Les Cahiers du GERAD ISSN: 0711–2440

Mean field games: Large sparse network limits and Laplexion dynamics

M. Huang, P. E. Caines

G–2026–12

March 2026

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2026
– Bibliothèque et Archives Canada, 2026

Legal deposit – Bibliothèque et Archives nationales du Québec, 2026
– Library and Archives Canada, 2026

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Mean field games: Large sparse network limits and Laplexion dynamics

Minyi Huang [a, b]

Peter E. Caines [c, b]

[a] *School of Mathematics and Statistics, Carleton University, Ottawa (On), Canada, K1S 5B6*

[b] *GERAD, Montréal (Qc), Canada, H3T 1J4*

[c] *Department of Electrical and Computer Engineering, McGill University, Montreal (Qc), Canada, H3A 2A7*

mhuang@math.carleton.ca
peter.caines@mcgill.ca

March 2026
Les Cahiers du GERAD
G–2026–12

Abstract : Dynamic games are considered with large subpopulations distributed over large sparse graphs, where each agent has mean field coupling with all agents within the same cluster and meanwhile receives impact from neighboring clusters. Tractable limit models are obtained via the network limits of graph Laplacian operators modelling second order interactions. The resulting mean field game (MFG) has what is called Laplexion dynamics. For ring and torus topology cases, solutions for linear quadratic Laplexion MFG systems with infinite populations on infinite node networks are obtained via a second order linear partial differential equation system. This work focusses on the relationship between the finite population model and the sparse network limit model. We prove an ϵ-Nash equilibrium property for the obtained decentralized strategies.

Keywords : Immediate neighbor interactions; Laplexion; large networks; mean field games; Nash equilibria

1 Introduction

1.1 GMFG and GXMFG theory

Graphon theory [30] provides a natural framework for modeling the interaction of agents distributed over large networks when the nodes can be indexed by the reals in $[0,1]$ furnished only with the topology of the node set $[0,1]$. Graphons are symmetric measurable functions on $[0,1]^2$, taking values in $[0,1]$ (interpreted as generalized adjacency matrices). The graphon framework has been widely used in the analysis of dynamics, control and games on large networks (see e.g. [2, 8, 14, 16, 20, 26, 31, 34]). Graphon Mean Field Games (GMFGs) [7, 8] are formulated so as to describe the limiting behavior of large populations of dynamical agents interacting over large networks, the key property of interest being the existence of Nash equilibria.

The GMFG framework is formulated in terms of graphons and hence is restricted to asymptotically dense networks. However, many networks of interest are not dense. Large sparse graphs occur, for instance, in biological neural networks [17] and electrical power grids [33]. When the sequence of graphs is sparse but the node degrees still have unbounded growth, graphon theory is extended by an L^p theory for sparse graph limits [4]. The reader is also referred to sparse limit graph dynamical models in [27] for related game models on sparse networks, and general convergence analyses of interacting particles over graphs, including sparse graphs, in [25, 28]. The graphexon MFG (GXMFG) construction generalizes the GMFG framework to the graphexon setting ([5]) and consequently applies to networks which are asymptotically a combination of sparse and dense graph limits. Unlike the graphon case, non-trivial subsequential graphexon limits are defined for all graph sequences and hence in particular for sparse graph sequences. The solution of the GXMFG [9] assigns coupling weights to other nodes using a graphexon measure instead of a graphon function in GMFG equations [6, 8].

1.2 MFG and Laplexion dynamics

Immediate neighbor based interactions have been addressed in the game theoretic literature [3, 24, 36] and arise in markets of retailing services or goods [3]. Such interaction structures are also adopted in deriving continuum dynamics as motivated by bacterial colonies and flocking [18], and interacting particle models in mathematical physics [1, 21].

To describe interactions of immediate neighboring clusters, our previous work considers graphexon limits of sparse networks and proposes first order interaction models [9] and second order interactions models [10] containing a mean field derivative term in the nodal parameter. This interaction structure accounts for the influence of neighboring subpopulations through the spatial variations of their behaviors, while the standard graphon network limit framework has no defined metric topology on the set of vertices in which to define the notion of locality. The more recent work [11] starts with a dynamical model via graph Laplacian operators and, taking their sparse network limit, provides a mathematical derivation of second order interaction terms in [10], which, due to their interpretation via limit Laplacians, are called Laplexions. Graph Laplacians have been widely used in other areas such as multi-agent coordination control [32] and machine learning [13]. For developing limit theory, the modeling in this paper starts with finite populations on stylized finite networks as either one dimensional or two dimensional lattices [1, 3, 18, 21]. Our network models are cases of extreme sparsity, and are not covered by the sparse limit theory as in [4, 25]. In a linear-quadratic (LQ) setting, this paper addresses the relationship between the finite population model and the sparse network limit model while some preliminary analysis has appeared in [12] without detailed proofs. Our analysis depends on techniques different from those employed in MFGs and GMFGs, primarily due to the second order (diffusive) interactions. Our techniques will combine mean field approximations for intra-cluster coupling and Laplexion approximations of the seconder order interactions for inter-cluster coupling. The Laplexion is closely related to hydrodynamic limit under scaling of diffusion type in asymptotic analysis of interacting particle systems [1, 21].

The paper is organized as follows. Section 2 introduces the dynamical modeling involving a graph Laplacian operator to capture the influence of subpopulations at neighbouring nodes, which leads to the so-called Laplexion mean field game (LMFG) in the large network limit. Section 3 derives the solution equation system of the LQ LMFG and the associated decentralized strategies. Section 4 presents an existence and uniqueness analysis of the partial differential equation (PDE) system of the LMFG. Section 5 gives the a priori cost bound for the decentralized strategies. Section 6 analyzes the Laplexion limit of the second order interaction terms. Section 5 establishes an asymptotic Nash equilibrium theorem.

Notation: We use N^ι (ι=1 for the ring, $\iota = 2$ for the torus) to denote the number of nodes in the graph, and N_1 to denote the subpopulation size at each node. We take $C, C_1, \cdots$ as generic constants, which change from place to place but do not depend on (N, N_1).

2 Dynamic games with subpopulations over sparse networks and the Laplexion limit

We start with some preliminaries on graph limits and graph Laplacian operators.

2.1 Large graphs and their limits

Let $(\mathcal{M}, \mathbf{g})$ be a compact, connected Riemannian manifold of dimension p, equipped with its geodesic distance $d_{\mathbf{g}}$ and Riemannian volume measure $dV_{\mathbf{g}}$. Let $\mathcal{P}(\mathcal{M})$ be all probability measures defined on $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$ with Borel σ-field $\mathcal{B}(\mathcal{M})$.

Definition 2.1 (Embedded graph sequence)**.** $G^N = (\mathcal{V}^N, \mathcal{E}^N)$ shall denote a non-empty finite graph without self-loops, with node set denoted $\mathcal{V}^N := \{1, \cdots, N\}$, and undirected edge set of unordered pairs denoted $\mathcal{E}^N \subseteq \mathcal{V}^N \times \mathcal{V}^N$. The sequence $\{G^N, \varphi^N\}_{N=1}^{\infty}$ is an *embedded graph sequence* if every G^N is embedded into $\mathcal{M}$ via a bijective node mapping $\varphi^N : \mathcal{V}^N \to \mathcal{M}$, assigning a location $\alpha_i^N = \varphi^N(i) \in \mathcal{M}$ to each node $i \in \mathcal{V}^N$.

Definition 2.2 (Empirical vertex and edge measures)**.** Let $G_N = (\mathcal{V}_N, \mathcal{E}_N)$ be a finite embedded graph on a Riemannian manifold $\mathcal{M}$ with an embedding mapping $\varphi^N : \mathcal{V}_N \to \mathcal{M}$, with $\alpha_i^N = \varphi^N(i)$.

1. The *empirical vertex measure* $V_N \in \mathcal{P}(\mathcal{M})$ is defined by

$$V_N \;:=\; \frac{1}{|\mathcal{V}_N|} \sum_{i \in \mathcal{V}_N} \delta_{\alpha_i^N},$$

   where $\delta_{\alpha_i^N} \in \mathcal{P}(\mathcal{M})$ denotes the Dirac probability measure on $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$ concentrated on α_i^N.
2. The *empirical graphexon measure* $W_N \in \mathcal{P}(\mathcal{M} \times \mathcal{M})$ is defined by

$$W_N \;:=\; \frac{1}{2|\mathcal{E}^N|} \sum_{\{i,j\} \in \mathcal{E}^N} \big(\delta_{(\alpha_i^N, \alpha_j^N)} + \delta_{(\alpha_j^N, \alpha_i^N)}\big),$$

   where $\delta_{(\alpha_i^N, \alpha_j^N)} \in \mathcal{P}(\mathcal{M} \times \mathcal{M})$ denotes the Dirac probability measure on $(\mathcal{M} \times \mathcal{M}, \mathcal{B}(\mathcal{M} \times \mathcal{M}))$ that concentrates on the ordered pair (α_i^N, α_j^N) whenever $\{i, j\} \in \mathcal{E}^N$.

Remark 2.3. The measure W_N assigns zero measure to the diagonal in $\mathcal{M} \times \mathcal{M}$ since the graphs under consideration have no self-loops.

Since $\mathcal{M}$ is a compact metric space, the space of probability measures $\mathcal{P}(\mathcal{M})$ is tight. Consequently, by Prokhorov's theorem, the space of probability measures $\mathcal{P}(\mathcal{M})$, endowed with the topology of weak convergence, is also compact. Consequently, the sequence of empirical vertex measures $\{V_N\}_{N \in \mathbb{N}}$ is relatively compact which guarantees the existence of convergent subsequences.

Definition 2.4 ([5, 9, 10]). Let (V_N, W_N) be the empirical vertex and edge measures from Definition 2.2. We equip $\mathcal{P}(\mathcal{M})$ with the topology of weak convergence of probability measures, and let $\mathcal{P}_s(\mathcal{M}\times\mathcal{M})$ denote the space of finite *symmetric* Borel measures on $\mathcal{M}\times\mathcal{M}$.

- A *vertexon* is any weak limit $V_\infty \in \mathcal{P}(\mathcal{M})$ of a subsequence $\{V_{N_k}\}_{k=1}^\infty$.
- Given such a vertexon V_∞ realized by $\{N_k\}_{k=1}^\infty$, an associated *graphexon* is any weak limit $W_\infty \in \mathcal{P}_s(\mathcal{M}\times\mathcal{M})$ of a further subsequence in $\{W_{N_k}\}_{k=1}^\infty$.

The limit V_∞ depends on the chosen subsequence $\{N_k\}_{k=1}^\infty$. Hence, both V_∞ and the associated W_∞ need not be unique.

2.2 The graph Laplacian operator

Consider an undirected graph $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{\alpha_1, \cdots, \alpha_N\}$ is the set of (distinct) nodes and $\mathcal{E}$ is the set of edges (α_i, α_j) with $i \neq j$. For convenience in the subsequent analysis, we label the node set in this form instead of using a set of integers. The finite graph Laplacian $(g_{\alpha_i\alpha_j})_{1\le i,j,\le N}$ is a symmetric matrix, where for $i \neq j$, $g_{\alpha_i\alpha_j} > 0$ if and only if (α_i, α_j) is an edge (in which case we denote $\alpha_i \sim \alpha_j$), and $g_{\alpha_i\alpha_j} = 0$ otherwise; $g_{\alpha_i\alpha_i} = -\sum_{j\neq i} g_{\alpha_i\alpha_j}$.

For a function ψ from $\mathcal{V}$ to $\mathbb{R}^n$, denote the finite graph Laplacian operator by

$$(\mathcal{L}_N\psi)(\alpha_i) := \sum_{\alpha_j:\alpha_j\sim\alpha_i} g_{\alpha_i\alpha_j}(\psi(\alpha_j) - \psi(\alpha_i)).$$

So $\mathcal{L}_N\psi$ is still a function from $\mathcal{V}$ to $\mathbb{R}^n$.

2.3 Agent dynamics and costs within a finite population

For the graph G in Section 2.2, suppose that each of the N nodes is occupied by a cluster (or called a subpopulation) of agents. At node α_k, $1 \le k \le N$, denote the agent states by $x^i_{\alpha_k}(t)$, $1 \le i \le N_1$. For notational simplicity, we assume equal subpopulation sizes N_1 although this restriction can be relaxed. Define the local empirical state distribution of the subpopulation at node α_k:

$$\mu^{(N_1)}_{\alpha_k}(t) = \frac{1}{N_1}\sum_{j=1}^{N_1} \delta_{x^j_{\alpha_k}(t)}, \quad 1 \le k \le N,$$

where δ_x is the Dirac measure at the point x.

While we are primarily interested in developing our methodology in an LQ setting, our exposition starts with dynamics and costs in a general form to make the notation more compact. For agent dynamics we introduce the stochastic differential equation (SDE):

$$\begin{aligned} dx^i_{\alpha_k}(t) =& f(x^i_{\alpha_k}(t), u^i_{\alpha_k}(t), \mu^{(N_1)}_{\alpha_k}(t))dt \\ &+ \sum_{m:m\neq k} g_{\alpha_k\alpha_m}\Big[\int_{\mathbb{R}^n}\phi(y)\mu^{(N_1)}_{\alpha_m}(t,dy) - \int_{\mathbb{R}^n}\phi(y)\mu^{(N_1)}_{\alpha_k}(t,dy)\Big]dt + \sigma dw^i_{\alpha_k}, \end{aligned} \tag{2.1}$$

where the state $x^i_{\alpha_k}(t)$, control $u^i_{\alpha_k}(t)$ and standard Brownian motion $w^i_{\alpha_k}(t)$ have values in $\mathbb{R}^n$, $\mathbb{R}^{n_1}$, and $\mathbb{R}^{n_2}$, respectively; $\phi = (\phi_1, \cdots \phi_n)^T : \mathbb{R}^n \to \mathbb{R}^n$ is a given coupling function; and σ is a constant coefficient matrix. The initial states $x^i_{\alpha_k}(0)$, $1 \le i \le N_1, 1 \le k \le N$, are independent. The standard Brownian motions are independent and also independent of the initial states. The above graph Laplacian based inter-cluster coupling is different from the dynamical coupling in [8] since in the present model we do not assume a dense graph sequence.

The individual cost is given by

$$J^i_{\alpha_k} = \mathbb{E}\int_0^T L(x^i_{\alpha_k}(t), u^i_{\alpha_k}(t), \mu^{(N_1)}_{\alpha_k}(t))dt + \mathbb{E}h(x^i_{\alpha_k}(T), \mu^{(N_1)}_{\alpha_k}(T)). \tag{2.2}$$

2.4 Limit dynamics of infinite sub-populations on a finite graph

As an approximation of (2.1)–(2.2) with $N_1 \to \infty$, we next introduce a mid-level model where the N nodes of the graph G are each occupied by an infinite subpopulation. At time t, the agents of the cluster residing at node $\beta \in \mathcal{V} = \{\alpha_1, \cdots, \alpha_N\}$ generate the local mean field $\mu_\beta(t)$, as the infinite population limit of $\mu_\beta^{(N_1)}(t)$ when $N_1 \to \infty$.

Given the coupling function ϕ, define $\bar\phi_i(t,\beta) = \int_y \phi_i(y)\mu_\beta(t,dy) \in \mathbb{R}^n$, $1 \le i \le n$, $\bar\phi = (\bar\phi_1, \cdots, \bar\phi_n)^T$ for $\beta \in \mathcal{V}$. Let $x_\alpha(t)$ denote the state of a representative agent at node $\alpha \in \mathcal{V}$. We consider the following dynamics

$$dx_\alpha(t) = f(x_\alpha(t), u_\alpha(t), \mu_\alpha(t))dt + (\mathcal{L}_N \bar\phi(t,\cdot))(\alpha)dt + \sigma dw_\alpha(t), \tag{2.3}$$

where $\mathcal{L}_N$ is the finite graph Laplacian operator.

We further introduce the cost of the representative agent

$$J_\alpha = \mathbb{E}\int_0^T L(x_\alpha(t), u_\alpha(t), \mu_\alpha(t))dt + \mathbb{E}h(x_\alpha(T), \mu_\alpha(T)).$$

2.5 The solution via infinite populations and large network limit

For asymptotic analysis with a suitable neighbor structure, we take the ring network $\mathbb{T} = [0,1)$ (where 1 is identified with 0) and the torus network $\mathbb{T}^2$ as prototype network limit models. They are intended to approximate large one and two dimensional networks while their periodic structure avoids the nuisance of boundary effects.

(I) **Discrete ring network** $\mathbb{T}_{1/N}$. Consider the partition of $[0,1)$ into N subintervals $\mathcal{I}_k = [(k-1)/N, k/N)$, $k = 1, \cdots, N$. Let the midpoint $\alpha_k = \frac{k-1}{N} + \frac{1}{2N}$ denote the nodal location. Denote the nodal set $\mathcal{V}_{1/N} = \{\alpha_k, 1 \le k \le N\}$ forming a ring, as a graph, denoted by $\mathbb{T}_{1/N}$ (see Fig. 1) so that $\alpha_1 \sim \alpha_N$. We define the graph Laplacian as an $N \times N$ matrix: for $j \ne i$,

$$g_{\alpha_i\alpha_j} = N^2, \quad \text{if } \alpha_j \sim \alpha_i,$$

and $g_{\alpha_i\alpha_j} = 0$ otherwise; and $g_{\alpha_i\alpha_i} = -2N^2$.

(II) **Discrete torus network** $\mathbb{T}^2_{1/N}$. The discrete torus is a graph consisting of N^2 nodes of the form $\alpha_{ij} \coloneqq (\alpha_i, \alpha_j)$ (an ordered pair), where $\alpha_i, \alpha_j \in \mathcal{V}_{1/N}$, so that each node α_{ij} has 4 neighbors. The nodal set is $\mathcal{V}^2_{1/N}$. The graph Laplacian is defined by the rule: for $\alpha_{ij} \ne \alpha_{i'j'}$, $g_{\alpha_{ij}\alpha_{i'j'}} = N^2$ if $\alpha_{ij} \sim \alpha_{i'j'}$, and $g_{\alpha_{ij}\alpha_{i'j'}} = 0$ otherwise; and $g_{\alpha_{ij}\alpha_{ij}} = -4N^2$ for all i,j.

In the mathematical physics literature, a scaling by N^2 for immediate neighbors is customarily called the scaling of diffusion type [21], which is used to derive the so-called hydrodynamic limits. We will analyze limits of the finite graph Laplacian operators in (2.3).

Recall that $\bar\phi_i(t,\alpha) = \int \phi_i(y)\mu_\alpha(t,dy)$ for $\alpha \in \mathcal{V}$. To analyze the large network limit, we take the continuum vertex set $\mathcal{V}_\infty$ as $\mathbb{T}^\iota$, $\iota = 1,2$, and let $\mu_\alpha(t)$ be defined for all $\alpha \in \mathcal{V}_\infty$, so that $\bar\phi_i(t,\alpha)$ is now defined for $\alpha \in \mathcal{V}_\infty = [0,1)^\iota$. We further identify $\bar\phi_i$ as a function defined on $[0,T] \times \mathbb{R}^\iota$ (still denoted as $\bar\phi_i$), with period 1 for each component of $\mathbb{R}^\iota$. A function $f(t,z)$ on $[0,T] \times \mathbb{R}^k$, $k \ge 1$, is said to have period 1 for each space variable if $f(t, z_1 + i_1, \cdots z_k + i_k) = f(t, z_1, \cdots, z_k)$ for all $z_1, \cdots, z_k \in \mathbb{R}$ and all $i_1, \cdots, i_k \in \mathbb{Z}$. If $\bar\phi_i$ has up to second order continuous partial derivatives in the variable $\alpha \in \mathbb{R}^\iota$, by elementary calculation, we derive

$$\lim_{N\to\infty} \sup_{\alpha \in \mathcal{V}^\iota_{1/N}} |\mathcal{L}_N \bar\phi(t,\cdot)(\alpha) - \Delta\bar\phi(t,\alpha)| = 0.$$

where

$$\Delta\bar\phi(t,\alpha) \coloneqq (\Delta\bar\phi_1(t,\alpha), \cdots, \Delta\bar\phi_n(t,\alpha))^T,$$

$$\Delta\bar{\phi}_i(t,\alpha) = \begin{cases} \partial^2_\alpha \bar{\phi}_i(t,\alpha) & \text{if } \alpha \in \mathbb{T}, \\ \sum_{k=1}^2 \partial^2_{\alpha^{(k)}} \bar{\phi}_i(t,\alpha) & \text{if } \alpha = (\alpha^{(1)}, \alpha^{(2)}) \in \mathbb{T}^2. \end{cases}$$

The following limit model for (2.3) takes a formal limit of the graph Laplacian operator and is called a *Laplexion Mean Field Dynamical System*

$$dx_\alpha(t) = f(x_\alpha(t), u_\alpha(t), \mu_\alpha(t))dt + \bar{\Phi}(t,\alpha)dt + \sigma dw_\alpha(t), \tag{2.4}$$

which specifies the dynamics of a representative agent $\mathcal{A}_\alpha$ at node $\alpha \in \mathcal{V}_\infty$ and where we suppose $\bar{\Phi}(t,\alpha) \coloneqq \Delta\bar{\phi}(t,\alpha)$ is well-defined. The cost function is given by

$$J_\alpha = \mathbb{E}\int_0^T L(x_\alpha(t), u_\alpha(t), \mu_\alpha(t))dt + \mathbb{E}h(x_\alpha(T), \mu_\alpha(T)). \tag{2.5}$$

The above Laplexion framework is different from the GMFG case which depends on a graphon function for coupling. For the ring case of $\mathbb{T}_{1/N}$, the vertexon W_∞, as the counterpart of a graphon, assigns a uniform probability distribution along the diagonal line of the unit square $[0,1)^2$, and as a result, the use of W_∞ alone cannot indicate interconnections between clusters. In contrast, the incorporation of the Laplacian operator enables us to capture the immediate neighbor interactions in the limit model (2.4).

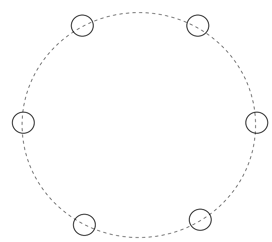

Figure 1: The ring network with two immediate neighbors

2.6 The Laplexion mean field game (LMFG) equations

Denote $\Sigma_w = \sigma\sigma^T$. Based on locally infinite populations and the network limit, the formal LMFG solution pair consists of the Hamilton–Jacobi–Bellman (HJB) equation and the Fokker–Planck–Kolmogorov (FPK) equation:

$$\begin{aligned} -\partial_t V_\alpha(t,x) &= \inf_u \Big\{\partial_x V_\alpha(t,x)[f(x,u,\mu_\alpha(t)) + \bar{\Phi}(t,\alpha)] + L(x,u,\mu_\alpha(t))\Big\} \\ &\quad + \frac{1}{2}\mathrm{Tr}(\Sigma_w \partial_x^2 V_\alpha(t,x)), \end{aligned} \tag{2.6}$$

$$V_\alpha(T,x) = h(x,\mu_\alpha(T)), \quad (t,x) \in [0,T]\times\mathbb{R}^n, \quad \alpha \in \mathcal{V}_\infty,$$

$$\partial_t p_\alpha(t,x) = -\mathrm{div}\{[f(x,u^0,\mu_\alpha(t)) + \bar{\Phi}(\alpha,t)]p_\alpha(t,x)\} + \frac{1}{2}\mathrm{Tr}(\Sigma_w \partial_x^2 p_\alpha(t,x)), \tag{2.7}$$

where $u^0 \coloneqq u^0(t,x|\mu_G)$ is the best response control law determined by Equation (2.6), $\bar{\Phi}(t,\alpha) = \Delta\int_{\mathbb{R}^n}\phi(y)p_\alpha(t,y)dy$ with coupling function ϕ, and $\mu_G = (\mu_\beta)_{0\le\beta<1}$. We use $p_\alpha(t,\cdot)$ to denote the probability density function of $\mu_\alpha(t)$ (as the local mean field at node α) if it exists. The term $\partial_x^2 V_\alpha$ stands for the Hessian matrix, and $\mathrm{div}(\cdot)$ for divergence.

As in standard mean field games, the HJB Equation (2.6) is derived by solving an optimal control problem of agent $\mathcal{A}_\alpha$ taking $\mu_\alpha(\cdot)$ as fixed. The solution of the HJB–FPK equation system then consists of the family of function pairs (V_α, p_α) defined on $[0,T]\times\mathbb{R}^n$, which are indexed by $\alpha \in \mathcal{V}_\infty$.

3 The Linear Quadratic (LQ) model

In all subsequent sections, we will study the model (2.1)–(2.2) in an LQ setting. The cases where the finite network is respectively either a discrete ring $\mathbb{T}_{1/N}$ or a discrete torus $\mathbb{T}^2_{1/N}$ will be considered: first, the ring has nodes $\{\alpha_1^N, \cdots, \alpha_{N^*}^N\}$, where $N^* = N$, while, second, the torus network is such that $N^* = N^2$. Let N_1 be the equal subpopulation size at all nodes. Denote the subpopulation average state $x^{(N_1)}_{\alpha_k^N} = \frac{1}{N_1}\sum_{j=1}^{N_1} x^j_{\alpha_k^N}$. The agent $\mathcal{A}^i_{\alpha_k^N}$, $1 \le i \le N_1$, at node α_k^N has dynamics

$$
\begin{aligned}
dx^i_{\alpha_k^N}(t) = {}& [Ax^i_{\alpha_k^N}(t) + Bu^i_{\alpha_k^N}(t)]dt + Dx^{(N_1)}_{\alpha_k^N}(t)dt \\
&+ \kappa N^2 \sum_{k':\alpha_{k'}^N \sim \alpha_k^N} (x^{(N_1)}_{\alpha_{k'}^N} - x^{(N_1)}_{\alpha_k^N})(t)dt + \sigma dw^i_{\alpha_k^N}(t),
\end{aligned} \tag{3.1}
$$

with initial state $x^i_{\alpha_k^N}(0)$, where $N^2\sum_{k':\alpha_{k'}^N \sim \alpha_k^N}(\cdot)$ is called the second order interaction term, involving the subpopulations at immediate neighboring nodes (see Fig. 1 for the ring network case). Here $x^i_{\alpha_k^N}(t)$, $u^i_{\alpha_k^N}(t)$, and $w^i_{\alpha_k^N}(t)$ have dimensions n, n_1, and n_2, respectively, and coefficient $\kappa > 0$. For a matrix $M \ge 0$, denote $\|z\|^2_M = z^T M z$. The cost of agent $\mathcal{A}^i_{\alpha_k^N}$ is

$$
\begin{aligned}
& J^i_{\alpha_k^N}(u^i_{\alpha_k^N}, u^{-i}_{\alpha_k^N}, \mathbf{u}_{-\alpha_k^N}) \\
&= \mathbb{E}\int_0^T \Big[\|x^i_{\alpha_k^N}(t) - \Gamma_1 x^{(N_1)}_{\alpha_k^N}(t) - \Gamma_2\eta\|^2_Q + \|u^i_{\alpha_k^N}(t)\|^2_R\Big]dt \\
&\qquad + \mathbb{E}\|x^i_{\alpha_k^N}(T) - \Gamma_{1f} x^{(N_1)}_{\alpha_k^N}(T) - \Gamma_{2f}\eta\|^2_{Q_f},
\end{aligned} \tag{3.2}
$$

for $R > 0$, $Q \ge 0$, $Q_f \ge 0$, and $\eta \in \mathbb{R}^n$, where $u^{-i}_{\alpha_k^N}$ stands for the strategies of all agents at node α_k^N except $\mathcal{A}^i_{\alpha_k^N}$, $\mathbf{u}_{\alpha^N_{-k}}$ for strategies of all agents not at node α_k^N. For notational simplicity, we consider constant parameters $(A, B, \cdots, \eta, Q_f)$. The extension to time-dependent parameters will be straightforward.

Within the large network limit, we take $\mathbb{T}^\iota$ $(\iota = 1, 2)$ as the node set, and $x_\alpha(t) \in \mathbb{R}^n$ as the state of the representative agent $\mathcal{A}_\alpha$ situated at node $\alpha \in \mathbb{T}^\iota$. At node α, the state average of the subpopulation is replaced by $m(t,\alpha) = \int_{\mathbb{R}^n} y\mu_\alpha(t, dy)$. We take $\phi(y) = \kappa y$, $\kappa > 0$, so that $\bar{\Phi}(t,\alpha)$ in (2.4) becomes $\kappa\Delta m(t,\alpha)$. Suppose all initial states have mean $m_\alpha(0)$ and finite second moment at $t = 0$.

For the best response control problem of agent $\mathcal{A}_\alpha$, we consider the dynamics

$$
dx_\alpha(t) = (Ax_\alpha(t) + Bu_\alpha(t) + Dm(t,\alpha) + \kappa\Delta m(t,\alpha))dt + \sigma dw_\alpha(t). \tag{3.3}
$$

The cost of agent $\mathcal{A}_\alpha$ is given by

$$
\begin{aligned}
J_\alpha = {}& \mathbb{E}\int_0^T \Big(||x_\alpha(t) - \Gamma_1 m(t,\alpha) - \Gamma_2\eta||^2_Q + u_\alpha^T(t) R u_\alpha(t)\Big)dt \\
&+ \mathbb{E}\|x_\alpha(T) - \Gamma_{1f} m(T,\alpha) - \Gamma_{2f}\eta\|^2_{Q_f}.
\end{aligned} \tag{3.4}
$$

Next, we introduce the Riccati ordinary differential equation (ODE):

$$
0 = \dot{P} + PA + A^T P - PBR^{-1}B^T P + Q, \qquad P(T) = Q_f,
$$

which has a unique solution P on $[0, T]$. The best response control law is given by

$$
\hat{u}_\alpha(t) = -R^{-1}B^T[P(t)x_\alpha(t) + S(t,\alpha)], \tag{3.5}
$$

where $S(t,\alpha)$ is given by an ODE when α and the function m are fixed. Subsequently, m is determined by the standard consistency condition of MFGs [23]: the ODE of $m(t,\alpha)$ is obtained by averaging the closed-loop equations for (3.3) of all agents residing at node α. We summarize the LMFG equation system:

$$\begin{aligned}\partial_t S(t,\alpha) = & -(A^T - P(t)BR^{-1}B^T)S(t,\alpha) \\ & -\kappa P(t)\Delta m(t,\alpha) + (Q\Gamma_1 - P(t)D)m(t,\alpha) + Q\Gamma_2\eta, \end{aligned} \tag{3.6}$$

$$\begin{aligned}\partial_t m(t,\alpha) = & \kappa\Delta m(t,\alpha) + (A - BR^{-1}B^T P(t) + D)m(t,\alpha) \\ & - BR^{-1}B^T S(t,\alpha), \qquad 0 < t < T, \quad \kappa > 0,\end{aligned} \tag{3.7}$$

where $\alpha \in \mathbb{T}^\iota$, $m(0,\alpha)$ is given, and $S(T,\alpha) = -Q_f(\Gamma_{1f} m(T,\alpha) + \Gamma_{2f}\eta)$. The equation system is a generalization of the forward-backward ODE system (see e.g. [23, Sec. IV]) in MFGs with symmetric players. The condition $\kappa > 0$ is important in existence analysis since (3.7) becomes a parabolic equation (with space variable α) subject to an initial condition.

4 Existence and uniqueness analysis

To apply results in the literature to the existence analysis for (3.6)–(3.7), we will identify S and m as functions defined on $[0,T] \times \mathbb{R}^\iota$ ($\iota = 1$ or 2) with period 1 in each space variable within α.

For a general analysis of classical solutions, we introduce certain Hölder spaces [19, 29]. Fix $\delta \in (0,1)$ as the Hölder exponent. On $\mathbf{Q} = [0,T] \times \mathbb{R}^k$ with $k \geq 1$, we define the parabolic distance between two points (t,x) and (s,y) as

$$d_p((t,x),(s,y)) = |t-s|^{1/2} + \max_{i\leq k} |x_i - y_i|.$$

Define the Hölder semi-norm for a function $v(t,x)$ with domain $\mathbf{Q}$:

$$[v]_{\delta/2,\delta;\mathbf{Q}} = \sup_{z_i\in\mathbf{Q}, z_1\neq z_2} \frac{|v(z_1) - v(z_2)|}{d_p(z_1,z_2)^\delta}.$$

We allow v to be a vector function. Denote $|v|_{0;\mathbf{Q}} = \sup_{z\in\mathbf{Q}} |v(z)|$. We further define the Hölder norms

$$\begin{aligned}|v|_{\delta/2,\delta;\mathbf{Q}} &= |v|_{0;\mathbf{Q}} + [v]_{\delta/2,\delta;\mathbf{Q}}, \\ |v|_{1+\delta/2,2+\delta;\mathbf{Q}} &= |v|_{0;\mathbf{Q}} + |\partial_t v|_{0;\mathbf{Q}} + \sum_i |\partial_{x_i} v|_{0;\mathbf{Q}} \\ &\quad + \sum_{i,j} |\partial_{x_i x_j} v|_{0;\mathbf{Q}} + [v]_{1+\delta/2,2+\delta;\mathbf{Q}},\end{aligned}$$

where $[v]_{1+\delta/2,2+\delta;\mathbf{Q}} = [\partial_t v]_{\delta/2,\delta;\mathbf{Q}} + \sum_{i,j} [\partial_{x_i x_j} v]_{\delta/2,\delta;\mathbf{Q}}$. We use $C^{\delta/2,\delta}(\mathbf{Q})$ (resp., $C^{1+\delta/2,2+\delta}(\mathbf{Q})$) to denote the space consisting of functions v with $|v|_{\delta/2,\delta;\mathbf{Q}} < \infty$ (resp., $|v|_{1+\delta/2,2+\delta;\mathbf{Q}} < \infty$). The Hölder space $C^{l+\delta}(\mathbb{R}^k)$ for $l = 0,1,2,\cdots$, is similarly defined without involving time variable t.

We give some analytical preparation on linear parabolic equations. Let $q(t,z)$ be an $\mathbb{R}^n$-valued function defined on $\mathbf{Q}$. Consider the following parabolic equation system [29, Theorem 5.1, Theorem 10.2]:

$$\begin{cases} \partial_t q(t,z) = \kappa\Delta q(t,z) + F(t,z)q(t,z) + g(t,z), \\ q(0,z) = \varphi(z), \end{cases} \tag{4.1}$$

where $\kappa > 0$, $\varphi \in C^{2+\delta}(\mathbb{R}^k)$, $F, g \in C^{\delta/2,\delta}(\mathbf{Q})$, and $|F|_{\delta/2,\delta;\mathbf{Q}} \leq C_F$. Here F and g are $\mathbb{R}^{n\times n}$- and $\mathbb{R}^n$-valued, respectively.

The following lemma is a special case of Theorem 10.2 in [29]. For notational brevity, we will omit the underlying domain for the norms.

Lemma 4.1. [29] Suppose $\mathbf{Q} = [0,T] \times \mathbb{R}^k$. Then there exists a unique solution $q \in C^{1+\delta/2,2+\delta}(\mathbf{Q})$ to (4.1) and

$$|q|_{1+\delta/2,2+\delta} \leq C_0(|g|_{\delta/2,\delta} + |\varphi|_{2+\delta}),$$

where the constant C_0 only depends on (κ, C_F, T, δ). □

Remark 4.2. C_0 in Lemma 4.1 can be selected such that it remains valid when T is replaced by any $T' < T$.

For existence analysis, we introduce the following function classes (with $\iota = 1,2$):

(i) $\mathcal{C}^{\delta}_{prd}(\mathbb{R}^\iota)$ consists of functions $f \in C^\delta(\mathbb{R}^\iota)$, having period 1 for each space variable.

(ii) $\mathcal{C}^{2+\delta}_{prd}(\mathbb{R}^\iota)$ consists of functions $f \in C^{2+\delta}(\mathbb{R}^\iota)$, having period 1 for each space variable (thus $\partial_{z_i z_j} f(z) \in C^\delta(\mathbb{R}^\iota)$ for all i,j). $C^{4+\delta}_{prd}(\mathbb{R}^\iota)$ is similarly defined.

(iii) $\mathcal{C}^{\delta/2,\delta}_{prd}([0,T] \times \mathbb{R}^\iota)$ consists of functions $f \in C^{\delta/2,\delta}([0,T] \times \mathbb{R}^\iota)$, having period 1 for each space variable.

(iv) $\mathcal{C}^{1+\delta/2,2+\delta}_{prd}([0,T] \times \mathbb{R}^\iota)$ consists of functions f in $C^{1+\delta/2,2+\delta}([0,T] \times \mathbb{R}^\iota)$, having period 1 for each space variable.

For convenience of subsequent analysis, we introduce another Hölder semi-norm. For two points x and y in $\mathbb{T}^\iota = [0,1)^\iota$, define the distance $d_{\mathbb{T}^\iota}(x,y) = \max_{i \leq \iota} \min\{1 - |x_i - y_i|, |x_i - y_i|\}$. Next, on the set $\mathbf{Q}^\circ = [0,T] \times \mathbb{T}^\iota$, we define the parabolic distance between two points (t,x) and (s,y) as

$$d_p^\circ((t,x),(s,y)) = |t-s|^{1/2} + d_{\mathbb{T}^\iota}(x,y).$$

The Hölder semi-norm for a function $v(t,x)$ defined on $\mathbf{Q}^\circ$ is given by

$$[v]_{\delta/2,\delta;\mathbf{Q}^\circ} = \sup_{z_i \in \mathbf{Q}^\circ, z_1 \neq z_2} \frac{|v(z_1) - v(z_2)|}{d_p^\circ(z_1,z_2)^\delta}.$$

Denote $|v|_{0;\mathbf{Q}^\circ} = \sup_{z \in \mathbf{Q}^\circ} |v(z)|$ and further define the Hölder space $C^{\delta/2,\delta}(\mathbf{Q}^\circ)$.

Lemma 4.3. Suppose $v \in C^{\delta/2,\delta}_{prd}(\mathbf{Q})$ with $\mathbf{Q} = [0,T] \times \mathbb{R}^\iota$. Then we have

$$[v]_{\delta/2,\delta;\mathbf{Q}} = [v]_{\delta/2,\delta;\mathbf{Q}^\circ}. \tag{4.2}$$

Proof. We start by showing $[v]_{\delta/2,\delta;\mathbf{Q}} \leq [v]_{\delta/2,\delta;\mathbf{Q}^\circ}$. Take any two points $(t,x),(t',x') \in \mathbf{Q}$, with $x, x' \in \mathbb{R}^\iota$. Using periodicity, we can find $x'' \in \mathbb{R}^\iota$ such that i) both x'' and x' are identified with the same point in $\mathbb{T}^\iota = [0,1)^\iota$, ii) $v(t',x') = v(t',x'')$, and moreover, iii) $|x_i - x_i''| \leq \min\{|x_i - x_i'|, 1/2\}$ for each $i \leq \iota$. Then we have

$$\frac{|v(t,x) - v(t',x')|}{d_p((t,x),(t',x'))^\delta} \leq \frac{|v(t,x) - v(t',x'')|}{d_p((t,x),(t',x''))^\delta} = \frac{|v(t,x) - v(t',x'')|}{d_p^\circ((t,x),(t',x''))^\delta} \leq [v]_{\delta/2,\delta;\mathbf{Q}^\circ}.$$

Hence $[v]_{\delta/2,\delta;\mathbf{Q}} \leq [v]_{\delta/2,\delta;\mathbf{Q}^\circ}$. We can similarly show $[v]_{\delta/2,\delta;\mathbf{Q}^\circ} \leq [v]_{\delta/2,\delta;\mathbf{Q}}$. □

Suppose $m(0,\cdot) \in C^{2+\delta}_{prd}(\mathbb{R}^\iota)$ in (3.7). We proceed to the existence analysis of (3.6)–(3.7) by formulating the following fixed point problem with $\mathbf{Q} = [0,T] \times \mathbb{R}^\iota$. Given any $\mathbb{R}^n$-valued function $S \in C^{\delta/2,\delta}_{prd}(\mathbf{Q})$ in (3.7), by Lemma 4.1, we obtain a unique solution m and define the following mapping Λ_1:

$$m = \Lambda_1(S),$$

which is well defined from $C^{\delta/2,\delta}_{prd}(\mathbf{Q})$ to $C^{1+\delta/2,2+\delta}(\mathbf{Q})$. By periodicity of $m(0,\cdot)$ and next uniqueness of the solution q in Lemma 4.1, we further have $m \in C^{1+\delta/2,2+\delta}_{prd}(\mathbf{Q})$.

Next, taking any $\mathbb{R}^n$-valued function $m \in C_{prd}^{1+\delta/2,2+\delta}(\mathbf{Q})$ in (3.6), we find a unique solution S by solving a linear ODE for each $\alpha \in \mathbb{R}^\iota$ and define the mapping

$$S = \Lambda_2(m). \tag{4.3}$$

Let $\Phi(t,\tau)$ be the fundamental solution matrix of the ODE $\dot{z}(t) = -(A^T - P(t)BR^{-1}B^T)z(t)$. Denote $|f|_0 = |f|_{0;[0,T]} = \sup_{t\in[0,T]} |f(t)|$. Define the constant

$$\begin{aligned} K_0 =&(T + T^{1-\delta/2} + T^{2-\delta/2}) \sup_{t,\tau\le T} \big[|\Phi(t,\tau)| + |\partial_t \Phi(t,\tau)|\big] \\ &\times (|Q\Gamma_1 - PD|_0 + \kappa |P|_0 + |Q_f \Gamma_{1f}|). \end{aligned}$$

The estimate in the specific form in the next lemma will be useful for identifying a contraction condition later for a fixed point problem.

Lemma 4.4. For $\mathbf{Q} = [0,T] \times \mathbb{R}^\iota$ and $m(0,\cdot) \in C_{prd}^{2+\delta}(\mathbb{R}^\iota)$, Λ_2 in (4.3) is a mapping from $C_{prd}^{1+\delta/2,2+\delta}(\mathbf{Q})$ to $C_{prd}^{\delta/2,\delta}(\mathbf{Q})$ with

$$\begin{aligned} |S|_{\delta/2,\delta;\mathbf{Q}} \le{}& K_0 |m|_{1+\delta/2,2+\delta;\mathbf{Q}} \\ &+ \sup_t |\Phi(t,T)Q_f\Gamma_{1f}| \cdot \Big(|m(0,\cdot)|_{0;\mathbb{R}^\iota} + \sum_i |\partial_{\alpha_i} m(0,\cdot)|_{0;\mathbb{R}^\iota}\Big) \\ &+ T^{1-\delta/2} \sup_t |\partial_t \Phi(t,T)| \cdot |Q_f \Gamma_{2f}\eta| + \sup_{t,\tau} |\Phi(t,\tau)| \cdot \big(T|Q\Gamma_2\eta| + |Q_f\Gamma_{2f}\eta|\big). \end{aligned} \tag{4.4}$$

Proof. See Appendix A. □

Now we introduce the following fixed point equation

$$S = \Lambda_2 \Lambda_1(S), \tag{4.5}$$

which determines a solution of the system (3.6)–(3.7).

Consider $m = \Lambda_1(S)$ and $m' = \Lambda_1(S')$, with the same initial condition $m(0,\alpha) = m'(0,\alpha)$, where S and S' are any functions in $C_{prd}^{\delta/2,\delta}(\mathbf{Q})$. By Lemma 4.1, for some fixed constant C_0, we have

$$|m - m'|_{1+\delta/2,2+\delta;\mathbf{Q}} \le C_0 |BR^{-1}B^T| \cdot |S - S'|_{\delta/2,\delta;\mathbf{Q}}. \tag{4.6}$$

Theorem 4.5. Given $m(0,\cdot) \in C_{prd}^{2+\delta}(\mathbb{R}^\iota)$, the PDE system (3.6)–(3.7) for (S,m) has a unique solution in $C_{prd}^{\delta/2,\delta}(\mathbf{Q}) \times C_{prd}^{1+\delta/2,2+\delta}(\mathbf{Q})$ if $C_0K_0|BR^{-1}B^T| < 1$, where C_0 is given in Lemma 4.1 taking $\mathbf{Q} = [0,T] \times \mathbb{R}^\iota$ and $F(t) = A - BR^{-1}B^TP(t) + D$.

Proof. For any two functions m, m' in $C_{prd}^{1+\delta/2,2+\delta}(\mathbf{Q})$, define $S = \Lambda_2(m)$, $S' = \Lambda_2(m')$. Then we follow the method of establishing (4.4) to similarly obtain

$$|S - S'|_{\delta/2,\delta;\mathbf{Q}} = |\Lambda_2(m) - \Lambda_2(m')|_{\delta/2,\delta;\mathbf{Q}} \le K_0 |m - m'|_{1+\delta/2,2+\delta;\mathbf{Q}}, \tag{4.7}$$

which together with (4.6) yields

$$\begin{aligned} |\Lambda_2\Lambda_1(S) - \Lambda_2\Lambda_1(S')|_{\delta/2,\delta;\mathbf{Q}} &\le K_0 |\Lambda_1(S) - \Lambda_1(S')|_{1+\delta/2,2+\delta;\mathbf{Q}} \\ &\le C_0K_0|BR^{-1}B^T| \cdot |S - S'|_{\delta/2,\delta;\mathbf{Q}}. \end{aligned}$$

We then apply a fixed point theorem on the Banach space $C_{prd}^{\delta/2,\delta}(\mathbf{Q})$. □

Remark 4.6. The contraction condition is satisfied when T is sufficiently small or the system has weak coupling, i.e., $|\Gamma_1| + |\Gamma_{1f}| + |D| + \kappa$ is sufficiently small.

Corollary 4.7. If $m(0,\cdot) \in C^{2+\delta}_{prd}(\mathbb{R}^\iota)$ in Theorem 4.5 is replaced by $m(0,\cdot) \in C^{4+\delta}_{prd}(\mathbb{R}^\iota)$, then we have $S \in C^{0,2}(\mathbf{Q})$ and $m \in C^{1,4}(\mathbf{Q})$, and moreover, S (resp., m) has all up to second order partial derivatives with respect to α belonging to $C^{\delta/2,\delta}_{prd}(\mathbf{Q})$ (resp., $C^{1+\delta/2,2+\delta}_{prd}(\mathbf{Q})$).

Proof. Consider the case of $\iota = 1$ while the case of $\iota = 2$ is similar. We apply a difference quotient approach [19, pp 72–74] by setting $S_h(t,\alpha) = h^{-1}[S(t,\alpha+h) - S(t,\alpha)]$ and $m_h(t,\alpha) = h^{-1}[m(t,\alpha+h) - m(t,\alpha)]$, $h > 0$, to obtain

$$\begin{aligned}\partial_t S_h(t,\alpha) =& -(A^T - P(t)BR^{-1}B^T)S_h(t,\alpha) \\ & - \kappa P(t)\Delta m_h(t,\alpha) + (Q\Gamma_1 - P(t)D)m_h(t,\alpha),\end{aligned} \tag{4.8}$$

$$\begin{aligned}\partial_t m_h(t,\alpha) =& \kappa\Delta m_h(t,\alpha) + (A - BR^{-1}B^T P(t) + D)m_h(t,\alpha) \\ & - BR^{-1}B^T S_h(t,\alpha), \qquad 0 < t < T, \quad \kappa > 0,\end{aligned} \tag{4.9}$$

where $S_h(T,\alpha) = -Q_f\Gamma_{1f}m_h(T,\alpha)$, with initial condition $m_h(0,\cdot)$. It is easy to show $\sup_h |m_h(0,\cdot)|_{2+\delta;\mathbb{R}^\iota} \leq C$. Under the contraction condition of Theorem 4.5, (S_h, m_h) is the unique solution of (4.8)–(4.9). In view of Lemma 4.1 and (4.4), we can show

$$\sup_h \left(|S_h|_{\delta/2,\delta;\mathbf{Q}} + |m_h|_{1+\delta/2,2+\delta;\mathbf{Q}}\right) \leq C.$$

By a standard subsequence argument [19, pp 72–74], we see that $(\partial_\alpha S, \partial_\alpha m)$ is a solution of (4.8)–(4.9) with initial condition $\partial_\alpha m(0,\alpha)$. Repeating this argument, we further obtain $(\partial^2_\alpha S, \partial^2_\alpha m)$ as a solution to (4.8)–(4.9) with initial condition $\partial^2_\alpha m(0,\alpha)$, so that $(\partial^2_\alpha S, \partial^2_\alpha m) \in C^{\delta/2,\delta}_{prd}(\mathbf{Q}) \times C^{1+\delta/2,2+\delta}_{prd}(\mathbf{Q})$. $\square$

4.1 Numerical example

Example 4.8. In the LQ model (3.1)–(3.2) we consider a ring network with $A = 0.5$, $B = 1$, $D = 0.2$, $\kappa = 0.5$, $Q = Q_f = 2$, $R = 1$, $\Gamma_1 = 0.8$, $\Gamma_2 = 0.2$, $\eta = 1$, $\Gamma_{1f} = 0.6$, $\Gamma_{2f} = 0.4$, $T = 2$, and $m(0,\alpha) = \psi(\alpha) := 256\alpha^4(1-\alpha)^4$, $0 \leq \alpha \leq 1$.

If we extend ψ on $\mathbb{R}$ with period 1, then it can be shown that $\psi \in C^{4+1/2}_{prd}(\mathbb{R})$. The system (3.6)–(3.7) is solved by a difference scheme, with stepwise $\Delta t = 0.0025$ and $\Delta\alpha = 0.05$. The numerical solution of (S,m) is displayed in Fig. 2.

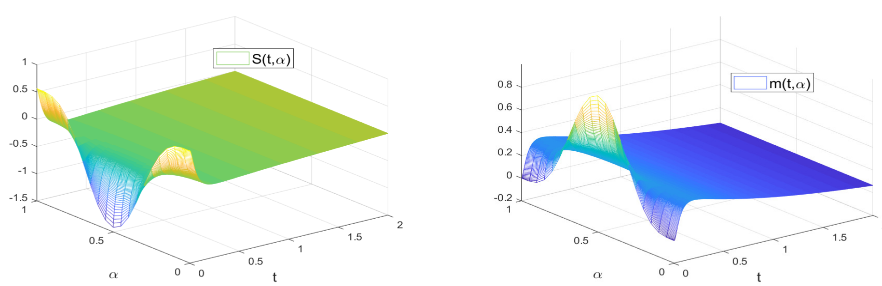

Figure 2: $S(t,\alpha)$ **and** $m(t,\alpha)$ **in** (3.6)–(3.7) **computed by iterations with** $t \in [0,2]$ **and** $\alpha \in [0,1]$.

5 Closed-loop analysis of the LQ model

In this section, the cases of finite networks $\mathbb{T}_{1/N}$ and $\mathbb{T}^2_{1/N}$ are considered. Let the LQ model be specified by (3.1)–(3.2). For tractability of the subsequent estimates, all remaining sections consider scalar states, control and Brownian motions in (3.1)–(3.2) so that $n = n_1 = n_2 = 1$.

We introduce the following assumptions (H1)–(H4), which are in effect throughout Sections 5, 6 and 7.

(H1) For a given function $m_0(\alpha)$ from $[0,1)^\iota$ to $\mathbb{R}$, we have $x^i_{\alpha^N_k}(0) = m_0(\alpha^N_k)$, for all i, k. The periodic extension of $m_0(\cdot)$ to the domain $\mathbb{R}^\iota$, denoted by $\widetilde{m}_0$, belongs to $C^{4+\delta}_{prd}(\mathbb{R}^\iota)$ for some $\delta \in (0,1)$.

(H2) There exists a unique solution (S, m) to (3.6)–(3.7) with $n = 1$ on $\mathbf{Q} = [0,T] \times \mathbb{R}^\iota$, such that $S \in C^{\delta/2,\delta}_{prd}(\mathbf{Q})$, $m \in C^{1+\delta/2,2+\delta}_{prd}(\mathbf{Q})$, $m(0,\alpha) = \widetilde{m}_0(\alpha)$.

(H3) The solution (S, m) to (3.6)–(3.7) is such that the 2nd (resp., 4th) order partial derivatives of $S(t,\alpha)$ (resp., $m(t,\alpha)$) with respect to α are continuous in $(t,\alpha) \in [0,T] \times \mathbb{R}^\iota$.

(H4) $\lim_{N\to\infty} N_1/N^3 = \infty$ for the ring network $\mathbb{T}_{1/N}$; $\lim_{N\to\infty} N_1/(N^2 \ln N) = \infty$ for the torus network $\mathbb{T}^2_{1/N}$.

Remark 5.1. (i) We assume equal initial states within the same cluster to streamline the presentation, and can deal with random initial states easily. (ii) For higher order regularity of the solution in (H3), Corollary 4.7 provides sufficient conditions. (iii) Since the second order interaction terms in (3.1) are random and scaled up by N^2, we need to average with fast growth of N_1 as in (H4) to reduce their fluctuations.

By assigning the agents' actual nodal positions to α in (3.5), the LMFG-based decentralized strategies for the finite population take the form

$$\hat{u}^i_{\alpha^N_k}(t) = -\, R^{-1}B[P(t)x^i_{\alpha^N_k}(t) + S(t,\alpha^N_k)], \quad 1 \le i \le N_1, \quad 1 \le k \le N^*. \tag{5.1}$$

In further analysis, for notational brevity, we will often drop the time argument in a process. For instance, we simply write $x^i_{\alpha^N_k}(t)$ as $x^i_{\alpha^N_k}$. The strategies in (5.1) yield the closed-loop dynamics:

$$\begin{aligned}
&\text{System (A)}: \\
&\quad d\hat{x}^i_{\alpha^N_k} = \{[A - B^2R^{-1}P(t)]\hat{x}^i_{\alpha^N_k} + D\hat{x}^{(N_1)}_{\alpha^N_k} - B^2R^{-1}S(t,\alpha^N_k)\}dt \\
&\qquad + \kappa N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} \big(\hat{x}^{(N_1)}_{\alpha^N_{k'}} - \hat{x}^{(N_1)}_{\alpha^N_k}\big)dt + \sigma dw^i_{\alpha^N_k}, \\
&\qquad\quad 1 \le i \le N_1, \quad 1 \le k \le N^*,
\end{aligned} \tag{5.2}$$

with $\hat{x}^i_{\alpha^N_k}(0) = x^i_{\alpha^N_k}(0)$ and $\hat{x}^{(N_1)}_{\alpha^N_k} = \frac{1}{N_1}\sum_{j=1}^{N_1} \hat{x}^j_{\alpha^N_k}$.

In the ϵ-Nash equilibrium analysis, only one agent applies a different strategy while all the other agents stick to the strategies in (5.1). Without loss of generality, we consider the strategy change of agent $\mathcal{A}^1_{\alpha^N_1}$ within the cluster residing at node α^N_1, and the system dynamics are given by

$$\begin{aligned}
&\text{System (B)}: \\
&\quad dx^1_{\alpha^N_1} = [Ax^1_{\alpha^N_1} + Bu^1_{\alpha^N_1} + Dx^{(N_1)}_{\alpha^N_1}]dt \\
&\qquad\qquad + \kappa N^2 \sum_{k:\alpha^N_k\sim\alpha^N_1} (x^{(N_1)}_{\alpha^N_k} - x^{(N_1)}_{\alpha^N_1})dt + \sigma dw^1_{\alpha^N_1}, \\
&\quad dx^i_{\alpha^N_k} = [Ax^i_{\alpha^N_k} + B\hat{u}^i_{\alpha^N_k} + Dx^{(N_1)}_{\alpha^N_k}]dt \\
&\qquad\qquad + \kappa N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} (x^{(N_1)}_{\alpha^N_{k'}} - x^{(N_1)}_{\alpha^N_k})dt + \sigma dw^i_{\alpha^N_k},
\end{aligned}$$

for $i \in \{1, \cdots, N_1\}$ if $k \neq 1$; and $i \in \{2, \cdots, N_1\}$ if $k = 1$.

The next theorem gives the a priori upper bound of the cost that is achieved by the decentralized strategies in (5.1).

Theorem 5.2. There exists a constant $\hat{C}_0$ such that

$$\sup_{N,N_1}\max_{k\le N^*,i\le N_1} J^i_{\alpha^N_k}(\hat{u}^i_{\alpha^N_k},\hat{u}^{-i}_{\alpha^N_k},\hat{\mathbf{u}}_{\alpha^N_{-k}})\le \hat{C}_0.$$

We postpone its proof, which is technical and needs a series of preparations. For the ϵ-Nash equilibrium analysis, Theorem 5.2 provides a benchmark performance level when one agent unilaterally applies an alternative strategy $u^i_{\alpha^N_k}$ for its own improvement. In this case it suffices to consider $u^i_{\alpha^N_k}$ satisfying

$$J^i_{\alpha^N_k}(u^i_{\alpha^N_k},\hat{u}^{-i}_{\alpha^N_k},\hat{\mathbf{u}}_{\alpha^N_{-k}})\le \hat{C}_0. \tag{5.3}$$

The restriction to such strategies $u^i_{\alpha^N_k}$ will allow us to obtain uniform error bounds when approximating $x^{(N_1)}_{\alpha^N_k}$ and $N^2\sum_{k':\alpha^N_{k'}\sim\alpha^N_k}(x^{(N_1)}_{\alpha^N_{k'}}-x^{(N_1)}_{\alpha^N_k})$, $1\le k\le N^*$, generated by $(u^i_{\alpha^N_k},\hat{u}^{-i}_{\alpha^N_k},\hat{\mathbf{u}}_{\alpha^N_{-k}})$.

By (5.2), we have

$$\begin{aligned} d\hat{x}^{(N_1)}_{\alpha^N_k} &= [A-B^2R^{-1}P(t)+D]\hat{x}^{(N_1)}_{\alpha^N_k}dt - B^2R^{-1}S(t,\alpha^N_k)dt \\ &\quad + \kappa N^2\sum_{k':\alpha^N_{k'}\sim\alpha^N_k}(\hat{x}^{(N_1)}_{\alpha^N_{k'}}-\hat{x}^{(N_1)}_{\alpha^N_k})dt+\sigma dw^{(N_1)}_{\alpha^N_k}, \end{aligned} \tag{5.4}$$

where $w^{(N_1)}_{\alpha^N_k}=\frac{1}{N_1}\sum_{j=1}^{N_1}w^j_{\alpha^N_k}$.

To find the cost bound in Theorem 5.2, a key step is to estimate the following quantities in System (A):

$$\sup_{N,N_1}\sup_{k,t}\mathbb{E}|\hat{x}^{(N_1)}_{\alpha^N_k}(t)|^2, \qquad \sup_{N,N_1}\sup_{k,t}N^2\mathbb{E}\Big|\sum_{k':\alpha^N_{k'}\sim\alpha^N_k}(\hat{x}^{(N_1)}_{\alpha^N_{k'}}-\hat{x}^{(N_1)}_{\alpha^N_k})(t)\Big|^2. \tag{5.5}$$

As it turns out below, the resulting analysis is much more difficult than in GMFGs with dense networks, which is due to the second order interaction term.

To facilitate the estimate of $\hat{x}^{(N_1)}_{\alpha^N_k}$, we introduce an auxiliary deterministic system:

$$\begin{aligned} &\text{System (C)}: \\ &\frac{d}{dt}\bar{x}_{\alpha^N_k} =[A-B^2R^{-1}P(t)+D]\bar{x}_{\alpha^N_k}-B^2R^{-1}S(t,\alpha^N_k) \\ &\qquad\qquad + \kappa N^2\sum_{k':\alpha^N_{k'}\sim\alpha^N_k}(\bar{x}_{\alpha^N_{k'}}-\bar{x}_{\alpha^N_k}), \quad 1\le k\le N^*, \end{aligned} \tag{5.6}$$

where $\bar{x}_{\alpha^N_k}(0)=m_0(\alpha^N_k)$.

Lemma 5.3. There exists a constant C such that $\sup_{N,k\le N^*,t\le T}|\bar{x}_{\alpha^N_k}(t)|\le C$.

Proof. For $t\in[0,T]$, define $x^N_{\max}(t)=\max_{1\le k\le N^*}\bar{x}_{\alpha^N_k}(t)$ which is then absolute continuous and so differentiable almost everywhere on $[0,T]$. If $x^N_{\max}(t)$ is differentiable at t, denote the set $\mathbf{s}^N_t=\{k: x^N_{\max}(t)=\bar{x}_{\alpha^N_k}(t)\}$. Consider the sequence $\{x^N_{\max}(t+1/j), j\ge j_0\}$ for a sufficiently large j_0 if $t<T$. There necessarily exists some $\hat{k}\in\mathbf{s}^N_t$ such that $x^N_{\max}(t+1/j)=\bar{x}_{\alpha^N_{\hat{k}}}(t+1/j)$ for an infinite number of values of j (we replace $t+1/j$ by $t-1/j$ if $t=T$). Therefore, we have $\frac{d}{dt}x^N_{\max}(t)=\frac{d}{dt}\bar{x}_{\alpha^N_{\hat{k}}}(t)$, which further implies that whenever $x^N_{\max}(t)$ is positive, its derivative, if existing, is at most $|A-B^2R^{-1}P(t)+D|x^N_{\max}(t)+c_1$ for $c_1:=B^2R^{-1}\sup_{\alpha,t}|S(t,\alpha)|$ since $\kappa>0$. To bound $x^N_{\max}(t)$ on $[0,T]$ from above, we construct the auxiliary ODE:

$$\dot{y}(t)=|A-B^2R^{-1}P(t)+D|y(t)+c_1, \quad y(0)=\max_{\alpha\in[0,1)^\iota}|m_0(\alpha)|. \tag{5.7}$$

Then by Lemma A.1 (i), we have $x^N_{\max}(t) \le y(t)$ for all $t \in [0,T]$.

Similarly, we apply Lemma A.1 (ii) to obtain a constant lower bound for $x^N_{\min}(t) \coloneqq \min_{1\le k\le N} \bar{x}_{\alpha^N_k}(t)$. This completes the proof. $\square$

Lemma 5.4. We have $\sup_{k\le N^*, t\le T} |\bar{x}_{\alpha^N_k}(t) - m(t,\alpha^N_k)| = o(1)$ as $N\to\infty$.

Proof. Taking $\alpha = \alpha^N_k$ in (3.7), we write the equation as an ODE:

$$\begin{aligned}\frac{d}{dt} m(t,\alpha^N_k) &= [A - B^2R^{-1}P(t) + D]m(t,\alpha^N_k) - B^2R^{-1}S(t,\alpha^N_k)\\ &\quad + \kappa N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} [m(t,\alpha^N_{k'}) - m(t,\alpha^N_k)] + \epsilon_{\alpha^N_k}(t),\end{aligned} \tag{5.8}$$

where

$$\epsilon_{\alpha^N_k}(t) = \kappa\Delta m(t,\alpha^N_k) - \kappa N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} [m(t,\alpha^N_{k'}) - m(t,\alpha^N_k)],$$

and $\sup_{k,t}|\epsilon_{\alpha^N_k}(t)| \to 0$ as $N\to\infty$ since $m \in C^{1+\delta/2, 2+\delta}_{prd}([0,T]\times\mathbb{R}^\iota)$. Denote $\zeta_{\alpha^N_k}(t) = \bar{x}_{\alpha^N_k}(t) - m(t,\alpha^N_k)$. By (5.6) and (5.8), we have

$$\begin{aligned}\frac{d}{dt}\zeta_{\alpha^N_k}(t) &= [A - B^2R^{-1}P(t) + D]\zeta_{\alpha^N_k}(t)\\ &\quad + \kappa N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} [\zeta_{\alpha^N_{k'}}(t) - \zeta_{\alpha^N_k}(t)] - \epsilon_{\alpha^N_k}(t), \quad k\le N^*.\end{aligned}$$

Following the method in proving Lemma 5.3, we construct the auxiliary ODE

$$\dot{y}(t) = |A - B^2R^{-1}P(t)|y(t) + \sup_{t,k}|\epsilon_{\alpha^N_k}(t)|, \quad y(0)=0,$$

and obtain an upper bound of the form $\max_{t,k}\zeta_{\alpha^N_k}(t) \le C\sup_{t,k}|\epsilon_{\alpha^N_k}(t)|$. Similarly we obtain a lower bound for $\min_{t,k}\zeta_{\alpha^N_k}(t)$, and conclude that $\max_{t,k}|\zeta_{\alpha^N_k}(t)| \le C\sup_{t,k}|\epsilon_{\alpha^N_k}(t)| = o(1)$ as $N\to\infty$. $\square$

Denote

$$\check{x}_{\alpha^N_k}(t) = \hat{x}^{(N_1)}_{\alpha^N_k}(t) - \bar{x}_{\alpha^N_k}(t), \qquad k\le N^*, \quad 0\le t\le T. \tag{5.9}$$

To obtain bounds for (5.5), we start by analyzing $\check{x}_{\alpha^N_k}(t)$, $1\le k\le N^*$, which will be stacked into a large vector by the following rule:

(i) For the ring network $\mathbb{T}_{1/N}$, define

$$\check{\mathbf{x}} = [\check{x}_{\alpha^N_1}, \cdots, \check{x}_{\alpha^N_N}]^T, \quad \mathbf{w} = [w^{(N_1)}_{\alpha^N_1}, \cdots, w^{(N_1)}_{\alpha^N_N}]^T, \tag{5.10}$$

where $\alpha^N_k = \frac{k-1}{N} + \frac{1}{2N}$, $1\le k\le N$ (see Section 2.5).

(ii) For the torus network $\mathbb{T}^2_{1/N}$, let the N^2 nodes be arranged into an $N\times N$ array. The entry at the ith row and jth column, to be called the (i,j)-th node, is $(\frac{i-1}{N} + \frac{1}{2N}, \frac{j-1}{N} + \frac{1}{2N})$, $1\le i,j\le N$. Each node has 4 neighbors. For instance, the $(1,1)$-th node has neighbors: the $(1,2)$-th node, the $(1,N)$-th node, the $(2,1)$-th node, and the $(N,1)$-th node. We list the N^2 nodes into a single column, starting with the first row, next the second row, and so forth until the last row. Following the above order, we now list the N^2 terms $\check{x}_{\alpha^N_k}$ (resp., $w^{(N_1)}_{\alpha^N_k}$) into a large vector denoted by $\check{\mathbf{x}}$ (resp., $\mathbf{w}$):

$$\check{\mathbf{x}} = [\check{x}_{\alpha^N_1}, \cdots, \check{x}_{\alpha^N_{N^2}}]^T, \quad \mathbf{w} = [w^{(N_1)}_{\alpha^N_1}, \cdots, w^{(N_1)}_{\alpha^N_{N^2}}]^T, \tag{5.11}$$

where α^N_1 and $\alpha^N_{N^2}$ are the $(1,1)$-th and the (N,N)-th nodes, respectively.

Our subsequent analysis will give details for the torus case, and the ring case can be similarly handled and is simpler. For the error estimate, we will derive an equation for $\check{\mathbf{x}}$. As a preparation, denote the $N \times N$ circulate matrix [15]

$$\mathcal{R}_N = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix},$$

which is symmetric and is the Laplacian matrix of the ring network $\mathbb{T}_{1/N}$.

For the ring network $\mathbb{T}_{1/N}$, letting $\hat{A}(t) = A - B^2R^{-1}P(t) + D$ and $\hat{A}_N(t) = \hat{A}(t) - 2\kappa N^2$, we introduce the following circulate matrix

$$\begin{aligned} \mathcal{M}_N^{rn}(t) &= \begin{bmatrix} \hat{A}_N & \kappa N^2 & 0 & \cdots & 0 & \kappa N^2 \\ \kappa N^2 & \hat{A}_N & \kappa N^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \kappa N^2 & 0 & 0 & \cdots & \kappa N^2 & \hat{A}_N \end{bmatrix} \\ &= \hat{A}(t) I_N - 2\kappa N^2 I_N + \kappa N^2 \mathcal{R}_N. \end{aligned} \tag{5.12}$$

where each row has 3 non-zero entries.

To analyze $\check{\mathbf{x}}$ for the torus network $\mathbb{T}^2_{1/N}$, define

$$\mathcal{T}_N = I_N \otimes \mathcal{R}_N + \mathcal{R}_N \otimes I_N, \tag{5.13}$$

which is the Laplacian of $\mathbb{T}^2_{1/N}$. An example of $\mathcal{T}_N$ with $N = 4$ is illustrated in Appendix B. Following the definition in [22], the right hand side of (5.13) is the Kronecker sum $\mathcal{R}_N \oplus \mathcal{R}_N$. We further denote

$$\mathcal{M}_N^{tr}(t) = I_{N^2} \otimes (\hat{A}(t) - 4\kappa N^2) + \kappa N^2 \mathcal{T}_N. \tag{5.14}$$

So $\mathcal{M}_N^{tr}$ is a symmetric matrix with dimension $N^2 \times N^2$.

We have the following representation of the process $\check{\mathbf{x}}$:

Lemma 5.5. Let $\check{x}_{\alpha_k^N}(t)$ be defined by (5.9). Then

(i) for ring network $\mathbb{T}_{1/N}$ and (5.10), we have

$$d\check{\mathbf{x}}(t) = \mathcal{M}_N^{rn} \check{\mathbf{x}}(t)dt + \sigma d\mathbf{w}(t), \tag{5.15}$$

with initial condition $\check{x}_{\alpha_k^N}(0)$, $1 \le k \le N$;

(ii) for the torus network $\mathbb{T}^2_{1/N}$ and (5.11), we have

$$d\check{\mathbf{x}}(t) = \mathcal{M}_N^{tr} \check{\mathbf{x}}(t)dt + \sigma d\mathbf{w}(t), \tag{5.16}$$

with initial condition $\check{x}_{\alpha_k^N}(0)$, $1 \le k \le N^2$.

Proof. By (5.4) and (5.6), it is straightforward to show

$$\begin{aligned} d\check{x}_{\alpha_k^N} =& [A - B^2R^{-1}P(t) + D]\check{x}_{\alpha_k^N} dt + \kappa N^2 \sum_{k': \alpha_{k'}^N \sim \alpha_k^N} (\check{x}_{\alpha_{k'}^N} - \check{x}_{\alpha_k^N})dt \\ &+ \sigma dw^{(N_1)}_{\alpha_k^N}, \quad 1 \le k \le N^*, \end{aligned} \tag{5.17}$$

where $\check{x}_{\alpha_k^N}(0) = 0$. Arranging $\check{x}_{\alpha_k^N}$ into a vector as in (5.10)–(5.11), and checking the coefficient matrix of each case, we derive the SDEs in (i) and (ii). $\square$

Define the polynomial $\varphi(z) = z + z^{N-1}$ for $\mathcal{R}_N$. By the results on general circulate matrices (see [15]), the eigenvalues of $\mathcal{R}_N$ are given by

$$\lambda_k = \varphi(e^{\mathbf{i}2\pi(k-1)/N}) = 2\cos\frac{2\pi(k-1)}{N}, \quad \mathbf{i}^2 = -1,\ 1 \le k \le N. \tag{5.18}$$

Denote $\omega_k = e^{\mathbf{i}2\pi(k-1)/N}$, $1 \le k \le N$, and the Fourier matrix

$$F = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega_N & \omega_N^2 & \cdots & \omega_N^{N-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \omega_N^{N-1} & \omega_N^{2(N-1)} & \cdots & \omega_N^{(N-1)^2} \end{bmatrix}.$$

By results on circulate matrices [15], F is unitary (i.e., $F^*F = I$) and

$$F^*\mathcal{R}_N F = \mathrm{diag}[\lambda_1, \cdots, \lambda_N], \tag{5.19}$$

where F^* is the conjugate transpose of F. Denote

$$\lambda_{jk} = \lambda_j + \lambda_k = 2\left(\cos\frac{2\pi(j-1)}{N} + \cos\frac{2\pi(k-1)}{N}\right), \quad 1 \le j,k \le N,$$

and the $N^2 \times N^2$ matrix

$$\Lambda_0 = \mathrm{diag}[\lambda_{11}, \cdots, \lambda_{1N}, \lambda_{21}, \cdots, \lambda_{2N}, \cdots, \lambda_{N1}, \cdots, \lambda_{NN}].$$

Lemma 5.6. The following properties hold:

(i) the matrix $\mathcal{M}_N^{rn}(t)$ in (5.12) is symmetric and $\mathcal{M}_N^{rn}(t)\mathcal{M}_N^{rn}(s) = \mathcal{M}_N^{rn}(s)\mathcal{M}_N^{rn}(t)$ for all $t, s \in [0,T]$;
(ii) $\mathcal{T}_N^T = \mathcal{T}_N$ for $\mathcal{T}_N$ defined by (5.13);
(iii) the matrix $\mathcal{M}_N^{tr}(t)$ in (5.14) is symmetric and $\mathcal{M}_N^{tr}(t)\mathcal{M}_N^{tr}(s) = \mathcal{M}_N^{tr}(s)\mathcal{M}_N^{tr}(t)$ for all $t, s \in [0,T]$;
(iv) we have $(F\otimes F)^*\mathcal{M}_N^{tr}(t)(F\otimes F) = (\hat{A}(t) - 4\kappa N^2)I_{N^2} + \kappa N^2\Lambda_0$, and $F\otimes F$ is a unitary matrix.

Proof. (i) holds since $\mathcal{R}_N$ is symmetric. Using the property $(M_1\otimes M_2)^T = M_1^T \otimes M_2^T$ for any two matrices M_1 and M_2, we obtain (ii), which further implies (iii). By basic properties of Kronecker products, it is easily shown that $F\otimes F$ is a unitary matrix. By [22, Theorem 4.4.5], the N^2 eigenvalues λ_{jk} (including algebraic multiplicities) of $\mathcal{T}_N$ are given by Λ_0, and we use properties of Kronecker products (see [22, Lemma 4.2.10]) and (5.19) to calculate

$$\begin{aligned}(F\otimes F)^*\mathcal{T}_N(F\otimes F) &= (F^*\otimes F^*)(I_N\otimes \mathcal{R}_N)(F\otimes F) \\ &\quad + (F^*\otimes F^*)(\mathcal{R}_N\otimes I_N)(F\otimes F) \\ &= (F^*F)\otimes(F^*\mathcal{R}_N F) + (F^*\mathcal{R}_N F)\otimes(F^*F) \\ &= \Lambda_0,\end{aligned}$$

which leads to (iv). □

By Lemma 5.6 (iv), all N^2 eigenvalues of $\mathcal{M}_N^{tr}(t)$ are given by

$$\hat{A}(t) - 2\kappa N^2\left(2 - \cos\frac{2\pi(j-1)}{N} - \cos\frac{2\pi(k-1)}{N}\right), \quad 1 \le j,k \le N. \tag{5.20}$$

Consider the ODEs

$$\dot{\mathbf{x}} = \mathcal{M}_N^{rn}(t)\mathbf{x}, \qquad \dot{\mathbf{x}} = \mathcal{M}_N^{tr}(t)\mathbf{x}, \tag{5.21}$$

where the dimension of $\mathbf{x}$ is compatible with the coefficient matrix. In view of Lemma 5.6 (i) and (iii), we apply Lemma A.2 to get the fundamental solution matrices

$$\Phi^{rn}(t,s) = e^{\int_s^t \mathcal{M}_N^{rn}(\tau)d\tau}, \quad \Phi^{tr}(t,s) = e^{\int_s^t \mathcal{M}_N^{tr}(\tau)d\tau}, \quad t,s \in [0,T], \tag{5.22}$$

which are symmetric for all given $t,s \in [0,T]$. We give some estimate for (5.9).

Lemma 5.7. (i) For the ring network $\mathbb{T}_{1/N}$, we have

$$\sup_{t,k} \mathbb{E}|\check{x}_{\alpha_k^N}(t)|^2 \le C/(NN_1).$$

(ii) For the torus network $\mathbb{T}^2_{1/N}$, we have

$$\sup_{t,k} \mathbb{E}|\check{x}_{\alpha_k^N}(t)|^2 \le C\ln N/(N^2N_1).$$

Proof. (i) For $\check{\mathbf{x}}$ given by (5.10), we have $\check{\mathbf{x}}(t) = \sigma\int_0^t \Phi^{rn}(t,\tau)d\mathbf{w}(\tau)$, so that

$$\mathbb{E}|\check{\mathbf{x}}(t)|^2 = \frac{\sigma^2}{N_1}\mathrm{Tr}\int_0^t e^{2\int_\tau^t \mathcal{M}_N^{rn}(r)dr}d\tau, \tag{5.23}$$

by symmetry of $\Phi^{rn}(t,\tau)$. By digitalization using (5.19), we have

$$\mathrm{Tr}\int_0^t e^{2\int_\tau^t \mathcal{M}_N^{tr}(r)dr}d\tau \le C\Big(1 + \sum_{1\le i\le N-1}\frac{1}{N^2(1-\cos\frac{2\pi i}{N})}\Big). \tag{5.24}$$

We estimate the sum $s_N := \sum_{k=1}^{N-1} N^{-2}\left(1-\cos\frac{2\pi k}{N}\right)^{-1}$. There exists a fixed small $\theta_0 > 0$ such that $1-\cos\theta \ge \theta^2/4$ for all $|\theta| \le \theta_0$. We write

$$s_N = \sum_{k:0<2\pi k/N\le\theta_0} + \sum_{k:\theta_0<2\pi k/N<2\pi-\theta_0} + \sum_{k:2\pi-\theta_0\le 2\pi k/N<2\pi} =: s_{N,1} + s_{N,2} + s_{N,3}, \tag{5.25}$$

where the summands are obvious and not displayed. We obtain

$$0 \le s_{N,1} \le \sum_{k:0<2\pi k/N\le\theta_0} 4/(2\pi k)^2 = O(1), \tag{5.26}$$

and similarly obtain $s_{N,3} = O(1)$ while clearly $s_{N,2} = O(1)$. Hence, $s_N = O(1)$ and $\mathbb{E}|\check{\mathbf{x}}(t)|^2 \le C/N_1$. By symmetry, we get the bound of each component.

(ii) In analogue to (5.23), we have

$$\mathbb{E}|\check{\mathbf{x}}(t)|^2 = \frac{\sigma^2}{N_1}\mathrm{Tr}\int_0^t e^{2\int_\tau^t \mathcal{M}_N^{tr}(r)dr}d\tau.$$

We have

$$\begin{aligned}
&\mathrm{Tr}\int_0^t e^{2\int_\tau^t \mathcal{M}_N^{tr}(r)dr}d\tau \qquad (5.27)\\
&= \mathrm{Tr}\int_0^t e^{2\int_\tau^t[\kappa N^2\Lambda_0 + (\hat{A}(r)-4\kappa N^2)I_{N^2}]dr}d\tau\\
&= \int_0^t \sum_{1\le i,j\le N} e^{2(t-\tau)[(\lambda_i+\lambda_j-4)\kappa N^2]}e^{2\int_\tau^t \hat{A}(r)dr}d\tau\\
&\le CT + C\sum_{i+j\ge1, 0\le i,j\le N-1}\frac{1}{N^2[(1-\cos\frac{2\pi i}{N}) + (1-\cos\frac{2\pi j}{N})]}
\end{aligned}$$

$$= O(\ln N),$$

where we have obtained the last line using the method in (5.25) and a double integral estimate. Finally for some constant C, we get $\mathbb{E}|\check{\mathbf{x}}(t)|^2 \le C \ln N/N_1$. By symmetry, we obtain the upper bound for each component. □

Lemma 5.8. For some constant C, we have

$$\sup_{N,N_1} \sup_{t,k} \mathbb{E}|\hat{x}^{(N_1)}_{\alpha^N_k}(t)|^2 \le C.$$

Proof. This lemma follows from (5.9), Lemmas 5.7 and 5.3. □

6 Estimate of the second order interaction term

This section analyzes the torus network $\mathbb{T}^2_{1/N}$ with detail. For α^N_k in the array of N^2 nodes of $\mathbb{T}^2_{1/N}$ (see Sec. 5), let $\alpha^N_{k_+}$ and $\alpha^N_{k_-}$ denote its two neighbors along the horizontal direction. Our next step is to approximate $\hat{x}^{(N_1)}_{\alpha^N_{k_+}} + \hat{x}^{(N_1)}_{\alpha^N_{k_-}} - 2\hat{x}^{(N_1)}_{\alpha^N_k}$ contained in (5.2) by the deterministic term $\bar{x}_{\alpha^N_{k_+}} + \bar{x}_{\alpha^N_{k_-}} - 2\bar{x}_{\alpha^N_k}$. We have the relation

$$\begin{aligned} &\hat{x}^{(N_1)}_{\alpha^N_{k_+}} + \hat{x}^{(N_1)}_{\alpha^N_{k_-}} - 2\hat{x}^{(N_1)}_{\alpha^N_k} - (\bar{x}_{\alpha^N_{k_+}} + \bar{x}_{\alpha^N_{k_-}} - 2\bar{x}_{\alpha^N_k}) \\ &= \check{x}_{\alpha^N_{k_+}} + \check{x}_{\alpha^N_{k_-}} - 2\check{x}_{\alpha^N_k}, \end{aligned}$$

where $\check{x}_{\alpha^N_k}$ is defined by (5.9). Now we set $v_{\alpha^N_k} = \check{x}_{\alpha^N_{k_+}} + \check{x}_{\alpha^N_{k_-}} - 2\check{x}_{\alpha^N_k}$ and $w^{(N_1)}_{\alpha^N_k} = \frac{1}{N_1}\sum_{i=1}^{N_1} w^i_{\alpha^N_k}$. By SDE (5.17) for $\check{x}_{\alpha^N_k}$, we have

$$\begin{aligned} dv_{\alpha^N_k} &= [A - B^2R^{-1}P(t) + D]v_{\alpha^N_k}dt + \kappa N^2 \sum_{j:\alpha^N_j \sim \alpha^N_k} (v_{\alpha^N_j} - v_{\alpha^N_k})dt \\ &\quad + \sigma dw^{(N_1)}_{\alpha^N_{k_+}} + \sigma dw^{(N_1)}_{\alpha^N_{k_-}} - 2\sigma dw^{(N_1)}_{\alpha^N_k}, \qquad 1 \le k \le N^2. \end{aligned}$$

Denote $\mathbf{v}(t) = [v_{\alpha^N_1}, \cdots, v_{\alpha^N_{N^2}}]^T$, which lists the N^2 entries by the same rule as in (5.11), and $\mathcal{D} = I_N \otimes \mathcal{R}_N - 2I_{N^2}$. We can show

$$d\mathbf{v}(t) = \mathcal{M}^{tr}_N \mathbf{v}(t)dt + \sigma \mathcal{D} d\mathbf{w}(t). \tag{6.1}$$

Using (6.1) we obtain the following lemma.

Lemma 6.1. We have $\sup_{k,0\le t\le T} E|v_{\alpha^N_k}(t)|^2 \le C \ln N/(N^2N_1)$.

Proof. The matrix $\mathcal{D}^2$ has all its eigenvalues (including algebraic multiplicities) given by the N values $(2-\lambda_1)^2, \cdots, (2-\lambda_N)^2$, each repeated N times (see [22, Theorem 4.2.12]), where λ_i is defined by (5.18). Hence $\mathcal{D}^2 \le 16 I_{N^2}$. By $\mathbf{v}(0) = 0$, we have $\mathbf{v}(t) = \sigma \int_0^t \Phi^{tr}(t,\tau)\mathcal{D}d\mathbf{w}(\tau)$, and

$$\mathbb{E}|\mathbf{v}(t)|^2 \le (16\sigma^2/N_1)\mathrm{Tr}\int_0^t e^{2\int_\tau^t \mathcal{M}^{tr}_N(r)dr} d\tau \le C \ln N/N_1, \tag{6.2}$$

where we have used (5.27) to get the last bound. By symmetry of the components in $\mathbf{v}$, the lemma follows. □

Remark 6.2. The bound in Lemma 6.1 is still valid after redefining $v_{\alpha^N_k}$ along the vertical direction of the array of nodes on $\mathbb{T}^2_{1/N}$.

Lemma 6.3. We have

$$\sup_{t,k} \mathbb{E}\Big|N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} (\hat{x}^{(N_1)}_{\alpha^N_{k'}} - \hat{x}^{(N_1)}_{\alpha^N_k}) - N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} (\bar{x}_{\alpha^N_{k'}} - \bar{x}_{\alpha^N_k})\Big|^2 \le CN^2 \ln N/N_1.$$

Proof. The lemma follows from Lemma 6.1 and Remark 6.2. □

Denote $\xi_{\alpha^N_k}(t) = N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} (\bar{x}_{\alpha^N_{k'}} - \bar{x}_{\alpha^N_k})(t)$. We will further check the asymptotic behavior of $\xi_{\alpha^N_k}(t)$. For convenience of analysis, we split the summation in $\xi_{\alpha^N_k}$ into two parts corresponding to neighbors along two directions of the array of nodes on $\mathbb{T}^2_{1/N}$. The estimate in Lemma 6.5 below will mean that Equation (5.6) may be regarded as a space discretized version of the parabolic equation of $m(t,\alpha)$. Along the horizontal direction of the array for $\mathbb{T}^2_{1/N}$, define

$$\theta_{\alpha^N_k}(t) := N^2(\bar{x}_{\alpha^N_{k_+}} + \bar{x}_{\alpha^N_{k_-}} - 2\bar{x}_{\alpha^N_k})(t). \tag{6.3}$$

For $\alpha \in [0,1)^2 \subset \mathbb{R}^2$, denote $\alpha = (\alpha^{(1)}, \alpha^{(2)})$. Write the second order partial derivatives $\partial^2_{\alpha^{(1)}} m(t,\alpha)$ and $\partial^2_{\alpha^{(2)}} m(t,\alpha)$. We proceed to show that $\theta_{\alpha^N_k}(t)$ approaches $\partial^2_{\alpha^{(1)}} m(t,\alpha^N_k)$ as $N\to\infty$.

Lemma 6.4. Let $\theta_{\alpha^N_k}(t)$, $1 \le k \le N^2$, be defined by (6.3). Then $(\theta_{\alpha^N_1}, \cdots, \theta_{\alpha^N_{N^2}})$ is the unique solution to the ODE system:

$$\begin{aligned}\frac{d}{dt}\theta_{\alpha^N_k}(t) = {}& [A - B^2R^{-1}P(t) + D]\theta_{\alpha^N_k}(t) \\ & - B^2R^{-1}N^2[S(t,\alpha^N_{k_+}) + S(t,\alpha^N_{k_-}) - 2S(t,\alpha^N_k)] \\ & + \kappa N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} (\theta_{\alpha^N_{k'}} - \theta_{\alpha^N_k})(t), \quad 1 \le k \le N^2,\end{aligned} \tag{6.4}$$

with initial condition $\theta_{\alpha^N_k}(0) = N^2[m_0(\alpha^N_{k_+}) + m_0(\alpha^N_{k_-}) - 2m_0(\alpha^N_k)]$.

Proof. We take $\alpha^N_k, \alpha^N_{k_+}$ and $\alpha^N_{k_-}$ in (5.6) to obtain three ODEs, which are multiplied by N^2 and combined together to derive (6.4). The solution to (6.4) is unique since it is a linear ODE system with bounded coefficients when S is given. □

Lemma 6.5. We have

$$\lim_{N\to\infty} \sup_{t\le T, k\le N^2} \big|\theta_{\alpha^N_k}(t) - \partial^2_{\alpha^{(1)}} m(t,\alpha)|_{\alpha=\alpha^N_k}\big| = 0.$$

Proof. To analyze (6.4), we introduce the parabolic equation

$$\begin{aligned}\partial_t\partial^2_{\alpha^{(1)}} m(t,\alpha) = {}& [A - B^2R^{-1}P(t) + D]\partial^2_{\alpha^{(1)}} m(t,\alpha) \\ & - B^2R^{-1}\partial^2_{\alpha^{(1)}} S(t,\alpha)dt + \kappa\Delta(\partial^2_{\alpha^{(1)}} m(t,\alpha)),\end{aligned} \tag{6.5}$$

where the initial condition is $\partial^2_{\alpha^{(1)}} m(0,\alpha)$ (periodic in α). We view $\partial^2_{\alpha^{(1)}} m(t,\alpha)$ as the solution of a second order parabolic equation obtained by differentiating (3.7) twice. There is a unique solution to the resulting linear parabolic equation subject to the initial condition $\partial^2_{\alpha^{(1)}} m(0,\alpha)$. Denoting $\chi(t,\alpha) = \chi_\alpha(t) = \partial^2_{\alpha^{(1)}} m(t,\alpha)$ and taking $\alpha = \alpha^N_k$ in (6.5), we obtain an ODE system with perturbation terms:

$$\begin{aligned}\frac{d}{dt}\chi_{\alpha^N_k}(t) = {}& [A - B^2R^{-1}P(t) + D]\chi_{\alpha^N_k} \\ & - B^2R^{-1}N^2[S(t,\alpha^N_{k_+}) + S(t,\alpha^N_{k_-}) - 2S(t,\alpha^N_k)] \\ & + \kappa N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} (\chi_{\alpha^N_{k'}} - \chi_{\alpha^N_k}) + \epsilon^o_{\alpha^N_k}, \qquad 1 \le k \le N^2,\end{aligned} \tag{6.6}$$

where

$$\begin{aligned}\epsilon^o_{\alpha_k^N}(t) = &- B^2R^{-1}\partial^2_{\alpha^{(1)}}S(t,\alpha)|_{\alpha=\alpha_k^N}\\ &+ B^2R^{-1}N^2[S(t,\alpha^N_{k_+}) + S(t,\alpha^N_{k_-}) - 2S(t,\alpha_k^N)]\\ &+ \kappa\Delta(\partial^2_{\alpha^{(1)}}m(t,\alpha))|_{\alpha=\alpha_k^N} - \kappa N^2\sum_{k':\alpha^N_{k'}\sim\alpha_k^N}(\chi_{\alpha^N_{k'}} - \chi_{\alpha_k^N})(t).\end{aligned}$$

We proceed to compare the solutions of (6.4) and (6.6). It is clear that $\epsilon^o_{\alpha_k^N}(t)$ is continuously differentiable on $[0,T]$ and $\sup_{t,k}|\epsilon^o_{\alpha_k^N}(t)| = o(1)$ as $N\to\infty$, which is due to the approximation of $\partial^2_{\alpha^{(1)}}S(t,\alpha)$ and $\Delta(\partial^2_{\alpha^{(1)}}\chi(t,\alpha))$ by the second order difference quotients under (H3). For (6.4) and (6.6), denote

$$\epsilon_{\alpha_k^N}(t) = \theta_{\alpha_k^N}(t) - \chi_{\alpha_k^N}(t), \qquad \epsilon_N = \sup_{k,t}|\epsilon_{\alpha_k^N}(t)|.$$

Then we have

$$\begin{aligned}\frac{d}{dt}\epsilon_{\alpha_k^N} = &[A - B^2R^{-1}P(t) + D]\epsilon_{\alpha_k^N}\\ &+ \kappa N^2\sum_{k':\alpha^N_{k'}\sim\alpha_k^N}(\epsilon_{\alpha^N_{k'}} - \epsilon_{\alpha_k^N}) - \epsilon^0_{\alpha_k^N}, \quad 1\le k\le N^2,\end{aligned}$$

where $\epsilon_{\alpha_k^N}(0) = N^2[m_0(\alpha^N_{k_+}) + m_0(\alpha^N_{k_-}) - 2m_0(\alpha_k^N)] - \partial^2_{\alpha^{(1)}}\widetilde{m}_0(\alpha_k^N)$.

Thus following the proof of Lemma 5.3, we may construct two auxiliary ODEs by setting $c_1 = \pm\sup_{k,t}|\epsilon^o_{\alpha_k^N}(t)|$ and initial condition $\pm\sup_k|\epsilon_{\alpha_k^N}(0)|$ in (5.7) to bound $\epsilon_{\alpha_k^N}(t)$ both from above and from below on $[0,T]$. So we have $\epsilon_N = o(1)$ in view of the estimate $\sup_{k,t}(|\epsilon^o_{\alpha_k^N}(t)| + |\epsilon_{\alpha_k^N}(0)|) = o(1)$ under (H1) and (H3). □

Remark 6.6. If we define $\theta'_{\alpha_k^N}(t)$ using two neighbors along the vertical direction instead of (6.3), then we similarly obtain

$$\lim_{N\to\infty}\sup_{k\le N^2, t\le T}|\theta'_{\alpha_k^N}(t) - \partial^2_{\alpha^{(2)}}m(t,\alpha)|_{\alpha=\alpha_k^N}| = 0.$$

Remark 6.7. For the ring network $\mathbb{T}_{1/N}$, Lemma 6.3 holds with an upper bound CN^3/N_1, and Lemma 6.5 still holds after replacing $\partial^2_{\alpha^{(1)}}m(t,\alpha)$ by $\partial^2_\alpha m(t,\alpha)$, and $k\le N^2$ by $k\le N$.

6.1 The proof of Theorem 5.2

Proof. We use Lemmas 6.3, 6.5, Remarks 6.6, 6.7, and (H4) to obtain

$$\sup_{N,N_1}\sup_{t,k}\mathbb{E}\Big|N^2\sum_{\alpha^N_{k'}\sim\alpha_k^N}(\hat{x}^{(N_1)}_{\alpha^N_{k'}} - \hat{x}^{(N_1)}_{\alpha_k^N})(t)\Big|^2 \le C \tag{6.7}$$

for some constant C, which combined with Lemma 5.8 yields $\sup_{t,k}\mathbb{E}|\hat{x}^i_{\alpha_k^N}(t)|^2 = O(1)$. Then we obtain the cost bound $\hat{C}_0$ in Theorem 5.2. □

7 The ϵ-Nash equilibrium

This section considers the network cases of both $\mathbb{T}_{1/N}$ and $\mathbb{T}^2_{1/N}$. To establish the ϵ-Nash equilibrium, consider alternative strategies for a single agent. Let the space $\mathcal{U}^c_{\rm ind}$ consist of individual feedback strategies $\varphi(t,\mathbf{x})$, $0\le t\le T$, with $\mathbf{x}$ being the state vector of the overall population (allowing centralized state information), such that φ is continuous in $(t,\mathbf{x})$ and Lipschitz continuous in $\mathbf{x}$. In particular, each decentralized strategy $\hat{u}^i_{\alpha_k^N}$ belongs to $\mathcal{U}^c_{\rm ind}$. If $u^l_{\alpha_j^N}\in\mathcal{U}^c_{\rm ind}$ for all j,l, the set of individual strategies

$(u^i_{\alpha^N_k}, u^{-i}_{\alpha^N_k}, \mathbf{u}_{-\alpha^N_k})$ ensures a well defined solution to (3.1), where $u^{-i}_{\alpha^N_k}$ and $\mathbf{u}_{-\alpha^N_k}$ follow the notation in Section 3. Define the strategy set $\mathcal{U}^{i,\hat{C}_0}_{\alpha^N_k}$ consisting of all $u^i_{\alpha^N_k} \in \mathcal{U}^c_{\text{ind}}$ such that $J^i_{\alpha^N_k}(u^i_{\alpha^N_k}, \hat{u}^{-i}_{\alpha^N_k}, \hat{\mathbf{u}}_{-\alpha^N_k})$ satisfies condition (5.3).

For the performance estimate, we need to approximate the two terms

$$x^{(N_1)}_{\alpha^N_k}(t), \quad N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} (x^{(N_1)}_{\alpha^N_{k'}} - x^{(N_1)}_{\alpha^N_k})(t) \quad \text{in System (B)},$$

respectively, by

$$\hat{x}^{(N_1)}_{\alpha^N_k}(t), \quad N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} (\hat{x}^{(N_1)}_{\alpha^N_{k'}} - \hat{x}^{(N_1)}_{\alpha^N_k})(t) \quad \text{in System (A)}.$$

Denote

$$\zeta_{\alpha^N_k}(t) = N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} (x^{(N_1)}_{\alpha^N_{k'}} - x^{(N_1)}_{\alpha^N_k})(t) - N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} (\hat{x}^{(N_1)}_{\alpha^N_{k'}} - \hat{x}^{(N_1)}_{\alpha^N_k})(t). \tag{7.1}$$

In the following, without loss of generality, we consider an alternative strategy for agent $\mathcal{A}^1_{\alpha^N_1}$ so that $(u^1_{\alpha^N_1}, \hat{u}^{-1}_{\alpha^N_1}, \hat{\mathbf{u}}_{-\alpha^N_1})$ is applied by the population.

Lemma 7.1. For System (B) with $(u^1_{\alpha^N_1}, \hat{u}^{-1}_{\alpha^N_1}, \hat{\mathbf{u}}_{-\alpha^N_1})$ and $u^1_{\alpha^N_1} \in \mathcal{U}^{1,\hat{C}_0}_{\alpha^N_1}$, we have

$$\sup_{u^1_{\alpha^N_1}\in\mathcal{U}^{1,\hat{C}_0}_{\alpha^N_1}} \sup_{t\le T} \mathbb{E}|x^{(N_1)}_{\alpha^N_1}(t) - \hat{x}^{(N_1)}_{\alpha^N_1}(t)|^2 = O(1/N_1^2), \tag{7.2}$$

$$\sup_{u^1_{\alpha^N_1}\in\mathcal{U}^{1,\hat{C}_0}_{\alpha^N_1}} \sup_{t\le T} \mathbb{E}|\zeta_{\alpha^N_1}(t)|^2 = O(N^4/N_1^2). \tag{7.3}$$

Proof. See Appendix A. □

By Lemmas 5.4, 5.7, 6.3, 6.5, and 7.1, Remarks 6.6 and 6.7, we further obtain the following corollary.

Corollary 7.2. For the cases of both $\mathbb{T}_{1/N}$ and $\mathbb{T}^2_{1/N}$, as $N \to \infty$, we have

$$\sup_{u^1_{\alpha^N_1}\in\mathcal{U}^{1,\hat{C}_0}_{\alpha^N_1}} \sup_{t} \mathbb{E}|x^{(N_1)}_{\alpha^N_1}(t) - m(t,\alpha^N_1)|^2 = o(1), \tag{7.4}$$

$$\sup_{u^1_{\alpha^N_1}\in\mathcal{U}^{1,\hat{C}_0}_{\alpha^N_1}} \sup_{t} \mathbb{E}\Big|N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_1} (x^{(N_1)}_{\alpha^N_{k'}} - x^{(N_1)}_{\alpha^N_1}) - \Delta m(t,\alpha^N_1)\Big|^2 = o(1). \tag{7.5}$$

Theorem 7.3. Under assumptions (H1), (H2), (H3), and (H4), the set of decentralized strategies $\{\hat{u}^i_{\alpha^N_k}, 1 \le i \le N_1, 1 \le k \le N^*\}$ in (5.1) is a ϵ_N-Nash equilibriums for the finite population of $N_1 N^*$ agents, i.e.,

$$J^i_{\alpha^N_k}(\hat{u}^i_{\alpha^N_k}, \hat{u}^{-i}_{\alpha^N_k}, \hat{\mathbf{u}}_{-\alpha^N_k}) \le \inf_{u^i_{\alpha^N_k}\in\mathcal{U}^c_{\text{ind}}} J^i_{\alpha^N_k}(u^i_{\alpha^N_k}, \hat{u}^{-i}_{\alpha^N_k}, \hat{\mathbf{u}}_{-\alpha^N_k}) + \epsilon_N, \quad \forall i,k,$$

where $\epsilon_N = o(1)$ as $N \to \infty$.

Proof. Without loss of generality, consider agent $\mathcal{A}^1_{\alpha^N_1}$ taking an alternative strategy $u^1_{\alpha^N_1}$. By Theorem 5.2, it suffices to consider $u^1_{\alpha^N_1} \in \mathcal{U}^{1,\hat{C}_0}_{\alpha^N_1}$. We rewrite the equation of $\mathcal{A}^1_{\alpha^N_1}$ in System (B) with $n = 1$:

$$dx^1_{\alpha^N_1} = [Ax^1_{\alpha^N_1} + Bu^1_{\alpha^N_1} + Dx^{(N_1)}_{\alpha^N_1}]dt \tag{7.6}$$

$$+\kappa N^2 \sum_{k:\alpha_k^N\sim\alpha_1^N}(x_{\alpha_k^N}^{(N_1)}-x_{\alpha_1^N}^{(N_1)})dt+\sigma dw_{\alpha_1^N}^1.$$

To estimate $J_{\alpha_k^N}^1$, we introduce an auxiliary optimal control problem with dynamics and cost:

$$dz(t)=[Az(t)+Bu^\dagger(t)+Dm(t,\alpha_1^N)]dt+\kappa\Delta m(t,\alpha_1^N)dt+\sigma dw_{\alpha_1^N}^1(t), \tag{7.7}$$

$$\begin{aligned}J^*(u^\dagger)&=\mathbb{E}\int_0^T\Big[\|z(t)-\Gamma_1 m(t,\alpha_1^N)-\Gamma_2\eta\|_Q^2+\|u^\dagger(t)\|_R^2\Big]dt\\&\quad+\mathbb{E}\|z(T)-\Gamma_{1f}m(T,\alpha_1^N)-\Gamma_{2f}\eta\|_{Q_f}^2.\end{aligned}$$

where $z(0)=x_{\alpha_1^N}^1(0)$. By the solution of the best response control problem in (3.3)–(3.4), the optimal control law in (7.7) is given by

$$\hat{u}^\dagger(t)=-BR^{-1}[P(t)z(t)+S(t,\alpha_1^N)].$$

For the set of decentralized strategies in (5.1), we use Corollary 7.2 and elementary L^2 error estimates to obtain

$$|J_{\alpha_1^N}^1(\hat{u}_{\alpha_1^N}^1,\hat{u}_{\alpha_1^N}^{-1},\hat{\mathbf{u}}_{-\alpha_1^N})-J^*(\hat{u}^\dagger)|=o(1),\quad \text{as } N\to\infty. \tag{7.8}$$

Throughout the proof, each $o(1)$ term may be estimated without depending on which agent has been selected for an alternative control. When $\hat{u}_{\alpha_1^N}^1$ is replaced by a strategy $u_{\alpha_1^N}^1\in\mathcal{U}_{\alpha_1^N}^{1,\hat{C}_0}$ to generate System (B), we need to show that the cost may be reduced by at most $o(1)$. We write

$$\begin{aligned}&J_{\alpha_1^N}^1(u_{\alpha_1^N}^1,\hat{u}_{\alpha_1^N}^{-1},\hat{\mathbf{u}}_{-\alpha_1^N})\\&=\mathbb{E}\int_0^T\Big(\|x_{\alpha_1^N}^1(t)-\Gamma_1x_{\alpha_1^N}^{(N_1)}(t)-\Gamma_2\eta\|_Q^2+\|u_{\alpha_1^N}^1(t)\|_R^2\Big)dt\\&\qquad+\mathbb{E}\|x_{\alpha_1^N}^1(T)-\Gamma_{1f}x_{\alpha_1^N}^{(N_1)}(T)-\Gamma_{2f}\eta\|_{Q_f}^2.\end{aligned}$$

We write

$$\begin{aligned}&\mathbb{E}\int_0^T\|x_{\alpha_1^N}^1(t)-\Gamma_1x_{\alpha_1^N}^{(N_1)}(t)-\Gamma_2\eta\|_Q^2dt\\&=\mathbb{E}\int_0^T\|z(t)-\Gamma_1m(t,\alpha_1^N)-\Gamma_2\eta\|_Q^2dt\\&\quad+\mathbb{E}\int_0^T\|z(t)-x_{\alpha_1^N}^1(t)+\Gamma_1[x_{\alpha_1^N}^{(N_1)}(t)-m(t,\alpha_1^N)]\|_Q^2dt\\&\quad-2\mathbb{E}\int_0^T[z(t)-\Gamma_1m(t,\alpha_1^N)-\Gamma_2\eta]Q\\&\qquad\cdot[z(t)-x_{\alpha_1^N}^1(t)+\Gamma_1(x_{\alpha_1^N}^{(N_1)}(t)-m(t,\alpha_1^N))]dt.\end{aligned} \tag{7.9}$$

When System (B) has generated $u_{\alpha_1^N}^1(t,\omega)$ as a random process, we set $u^\dagger(t,\omega)=u_{\alpha_1^N}^1(t,\omega)$ in (7.7). By comparing $z(t)$ in (7.7) and $x_{\alpha_1^N}^1(t)$ in (7.6) under the same control process $u_{\alpha_1^N}^1(t,\omega)$ on $[0,T]$, we use Corollary 7.2 to obtain

$$\sup_{u_{\alpha_1^N}^1\in\mathcal{U}_{\alpha_1^N}^{1,\hat{C}_0}}\ \sup_{0\le t\le T}\mathbb{E}|z(t)-x_{\alpha_1^N}^1(t)|^2=o(1), \tag{7.10}$$

as $N \to \infty$. By (7.9), (7.10) and Corollary 7.2, we have

$$\sup_{u^1_{\alpha_1^N} \in \mathcal{U}^{1,\hat{C}_0}_{\alpha_1^N}} \Big| \mathbb{E} \int_0^T \|x^1_{\alpha_1^N}(t) - \Gamma_1 x^{(N_1)}_{\alpha_1^N}(t) - \Gamma_2 \eta\|^2_Q dt - \mathbb{E} \int_0^T \|z(t) - \Gamma_1 m(t, \alpha_1^N) - \Gamma_2 \eta\|^2_Q dt \Big| = o(1).$$

We similarly have

$$\sup_{u^1_{\alpha_1^N} \in \mathcal{U}^{1,\hat{C}_0}_{\alpha_1^N}} \Big| \mathbb{E}\|x^1_{\alpha_1^N}(T) - \Gamma_{1f} x^{(N_1)}_{\alpha_1^N}(T) - \Gamma_{2f}\eta\|^2_{Q_f} - \mathbb{E}\|z(T) - \Gamma_{1f} m(T, \alpha_1^N) - \Gamma_{2f}\eta\|^2_{Q_f} \Big| = o(1).$$

Therefore, we have

$$\begin{aligned} J^1_{\alpha_1^N}(u^1_{\alpha_1^N}, \hat{u}^{-1}_{\alpha_1^N}, \hat{\mathbf{u}}_{-\alpha_1^N}) &\geq \mathbb{E} \int_0^T \Big[\|z(t) - \Gamma_1 m(t, \alpha_1^N) - \Gamma_2 \eta\|^2_Q + \|u^1_{\alpha_1^N}(t)\|^2_R \Big] dt \\ &\quad + \mathbb{E}\|z(T) - \Gamma_{1f} m(T, \alpha_1^N) - \Gamma_{2f}\eta\|^2_{Q_f} - o(1) \\ &\geq J^*(\hat{u}^\dagger) - o(1), \end{aligned} \tag{7.11}$$

where $z(t)$ is generated by (7.7) using the process $u^1_{\alpha_1^N}(t, \omega)$. By (7.8) and (7.11), it follows that

$$J^1_{\alpha_1^N}(u^1_{\alpha_1^N}, \hat{u}^{-1}_{\alpha_1^N}, \hat{\mathbf{u}}_{-\alpha_1^N}) \geq J^1_{\alpha_1^N}(\hat{u}^1_{\alpha_1^N}, \hat{u}^{-1}_{\alpha_1^N}, \hat{\mathbf{u}}_{-\alpha_1^N}) - o(1).$$

This completes the proof. □

8 Conclusion

Future work will address general existence and uniqueness analyses for nonlinear Laplexion mean field game equation systems.

Appendix A

Proof of Lemma 4.4. We first show that Λ_2 has its image in $C^{\delta/2,\delta}_{prd}(\mathbf{Q})$. Take an arbitrary function $m \in C^{1+\delta/2,2+\delta}_{prd}(\mathbf{Q})$ in (3.6), which is solved as an ODE to give S as a function with period 1 in each space variable:

$$\begin{aligned} S(t,\alpha) =& \Lambda_2(m)(t) \\ =& -\int_t^T \Phi(t,\tau)[(Q\Gamma_1 - P(\tau)D)m(\tau,\alpha) - \kappa P(\tau)\Delta m(\tau,\alpha) + Q\Gamma_2\eta] d\tau \\ & - \Phi(t,T)Q_f[\Gamma_{1f} m(T,\alpha) + \Gamma_{2f}\eta]. \end{aligned} \tag{A.1}$$

To estimate $|S|_{\delta/2,\delta;\mathbf{Q}}$ using $[S]_{\delta/2,\delta;\mathbf{Q}^\circ}$ according to Lemma 4.3, we observe

$$\begin{aligned} |S(t,\alpha) - S(t,\alpha')| \leq& \int_t^T |\Phi(t,\tau)| \cdot |Q\Gamma_1 - P(\tau)D| \cdot |m(\tau,\alpha) - m(\tau,\alpha')| d\tau \\ & + \int_t^T |\Phi(t,\tau)| \cdot \kappa |P(\tau)| \cdot |\Delta m(\tau,\alpha) - \Delta m(\tau,\alpha')| d\tau \\ & + |\Phi(t,T)Q_f\Gamma_{1f}| \cdot |m(T,\alpha) - m(T,\alpha')|. \end{aligned} \tag{A.2}$$

By Lemma 4.3 and periodicity of S, it suffices to consider $\alpha, \alpha' \in \mathbb{R}^\iota$ with $d_{max}(\alpha, \alpha') := \max_{i\le\iota} |\alpha_i - \alpha_i'| \le 1/2$. Hence by (A.2), we have

$$\begin{aligned}
&\sup_{\alpha\neq\alpha', t\in[0,T]} d_{max}(\alpha,\alpha')^{-\delta}|S(t,\alpha)-S(t,\alpha')| \\
&\le T \sup_{t,\tau\in[0,T]} |\Phi(t,\tau)| \cdot (|Q\Gamma_1 - PD|_0 + \kappa|P|_0) \\
&\quad \times \sup_{\tau,\alpha,\alpha'} d_{max}(\alpha,\alpha')^{-\delta}\Big(|m(\tau,\alpha)-m(\tau,\alpha')| + |\Delta m(\tau,\alpha) - \Delta m(\tau,\alpha')|\Big) \\
&\quad + \sup_t |\Phi(t,T)Q_f\Gamma_{1f}| \cdot \sup_{\alpha,\alpha'} d_{max}(\alpha,\alpha')^{-\delta}|m(T,\alpha)-m(T,\alpha')| \\
&\le T\Big\{\sup_{t,\tau}|\Phi(t,\tau)| \cdot (|Q\Gamma_1 - PD|_0 + \kappa|P|_0)\Big\} \\
&\quad \times \Big(\sum_i |\partial_{\alpha_i} m|_{0;\mathbf{Q}} + \sup_{\tau,\alpha,\alpha'} d_{max}(\alpha,\alpha')^{-\delta}|\Delta m(\tau,\alpha)-\Delta m(\tau,\alpha')|\Big) \\
&\quad + T\sup_t |\Phi(t,T)Q_f\Gamma_{1f}| \sup_{\tau,\alpha,\alpha'} d_{max}(\alpha,\alpha')^{-\delta}|\partial_\tau m(\tau,\alpha) - \partial_\tau m(\tau,\alpha')| \\
&\quad + \sup_t |\Phi(t,T)Q_f\Gamma_{1f}| \cdot \sum_i |\partial_{\alpha_i} m(0,\cdot)|_{0;\mathbb{R}^\iota}.
\end{aligned} \tag{A.3}$$

Next for $0 \le t < t' \le T$, we have

$$\begin{aligned}
S(t,\alpha) - S(t',\alpha) = &-\int_t^{t'} \Phi(t,\tau)[(Q\Gamma_1 - P(\tau)D)m(\tau,\alpha) - \kappa P(\tau)\Delta m(\tau,\alpha)]d\tau \\
&+ \int_{t'}^T [\Phi(t',\tau) - \Phi(t,\tau)] \cdot [(Q\Gamma_1 - P(\tau)D)m(\tau,\alpha) - \kappa P(\tau)\Delta m(\tau,\alpha)]d\tau \\
&+ [\Phi(t',T) - \Phi(t,T)]Q_f[\Gamma_{1f} m(T,\alpha) + \Gamma_{2f}\eta].
\end{aligned}$$

We use the differentiability of Φ to estimate

$$\begin{aligned}
&|t-t'|^{-\delta/2} \cdot |S(t,\alpha) - S(t',\alpha)| \\
&\le \Big(T^{1-\delta/2}\sup_{t,\tau}|\Phi(t,\tau)| + T^{2-\delta/2}\sup_{t,\tau}|\partial_t\Phi(t,\tau)|\Big) \\
&\qquad \times (|Q\Gamma_1 - PD|_0 + \kappa|P|_0)(|m|_{0;\mathbf{Q}} + |\Delta m|_{0;\mathbf{Q}}) \\
&\quad + T^{1-\delta/2}\sup_t |\partial_t \Phi(t,T)| \cdot (|Q_f\Gamma_{1f}| \cdot |m|_{0;\mathbf{Q}} + |Q_f\Gamma_{2f}\eta|).
\end{aligned} \tag{A.4}$$

Hence we conclude $S = \Lambda_2(m) \in C^{\delta/2,\delta}_{prd}(\mathbf{Q})$.

Now it suffices to consider $z = (t,\alpha)$, $z = (t',\alpha')$ such that $d_{max}(\alpha,\alpha') \le 1/2$. We have

$$\begin{aligned}
\frac{|S(z) - S(z')|}{d_p(z,z')^\delta} &\le \frac{|S(z) - S(t,\alpha') + S(t,\alpha') - S(z')|}{d_p(z,z')^\delta} \\
&\le d_{max}(\alpha,\alpha')^{-\delta}|S(z) - S(t,\alpha')| + |t-t'|^{-\delta/2}|S(t,\alpha') - S(z')|,
\end{aligned}$$

which combined with (A.3), (A.4) and an estimate of $|S|_{0;\mathbf{Q}}$ using (A.1) yields (4.4). □

We give a comparison lemma of differential inequalities, which is used in proving Lemma 5.3. For similar comparison results, see [35].

Lemma A.1. Suppose $c(t)$ is a nonnegative continuous function on $[0,T]$, $c_0 \ge 0$ is a constant, and the function $z(t)$ is absolute continuous on $[0,T]$.

(i) If $\dot z(t) \le c(t)z(t) + c_0$ holds almost everywhere on the set $\{t : z(t) > 0\}$, and $y(t)$ satisfies the ODE $\dot y(t) = c(t)y(t) + c_0$ with $y(0) \ge \max\{z(0), 0\}$. Then $z(t) \le y(t)$ for all $t \in [0,T]$.

(ii) If $\dot z(t) \geq c(t)z(t) - c_0$ holds almost everywhere on the set $\{t : z(t) < 0\}$, and $y(t)$ satisfies the ODE $\dot y(t) = c(t)y(t) - c_0$ with $y(0) \leq \min\{z(0), 0\}$. Then $z(t) \geq y(t)$ for all $t \in [0, T]$.

Proof. For case (i), we may assume non-empty $\{t : z(t) > 0\}$ since otherwise the conclusion is obvious. We prove by contradiction and assume there exists $t_1 \in (0, T]$ such that $z(t_1) > y(t_1)$, which implies for some $t_0 \in [0, t_1)$, $z(t) \geq y(t) \geq 0$ holds on $[t_0, t_1]$ with $z(t_0) = y(t_0)$. We have

$$y(t_1) - z(t_1) = \int_{t_0}^{t_1} [\dot y(s) - \dot z(s)]ds \geq \int_{t_0}^{t_1} c(s)[y(s) - z(s)]ds \geq 0,$$

leading to a contradiction. Therefore we get the desired comparison inequality. The proof of part (ii) is similar and omitted here. □

Lemma A.2. Suppose $F(t)$ is a continuous $\mathbb{R}^{k\times k}$-valued function defined on $[0, T]$, and $F(t)F(s) = F(s)F(t)$ for all $s, t \in [0, T]$. Then $\Phi(t, s) = e^{\int_s^t F(r)dr}$ is the fundamental solution matrix of the linear ODE $\dot x = F(t)x$, i.e., $\frac{\partial}{\partial t}\Phi(t, s) = F(t)\Phi(t, s)$ with $\Phi(s, s) = I_k$.

Proof. We have $\frac{1}{\epsilon}[\Phi(t + \epsilon, s) - \Phi(t, s)] = \frac{1}{\epsilon}(e^{\int_t^{t+\epsilon} F(r)dr} - I)e^{\int_s^t F(r)dr} \to F(t)\Phi(t, s)$ as $\epsilon \to 0$, where the equality follows from the commutativity condition $F(t)F(s) = F(s)F(t)$ for all s, t. □

Proof of Lemma 7.1. For System (B), we write

$$\begin{aligned} dx^{(N_1)}_{\alpha_1^N} =&[A - B^2R^{-1}P(t) + D]x^{(N_1)}_{\alpha_1^N}dt - B^2R^{-1}S(t, \alpha_1^N)dt \\ &+ \kappa N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_1} (x^{(N_1)}_{\alpha^N_{k'}} - x^{(N_1)}_{\alpha_1^N})dt + (B/N_1)\tilde u^1_{\alpha_1^N}dt + \sigma dw^{(N_1)}_{\alpha_1^N}, \end{aligned} \tag{A.5}$$

$$\begin{aligned} dx^{(N_1)}_{\alpha_k^N} =&[A - B^2R^{-1}P(t) + D]x^{(N_1)}_{\alpha_k^N}dt - B^2R^{-1}S(t, \alpha_k^N)dt \\ &+ \kappa N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} (x^{(N_1)}_{\alpha^N_{k'}} - x^{(N_1)}_{\alpha_k^N})dt + \sigma dw^{(N_1)}_{\alpha_k^N}, \quad 2 \leq k \leq N^*, \end{aligned} \tag{A.6}$$

where $x^{(N_1)}_{\alpha_j^N}(0) = m_0(\alpha_j^N)$ for all $1 \leq j \leq N^*$, and

$$\tilde u^1_{\alpha_1^N}(t) = u^1_{\alpha_1^N}(t) + BR^{-1}P(t)x^1_{\alpha_1^N}(t), \tag{A.7}$$

which is interpreted as a random process generated by the closed-loop System (B). Denote $\tilde z_{\alpha_k^N}(t) = x^{(N_1)}_{\alpha_k^N}(t) - \hat x^{(N_1)}_{\alpha_k^N}(t)$, $\tilde{\mathbf{z}} = [\tilde x_{\alpha_1^N}, \cdots, \tilde x_{\alpha^N_{N^*}}]^T$ and $\tilde{\mathbf{u}} = [\tilde u^1_{\alpha_1^N}, 0, \cdots, 0]^T$.

(i) The case of $\mathbb{T}^2_{1/N}$. For $u^1_{\alpha_1^N} \in \mathcal{U}^{1,\hat C_0}_{\alpha_1^N}$, we will give some estimates in terms of $\mathbb{E}\int_0^T |\tilde u^1_{\alpha_1^N}(t)|^2dt$, which will be shown later to be bounded. By (A.5) and (A.6), and in analogue to (5.16), we have

$$d\tilde{\mathbf{z}} = \mathcal{M}^{tr}_N\tilde{\mathbf{z}}(t)dt + (B/N_1)\tilde{\mathbf{u}}(t)dt,$$

which gives

$$\tilde z_{\alpha_k^N}(t) = [0, \cdots, 1, \cdots, 0]\int_0^t \Phi^{tr}(t, \tau)(B/N_1)\tilde{\mathbf{u}}(\tau)d\tau, \tag{A.8}$$

where 1 is the k-th entry of the row. Let $(M)_{ij}$ denote the (i, j)-th entry of a matrix M. By (5.22) and an orthogonal transformation we can show $\sup_N \sup_{t,s} |(\Phi^{tr}(t, s))_{ij}| \leq C$, so that (A.8) gives

$$\sup_{t,k} \mathbb{E}|\tilde z_{\alpha_k^N}(t)|^2 \leq (C/N_1^2)\mathbb{E}\int_0^T |\tilde u_{\alpha_1^N}(\tau)|^2d\tau, \tag{A.9}$$

which, in view of Lemma 5.8, implies that

$$\sup_{t,k} \mathbb{E}|x^{(N_1)}_{\alpha^N_k}(t)|^2 \le (C/N_1^2)\mathbb{E}\int_0^T |\tilde{u}_{\alpha^N_1}(\tau)|^2 d\tau + C_1. \tag{A.10}$$

We further estimate $\zeta_{\alpha^N_1}$ in (7.1). We have

$$\begin{aligned}
d\tilde{z}_{\alpha^N_1} &= [A - B^2R^{-1}P(t) + D]\tilde{z}_{\alpha^N_1}dt + \kappa N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_1} (\tilde{z}_{\alpha^N_{k'}} - \tilde{z}_{\alpha^N_1})dt \\
&\quad + (B/N_1)\tilde{u}^1_{\alpha^N_1}dt, \\
d\tilde{z}_{\alpha^N_k} &= [A - B^2R^{-1}P(t) + D]\tilde{z}_{\alpha^N_k}dt + \kappa N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} (\tilde{z}_{\alpha^N_{k'}} - \tilde{z}_{\alpha^N_k})dt, \\
&\quad 2 \le k \le N^2,
\end{aligned}$$

Let $\alpha^N_{i_+}$ and $\alpha^N_{i_-}$ be the right and left neighbors of node α^N_i, respectively. Denote

$$\chi_{\alpha^N_i} = \tilde{z}_{\alpha^N_{i_+}}(t) + \tilde{z}_{\alpha^N_{i_-}}(t) - 2\tilde{z}_{\alpha^N_i}(t), \quad \chi = [\chi_{\alpha^N_1}, \cdots, \chi_{\alpha^N_{N^2}}]^T.$$

Then by the method of getting (6.1), we derive

$$d\chi = \mathcal{M}^{tr}_N \chi dt + (B/N_1)\tilde{\mathbf{u}}^\star dt,$$

where the entry within $\tilde{\mathbf{u}}^\star$ at the row corresponding to node α^N_j is nonzero if and only if α^N_j is equal to α^N_1 or is as its horizontal neighbor. So $\tilde{\mathbf{u}}^\star$ contains a total of three nonzero entries: $\tilde{u}^1_{\alpha^N_1}$, $\tilde{u}^1_{\alpha^N_1}$, and $-2\tilde{u}^1_{\alpha^N_1}$. We have

$$\chi_{\alpha^N_1}(t) = (B/N_1)\int_0^T [1, 0, \cdots, 0]\Phi^{tr}(t,\tau)\tilde{\mathbf{u}}^\star(\tau)d\tau.$$

Hence, by the method of establishing (A.9), we obtain

$$\mathbb{E}|\chi_{\alpha^N_1}(t)|^2 = (C/N_1^2)\mathbb{E}\int_0^T |\tilde{u}_{\alpha^N_1}(\tau)|^2 d\tau. \tag{A.11}$$

We again get a bound of the form (A.11) when two neighbours along the vertical direction are used, so that we get the bound for $\zeta_{\alpha^N_1}$ in (7.1):

$$\sup_t \mathbb{E}|\zeta_{\alpha^N_1}(t)|^2 \le (CN^4/N_1^2)\mathbb{E}\int_0^T |\tilde{u}_{\alpha^N_1}(\tau)|^2 d\tau, \tag{A.12}$$

which combined with (6.7) yields

$$\sup_{t,k} \mathbb{E}\Big|N^2 \sum_{k':\alpha^N_{k'}\sim\alpha^N_k} (x^{(N_1)}_{\alpha^N_{k'}} - x^{(N_1)}_{\alpha^N_k})(t)\Big|^2 \le C + (CN^4/N_1^2)\mathbb{E}\int_0^T |\tilde{u}_{\alpha^N_1}(\tau)|^2 d\tau. \tag{A.13}$$

Now by (A.10), (A.13), and System (B), we further obtain

$$\sup_t \mathbb{E}|x_{\alpha^N_1}(t)|^2 \le C + (CN^4/N_1^2)\mathbb{E}\int_0^T |\tilde{u}_{\alpha^N_1}(\tau)|^2 d\tau. \tag{A.14}$$

By (A.7), (A.14) and $u^1_{\alpha^N_1} \in \mathcal{U}^{1,\hat{C}_0}_{\alpha^N_1}$, we derive

$$\mathbb{E}\int_0^T |\tilde{u}_{\alpha^N_1}(\tau)|^2 d\tau \le C + (CN^4/N_1^2)\mathbb{E}\int_0^T |\tilde{u}_{\alpha^N_1}(\tau)|^2 d\tau,$$

so that under (H4),

$$\sup_N \sup_{u^1_{\alpha_1^N}\in\mathcal{U}^{1,\hat{C}_0}_{\alpha_1^N}} \mathbb{E}\int_0^T |\tilde{u}_{\alpha_1^N}(\tau)|^2 d\tau \le C,$$

which combined with (A.9) and (A.12) implies (7.2) and (7.3).

(ii) The case of $\mathbb{T}_{1/N}$. The estimate is similar and we omit the detail. □

Appendix B

The following matrix illustrates $\mathcal{T}_N$ in (5.13) with $N = 4$:

$$\mathcal{T}_4 = \left[\begin{array}{cccc|cccc|cccc|cccc}
0&1&0&1&1&0&0&0&0&0&0&0&1&0&0&0\\
1&0&1&0&0&1&0&0&0&0&0&0&0&1&0&0\\
0&1&0&1&0&0&1&0&0&0&0&0&0&0&1&0\\
1&0&1&0&0&0&0&1&0&0&0&0&0&0&0&1\\
\hline
1&0&0&0&0&1&0&1&1&0&0&0&0&0&0&0\\
0&1&0&0&1&0&1&0&0&1&0&0&0&0&0&0\\
0&0&1&0&0&1&0&1&0&0&1&0&0&0&0&0\\
0&0&0&1&1&0&1&0&0&0&0&1&0&0&0&0\\
\hline
0&0&0&0&1&0&0&0&0&1&0&1&1&0&0&0\\
0&0&0&0&0&1&0&0&1&0&1&0&0&1&0&0\\
0&0&0&0&0&0&1&0&0&1&0&1&0&0&1&0\\
0&0&0&0&0&0&0&1&1&0&1&0&0&0&0&1\\
\hline
1&0&0&0&0&0&0&0&1&0&0&0&0&1&0&1\\
0&1&0&0&0&0&0&0&0&1&0&0&1&0&1&0\\
0&0&1&0&0&0&0&0&0&0&1&0&0&1&0&1\\
0&0&0&1&0&0&0&0&0&0&0&1&1&0&1&0
\end{array}\right].$$

References

[1] S. Banerjee, A. Budhiraja, and M. Perlmutte. A new approach to large deviations for the Ginzburg–Landau model, Electronic J. Probability, vol. 25, no. 26, 1–51, 2020. doi: 10.1214/20-EJP434.

[2] E. Bayraktar, S. Chakraborty, and R. Wu. Graphon mean field systems, Ann. Appl. Probab., vol. 33, pp. 3587–3619, 2023.

[3] L. E. Blume. The statistical mechanics of strategic interaction, Games and Economic Behavior, vol. 5, pp. 387–424, 1993.

[4] C. Borgs, J. T. Chayes, H. Cohn, and Y. Zhao. An L^p theory of sparse graph convergence I: Limits, sparse random graph models, and power law distributions. Transactions of the American Mathematical Society, vol. 372, no.5, pp. 3019–3062, 2019.

[5] P. E. Caines. Embedded vertexon-graphons and embedded GMFG systems, *Proc. 61st IEEE CDC*, Cancun, Mexico, Dec 2022, pp. 5550–5557.

[6] P. E. Caines, D. Ho, M. Huang, J. Jian, and Q. Song. On the graphon mean field game equations: Individual agent affine dynamics and mean field dependent performance functions, *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 28, 2022.

[7] P. E. Caines and M. Huang. Graphon mean field games and the GMFG equations, *Proc. 57th IEEE CDC*, Miami Beach, FL, USA, 2018, pp. 4129–4134.

[8] P. E. Caines and M. Huang. Graphon mean field games and their equations, *SIAM J. Contr. Optim.*, vol. 59, no. 6, pp. 4373–4399, 2021.

[9] P. E. Caines and M. Huang. Mean field games on dense and sparse networks: the graphexon MFG equations, *Proc. American Control Conference*, Toronto, Canada, Jul. 2024, pp. 4230–4235.

[10] P. E. Caines and M. Huang. Sparse network mean field games: ring structures and related topologies, *Proc. the 63th IEEE CDC*, Milan, Italy, 2024, pp. 2584–2590.

[11] P. E. Caines and M. Huang. Mean field games on large sparse network limits: Laplexion dynamics on graphexons, Proc. 10th IFAC Conf on Networked Systems, Hong Kong, China, June 2025, pp. 43–48.

[12] P. E. Caines and M. Huang. Mean field games on sparse network limits: Laplexion dynamics and ϵ-Nash equilibria, Proc. 64th IEEE CDC, Rio de Janeiro, Brazil, Dec 2025.

[13] J. Calder and D. Slepcev. Properly-weighted graph Laplacian for semi-supervised learning, Appl. Math. Optim., vol. 82, pp. 1111–1159, 2020.

[14] R. Carmona, D. B. Cooney, C. V. Graves, and M. Lauriere. Stochastic graphon games: I. the static case, *Math. Oper. Res.*, vol. 47, pp. 750–778, 2022.

[15] P. J. Davis. Circulant Matrices, New York: John Wiley, 1979.

[16] A. de Crescenzo, F. de Feo, and H. Pham. Linear-quadratic optimal control for non-exchangeable mean-field SDEs and applications to systemic risk, arXiv:2503.03318v2, 2025.

[17] Sven Dorkenwal, Arie Matsliah, et al. Neuronal wiring diagram of an adult brain, Nature, 634, p. 124–138, 2024.

[18] M. Fischer, L. Kanzler, and C. Schmeiser. One-dimensional short-range nearest-neighbor interaction and its nonlinear diffusion limit, *SIAM J. Appl. Math.*, vol. 84, no. 1, pp. 1–18, 2024.

[19] A. Friedman. Partial differential equations of parabolic type, Mileona, New York: Dover Publications, 2008.

[20] S. Gao and P. E. Caines. Graphon control of large-scale networks of linear systems, *IEEE Transactions on Automatic Control*, vol. 65, no. 10, pp. 4090–4105, 2020.

[21] M. Z. Guo, G. C. Papanicolaou and S. R. S. Varadhan. Nonlinear diffusion limit for a system with nearest neighbor interactions, Commun. Math. Phys., vol. 118, pp. 31–59, 1988.

[22] R. A. Horn and C. R. Johnson. Topics in matrix analysis, Cambridge University Press, Cambridge, 1991.

[23] M. Huang and M. Zhou. Linear quadratic mean field games: Asymptotic solvability and relation to the fixed point approach, IEEE Transactions on Automatic Control, vol. 65, no. 4, pp. 1397–1412, Apr 2020

[24] Y. M. Ioannides. Topologies of social interactions, Econ. Theory, vol. 28, pp. 559–584, 2006.

[25] D. Kaliuzhnyi-Verbovetskyi and G. S. Medvedev. The semilinear heat equation on sparse random graphs, *SIAM Journal on Mathematical Analysis*, vol. 49, no. 2, pp. 1333–1355, 2017.

[26] C. Kuehn and C. Xu. Vlasov equations on digraph measures, *Journal of Differential Equations*, vol. 339, pp. 261–349, 2022.

[27] D. Lacker and A. Soret. A case study on stochastic games on large graphs in mean field and sparse regimes, *Math. Oper. Res.*, vol. 47, no. 2, pp. 1530–1565, 2021.

[28] D. Lacker, K. Ramanan, and R. Wu. Local weak convergence for sparse networks of interacting processes, *The Annals of Applied Probability*, vol. 33, no. 2, pp. 843–888, 2023.

[29] O. A. Ladyzhenskaya, N. Ural'ceva, and V. Solonnikov. Linear and quasi-linear equations of parabolic type, AMS, 1968.

[30] L. Lovász, *Large networks and graph limits*, AMS, 2012.

[31] G. S. Medvedev. The nonlinear heat equation on w-random graphs, *Arch. Rat. Mech. Anal.*, vol. 212, no. 3, pp. 781–803, 2014.

[32] R. Olfati-Saber, J. A. Fax, and R. M. Murray, Consensus and cooperation in networked multi-agent systems, Proceedings of IEEE, vol. 95, no. 1, pp. 215–233, 2007.

[33] G. A. Pagani and M. Aiello. The power grid as a complex network: a survey, Physica A, vol. 392, pp. 2688–2700, 2013.

[34] F. Parise and A. Ozdaglar. Graphon games, *Proc. 2019 ACM Conference on Economics and Computation*, 2019, pp. 457–458.

[35] G. Teschl. Ordinary differential equations and dynamical systems, AMS, 2012.

[36] H. P. Young. Individual strategy and social structure: an evolutionary theory of institutions. Princeton, NJ: Princeton Univ. Press, 1998.