P-1}\exp\left(-\frac{\kappa L_3 t^*_n}{n}\right)\right].
\end{aligned} \tag{5.19}$$

• Step C (Decay estimate of $\mathcal{D}(\widehat{Q}^R(t))$): it follows from Lemma 5.2 that

$$a_n = \frac{b_n}{n} = \frac{t^*_n}{n\tau} \geq 1.$$

We substitute $n = \frac{t^*_n}{a_n\tau}$ in the R.H.S. of (5.19) to get

$$\mathcal{D}(\widehat{Q}^R(t)) \leq \mathcal{D}(\widehat{Q}^R(t^*_n)) \leq \mathcal{D}(\widehat{Q}^R(0))\exp\left[-\frac{\kappa L_1}{P-1}\frac{\exp\left(-\kappa L_3\tau a_n\right)}{a_n}\frac{t^*_n}{t}t\right].$$

It follows from Lemma 5.2 that

$$a_n \to \mathbb{E}(b_1) \quad \text{a.s. as } n \to \infty.$$

On the other hand, we set

$$c_n := t/(n\tau) \quad \text{for } t \in [t^*_n, t^*_{n+1}).$$

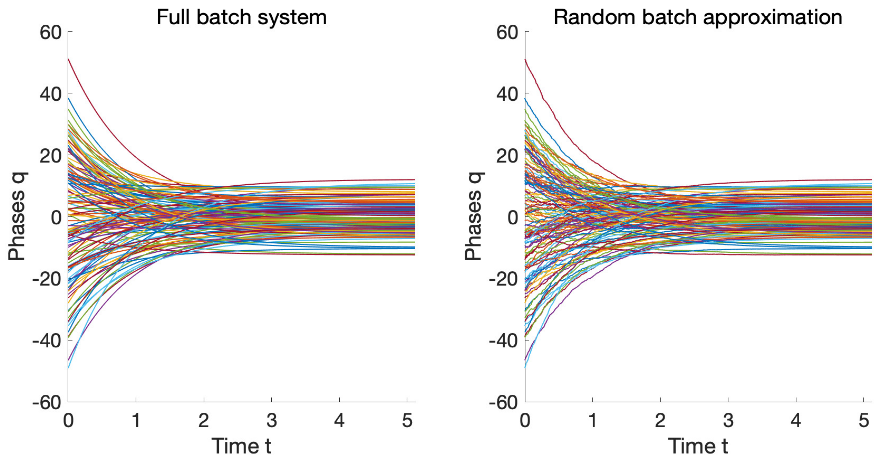

FIGURE 1. Trajectories of $N = 128$ particles along time. **Left:** solution to the full batch system (1.1), **right:** solution to the random batch system (1.3) with $P = 2$ and $\tau = 0.01 = \Delta t$.

Then, one has

$$a_n \le c_n \le a_{n+1}\frac{n+1}{n} \quad \text{so that} \quad 1 \le \frac{t}{t_n^*} = \frac{c_n}{a_n} \le \frac{(n+1)a_{n+1}}{na_n} \to 1 \quad \text{as} \quad n \to \infty, \quad a.s.$$

Hence, for any $\varepsilon > 0$, there exists a constant $T > 0$ such that

$$\mathcal{D}(\widehat{Q}^R(t)) \le \mathcal{D}(\widehat{Q}^R(0)) \exp\left[-\left(\frac{\kappa L_1 \exp\left(-\kappa L_3 \tau \mathbb{E}(b_1)\right)}{(P-1)\mathbb{E}(b_1)} - \varepsilon\right) t\right], \quad \text{a.s. } t \ge T.$$

This completes the proof of Theorem 3.2.

6. Numerical Simulations

In this section, we provide several numerical simulations on the emergent behaviors of the RBM approximated model to compare with the analytical results presented in Section 3. For simplicity, we use a linear model to check the decay rate:

$$\begin{cases} \dfrac{dq_i}{dt} = \dfrac{\kappa}{N-1}\displaystyle\sum_{j\neq i}((q_j - q_i) - (q_j^\infty - q_i^\infty)), \quad t > 0, \\ q_i(0) = q_i^{in} \in \mathbb{R}^1, \quad i = 1, \cdots, N. \end{cases} \tag{6.1}$$

Here, the Lipschitz constants L_1 and L_2 are both 1, and the convergence of the solutions to Q^∞ is guaranteed for any initial data.

For the simulation of the time evolution, we used the forward Euler method to integrate RBM in the simplest setting in which the final time, time-discretization, number of particles, batch size and the coupling strength are chosen as follows:

$$T = 5.12, \quad \Delta t = 0.01, \quad N = 128, \quad P = 2 \quad \text{and} \quad \kappa = 1.$$

In Figure 1, one can see the convergence of the linear system. The left figure visualizes the trajectories of 128 particles in the original full batch system, where the initial configuration and the equilibrium are given from centered normal distributions with standard derivations 20 and 5, respectively. We fixed these data for entire simulations to get consistent results.

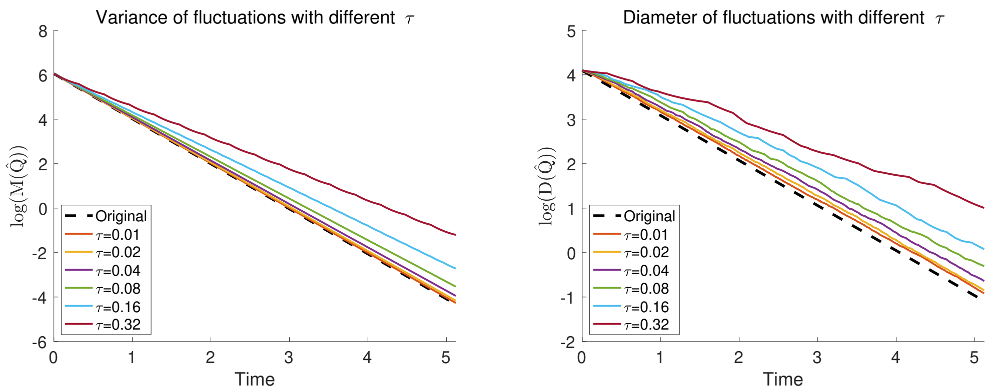

FIGURE 2. Time-evolution of the second moment $\mathbb{E}[\mathcal{M}_2(\widehat{Q})]$ (left) and the maximal diameter $\max[\mathcal{D}(\widehat{Q})]$ (right) for each time-step τ, $\tau = 0.01, \dots, 0.32$ from 200 simulations.

Variance	Decay rate	Diameter	Decay rate
Original	2.0260	Original	1.0130
RBM, $\tau = 0.01$	2.0147	RBM, $\tau = 0.01$	0.9783
RBM, $\tau = 0.02$	1.9942	RBM, $\tau = 0.02$	0.9650
RBM, $\tau = 0.04$	1.9525	RBM, $\tau = 0.04$	0.9252
RBM, $\tau = 0.08$	1.8709	RBM, $\tau = 0.08$	0.8594
RBM, $\tau = 0.16$	1.7116	RBM, $\tau = 0.16$	0.7847
RBM, $\tau = 0.32$	1.4160	RBM, $\tau = 0.32$	0.6037

TABLE 1. Average decay rate from $t = 0$ to $t = 5.12$ computed from Figure 2.

In the right figure, an RBM approximated solution is presented with $P = 2$ and $\tau = 0.01$. Note that the time-step in RBM is equal to the time-discretization of the numerical scheme. The approximated trajectory looks quite similar to the full batch system, although in the beginning (for $t \in [0, 1]$) noisy behaviors can be observed. Recall that the RBM approximated solution has much less computational cost, $\mathcal{O}(N)$, compared to $\mathcal{O}(N^2)$ of the original system (6.1).

In Figure 2, we observe the decay of fluctuations in terms of the averaged variance $\mathbb{E}[\mathcal{M}_2(\widehat{Q})]$ and the maximal diameter $\max[\mathcal{D}(\widehat{Q})]$, which corresponds to Theorems 3.1 and 3.2, respectively. In the left figure, the variance $\mathcal{M}_2(\widehat{Q}(t))$ is computed for each time-step $\tau = 0.01,\ 0.02,\ 0.04,\ 0.08,\ 0.16$ and 0.32, which is averaged over 200 simulations. To compare with RBM approximations, the variance of the original (full batch) trajectories is also presented with dashed line. Note that the decay rate from the full batch system is nearly 2 which coincides with Remark 2.2 (with $\kappa = L_1 = 1$). In the time-evolution of $\mathbb{E}[\mathcal{M}_2(\widehat{Q})]$, the RBM approximated model with $\tau = 0.01$ and the original full-batch system show similar decay rates. The RBM approximation shows quite accurate collective behaviors in terms of variance of fluctuations, which is much better than our analytic estimate in Theorem 3.1.

In the right figure of Figure 2, the diameter $\mathcal{D}(\widehat{Q}(t))$ is computed with 200 simulations for each time-step. Instead of averaging it, we choose the maximal value of diameters from 200 independent RBM trajectories. Though the decay is weakened in the RBM approximation,

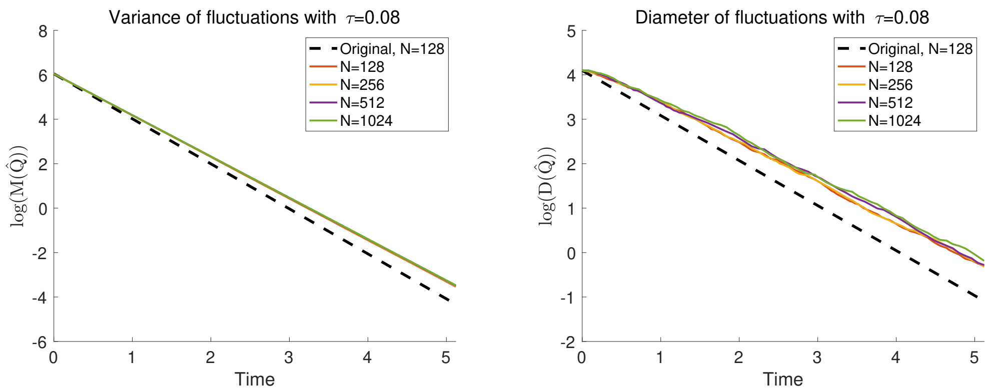

FIGURE 3. Time-evolution of the second moment $\mathbb{E}[\mathcal{M}_2(\widehat{Q})]$ (left) and the diameter $\max[\mathcal{D}(\widehat{Q})]$ (right) for $\tau = 0.08$ with varying N, $N = 128, 256, 512, 1024$ from 200 simulations. The simulations are done with RBM ($P = 2$), while the original full-batch dynamics is drawn for comparison.

Variance	Decay rate	Diameter	Decay rate
Full batch	2.0260	Full batch	1.0130
RBM, $N = 128$	1.8709	RBM, $N = 128$	0.8594
RBM, $N = 256$	1.8658	RBM, $N = 256$	0.8623
RBM, $N = 512$	1.8610	RBM, $N = 512$	0.8548
RBM, $N = 1024$	1.8585	RBM, $N = 1024$	0.8365

TABLE 2. Average decay rate from $t = 0$ to $t = 5.12$ computed from Figure 3.

it also shows quite similar decay rate with $\tau = 0.01$. The averaged decay rates are presented in Table 1.

In Figure 3, instead of different time-steps τ, we present the corresponding simulations with varying N, the number of particles. To observe the dependency of the decay rates Λ_1 and Λ_2 on N, the averaged variances and the maximal diameters are drawn from 200 simulations are done for each $N = 128$, 256, 512 and 1024.

One of the difficulty of varying N is the choice of the initial data. In order to compare the trajectories from different number of particles, we set duplicated initial data; a pair of particles is placed for each initial position shown in Figure 1 ($N = 128$) to operate numerical simulations with $N = 256$. We set 4 or 8 particles in the same position for $N = 512$ and 1024, respectively. From this setting, the initial variance $\mathcal{M}_2(\widehat{Q}(0))$ and the maximal diameter $\mathcal{D}(\widehat{Q}(0))$ are exactly the same with different N. The results in Figure 3 and Table 2 show that the decay rates are similar with different N, which is not proportional to $1/N$ or $1/N^2$ as in Theorem 3.1 or 3.2.

7. CONCLUSION

In this paper, we have studied asymptotic convergence toward the equilibrium of the first-order nonlinear consensus model with random batch interactions. Our first-order consensus model is general enough to include well-studied aggregation models such as a linear consensus model and the Kuramoto model for synchronization. Since the RBM model is approximate in nature, its asymptotic behaviors can sometimes be qualitatively different from the original full batch model as discussed in [25]. In fact, solutions to the RBM model can fluctuate around an equilibrium state of the full batch model due to random interactions. Hence, the convergence should hold only in the sense of time averaging. To overcome this difficulty, we adopt a variant [28] of the RBM which assumes a priori knowledge on the equilibrium point. Then, the approximated system converges toward the equilibrium of the full batch model exponentially fast, which is established in this article.

For the convergence analysis, we use the contractive property of the RBM-approximate model and the linear approximation to estimate the exponential decay rate. In this work, we provided an exponential decay of quantity of interest in two different approaches; one is from the variance of fluctuations, and the other is to use the diameter of fluctuations almost surely. Both estimations show that the decay exponent is much smaller than the original system in terms of N, the number of interacting particles. We think that the dependence on N in decay rate is due to the analytic techniques that we used in this paper, where we can only check the contraction of the fluctuation of each batch. However, numerical simulations seem to suggest that decay exponents are independent of N. Thus, our analytical decay estimates are still far from the optimal decay rate. Analysis on sharp decay rate will be an interesting and challenging open problem for a future work.

All the decay estimate that we have considered assume that the original full batch model has all-to-all interactions. Thus, the consensus model with a general connection topologies can cause a lot of technical difficulties in convergence analysis. For example, the second moment of one random batch is no more guaranteed to decay exponentially even for a time interval $[t_m, t_{m+1}]$ due to the connectivity inside a batch. Moreover, switching network arguments in [14] rely on the ergodicity coefficient, which is not yet developed under a general network topology. We leave this issue for a future work as well.

References

[1] Acebron, J. A., Bonilla, L. L., Pérez Vicente, C. J. P., Ritort, F. and Spigler, R.: *The Kuramoto model: A simple paradigm for synchronization phenomena.* Rev. Mod. Phys. **77** (2005), 137-185.

[2] Albi, G., Bellomo, N., Fermo, L., Ha, S.-Y., Kim, J., Pareschi, L., Poyato, D. and Soler, J.: *Vehicular traffic, crowds and swarms: From kinetic theory and multiscale methods to applications and research perspectives.* Math. Models Methods Appl. Sci. **29** (2019), 1901-2005.

[3] Alpin, Yu. A. and Gabassov, N. Z.: *A remark on the problem locating the eigenvalues of real matrices.* Izv. Vyssh. Uchebn. Zaved. Mat. **11** (1976), 98–100.

[4] Bellomo, N. and Ha, S.-Y.: *A quest toward a mathematical theory of the dynamics of swarms.* Math. Models Methods Appl. Sci. **27** (2017), 745-770.

[5] Bellomo, N. and Soler, J.: *On the mathematical theory of the dynamics of swarms viewed as complex systems.* Math Models Methods Appl. Sci. **22** (2012), 1140006, 29 pp.

[6] Bronski, J., Carty, T. and Simpson, S.: *A matrix valued Kuramoto model.* J. Stat. Phys. **178** (2020), 595-624.

[7] Buck, J. and Buck, E.: *Biology of synchronous flashing of fireflies.* Nature **211** (1966), 562.

[8] Carrillo, J. A., Jin, S., Li, L. and Zhu, Y.: *A consensus-based global optimization method for high dimensional machine learning problems.* ESAIM Control Optim. Calc. Var. **27** (2021), S5.

[9] Chi, D., Choi, S.-H. and Ha, S.-Y.: *Emergent behaviors of a holonomic particle system on a sphere.* J. Math. Phys. **55** (2014), 052703.

[10] Choi, Y., Ha, S.-Y., Jung, S. and Kim, Y.: *Asymptotic formation and orbital stability of phase-locked states for the Kuramoto model.* Physica D **241** (2012), 735-754.

[11] Chopra, N. and Spong, M. W.: *On exponential synchronization of Kuramoto oscillators.* IEEE Trans. Automatic Control **54** (2009), 353-357.

[12] Cucker, F. and Smale S.: *Emergent behavior in flocks.* IEEE Trans. Automat. Control **52** (2007), 852-862.

[13] Degond, P., Frouvelle, A., Merino-Aceituno, S. and Trescases, A.: *Quaternions in collective dynamics.* Multiscale Model. Simul. **16** (2018), 28–77.

[14] Dong, J.-G., Ha, S.-Y., Jung, J. and Kim, D.: *On the stochastic flocking of the Cucker-Smale flock with randomly switching topologies.* SIAM J. Control Optim. **58** (2020), 2332-2353.

[15] Dörfler, F. and Bullo, F.: *Synchronization in complex networks of phase oscillators: A survey.* Automatica **50** (2014), 1539-1564.

[16] Golse, F., Jin, S. and Paul, T.: *The random batch method for N-body quantum dynamics.* Found. Comut. Math., to appear.

[17] Ha, S. Y., Ha, T. and Kim, J. H.: Emergent behavior of a Cucker-Smale type particle model with nonlinear velocity couplings. IEEE Trans. Autom. Control, **55** (2010), 1679-1683.

[18] Ha, S.-Y., Jin, S., Kim, D. and Ko, D.: *Uniform-in-time error estimate of the random batch method for the Cucker-Smale model.* Math. Models Methods Appl. Sci., to appear.

[19] Ha, S.-Y., Ko, D., Park, J. and Zhang, X.: *Collective synchronization of classical and quantum oscillators.* EMS Surv. Math. Sci. **3** (2016), 209-267.

[20] Ha, S.-Y., Kim, H. W. and Ryoo, S. W.: *Emergence of phase-locked states for the Kuramoto model in a large coupling regime.* Commun. Math. Sci. **14** (2016), 1073-1091.

[21] Ha, S.-Y., Jin, S. and Kim, D.: *Convergence of a first-order consensus-based global optimization algorithm.* To appear in Math. Models Methods Appl. Sci. Available at arXiv:1910.08239 (2019).

[22] Ha, S.-Y. and Park, H.: *From the Lohe tensor model to the Lohe Hermitian sphere model and emergent dynamics.* SIAM J. Appl. Dyn. Syst. **19** (2020), 1312-1342.

[23] Ha, S.-Y. and Park, H.: *Emergent behaviors of Lohe tensor flock.* J. Stat. Phys. **178** (2020), 1268-1292.

[24] Jin, S. and Li, Lei.: *On the mean field limit of Random Batch Method for interacting particle systems*, Available at arXiv:2005.11740 (2020).

[25] Jin, S., Li, L. and Liu, J.-G.: *Random batch methods (RBM) for interacting particle systems.* J. of Computational Physics **400** (2020), 108877.

[26] Jin, S., Li, L. and Liu, J.-G.: *Convergence of Random Batch Method for interacting particles with disparate species and weights,* SIAM J. Numer. Anal., to appear.

[27] Kang, J.-H., Ha, S.-Y., Kang, K. and Jeong, E.: *How do cultural classes emerge from assimilation and distinction? An extension of the Cucker-Smale flocking model.* J. Math. Sociol. **38** (2014), 47-71.

[28] Ko, D., Ha, S.-Y., Jin, S. and Kim, D.: *Uniform error estimates for the random batch method to the first-order consensus models with anti-symmetric interaction kernels.* To appear in Stud. Appl. Math.

[29] Ko, D. and Zuazua, E.: *Model predictive control with random batch methods for a guiding problem.* Available at arXiv:2004.14834 (2020).

[30] Kuramoto, Y.: *Self-entrainment of a population of coupled non-linear oscillators.* In International symposium on mathematical problems in theoretical physics (pp. 420-422) Springer, Berlin, Heidelberg (1975).

[31] Li, L., Liu, J. G. and Tang, Y.: *A direct simulation approach for the Poisson-Boltzmann equation using the random batch method.* Available at arXiv:2004.05614 (2020).

[32] Lohe, M. A.: *Non-abelian Kuramoto model and synchronization.* J. Phys. A: Math. Theor. **42** (2009), 395101.

[33] Motsch, S. and Tadmor, E.: *Heterophilious dynamics enhances consensus.* SIAM Rev. **56** (2014), 577–621.

[34] Olfati-Saber, R. and Murray, R. M.: Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Automat. Control*, **49** (2004), 1520-1533.

[35] Peskin, C. S.: *Mathematical aspects of heart physiology.* Courant Institute of Mathematical Sciences, New York, 1975.

[36] Pikovsky, A., Rosenblum, M. and Kurths, J.: *Synchronization: A universal concept in nonlinear sciences.* Cambridge University Press, Cambridge, 2001.

[37] Rokhlin, V.: *Rapid solution of integral equations of classical potential theory.* Journal of computational physics, **60** (1985), 187–207.

[38] Spanos, D. P., Olfati-Saber, R., and Murray, R. M.: *Dynamic consensus on mobile networks.* IFAC world congress. Czech Republic: Prague, (2005), 1-6.

[39] Toner, J. and Tu, Y.: *Flocks, herds, and schools: A quantitative theory of flocking.* Phys. Rev. E **58** (1998), 4828-4858.
[40] Topaz, C. M. and Bertozzi, A. L.: *Swarming patterns in a two-dimensional kinematic model for biological groups.* SIAM J. Appl. Math. **65** (2004), 152-174.
[41] Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I. and Schochet, O.: *Novel type of phase transition in a system of self-driven particles.* Phys. Rev. Lett. **75** (1995), 1226-1229.
[42] Vicsek, T. and Zefeiris, A.: *Collective motion.* Phys. Rep. **517** (2012), 71-140.
[43] Winfree, A. T.: *The geometry of biological time.* Springer, New York, 1980.

(Ha)
Department of Mathematical Sciences and Research Institute of Mathematics,
Seoul National University, Seoul 08826 and
School of Mathematics, Korea Institute for Advanced Study,
Hoegiro 85, Seoul 02455, Republic of Korea
Email address: syha@snu.ac.kr

(Jin)
School of Mathematical Sciences, Institute of Natural Sciences, and MOE-LSC,
Shanghai Jiao Tong University, Shanghai 200240, China
Email address: shijin-m@sjtu.edu.cn

(Kim)
School of Mathematics, Korea Institute for Advanced Study,
Hoegiro 85, Seoul 02455, Republic of Korea
Email address: doheonkim@kias.re.kr

(Ko)
Department of Mathematics, The Catholic University of Korea,
Jibongro 43, Bucheon, Gyeonggido 14662, Republic of Korea
Email address: dongnamko@catholic.ac.kr