J Syst Sci Complex (2025) 38(1): 349–368

A Graphon Mean Field Convex Duality Approach to Shaping the Terminal Probability Distribution of a Network of Nonlinear Multi-Agent Systems

PAKNIYAT Ali

DOI: 10.1007/s11424-025-4507-7
Received: 17 October 2024

Abstract In order to steer the population distribution of a large number of agents interacting over a large-scale complex network towards a set of desired probability distributions for each sub-population, an approximate control scheme is proposed and developed by the use of Graphon Mean Field theory and Convex Duality Optimal Control. For a general class of multi-agent nonlinear systems interacting over large networks, the original problem for a finite population over a finite network is reformulated as an optimal control problem for an infinite population over an infinite network by letting the number of nodes in the graph and the number of agents within each cluster approach infinity. Subsequently, the associated control problem for the graphon limit system is reformulated as a linear program over the space of Radon measures and is solved using the duality relationship between the space of measures and that of continuous functions. A numerical example of a network with randomly sampled weightings is presented to illustrate the effectiveness of the graphon control probability assignment methodology.

Keywords Complex networks, convex duality optimal control, covariance control, graphon control, graphons, infinite dimensional systems, large networks, probability assignment.

1 Introduction

One effective strategy to overcome the computational intractability of large-scale systems is to pass to an appropriately formulated infinite limit[1]. This approach has a distinguished history, as can be traced within the conceptual principle underlying the celebrated Boltzmann equation of statistical mechanics and that of the fundamental Navier-Stokes equation of fluid mechanics, as well as the Fokker-Planck-Kolmogorov (FPK) equations for the macroscopic flow of probabilities[1]. These equations have played a crucial role in modeling the behavior of systems with a large number of interacting particles, allowing for the derivation of macroscopic

PAKNIYAT Ali
PO BOX 870276 *Mechanical Engineering; University of Alabama, Tuscaloosa, AL* 35487-0001*, USA.*
Email: apakniyat@ua.edu.
◇ *This paper was recommended for publication by Guest Editor ZHANG Ji-Feng.*
This work is humbly dedicated to Professor Peter E. Caines, in honor of his 80th birthday.

properties from microscopic dynamics with the consideration of the limit as the number of particles approaches infinity.

Mean field games (MFG) theory, introduced by Huang, et al.[2, 3], and independently by Lasry and Lions[4–6], represents another significant application of the infinite limit approach in control theory. By considering the limit as the number of agents approaches infinity, MFG transforms intractable multi-agent control problems into tractable ones by replacing the large number of individual interactions between agents with an aggregated mean field effect on a representative agent; thus significantly simplifying the analysis and the computation of optimal strategies. The resulting MFG system, typically consisting of a forward Fokker-Planck equation and a backward Hamilton-Jacobi-Bellman (HJB) equation, provides a powerful tool for studying large-scale systems in various fields, including economics[7], finance[8], and social sciences[9].

Classical Mean Field Game theory, while powerful, has a significant limitation: It assumes that every agent is equally influenced by all other agents in the population, whereas in many real-world scenarios, agents interact with only a subset of the population, and these interactions often have different strengths. This underscores the need for a more sophisticated framework capable of accommodating intricate interaction structures while maintaining analytical tractability. Traditional graph theory provides a natural way to represent such structured interactions, but it becomes computationally intractable for very large systems. To overcome this challenge, we turn to the concept of graphons—limit objects that emerge as the number of nodes in a sequence of graphs approaches infinity[10]. Graphons provide a continuous representation of large-scale networks, enabling the application of analytical techniques that would be prohibitively complex for finite graphs of comparable size.

Graphon mean field games (GMFG)[1, 11–16] provide a powerful framework for analyzing optimal control and game-theoretic problems in systems where a large population of agents interact through complex network structures. By extending MFG theory to graphon structures, GMFG enables the analysis of interactions among vast numbers of agents with diverse connection patterns and heterogeneous interaction strengths, while maintaining analytical tractability. The GMFG equations are of significant generality, permitting the study of both highly interconnected (dense) and weakly connected (sparse) networks of dynamical agents.

However, the existing GMFG framework lacks the capability to impose constraints on the associated probability distributions, limiting its applicability to scenarios involving desired distributional outcomes from large networks of multi-agent systems. For instance, while the GMFG framework can model power distribution networks, the management of energy consumption patterns often requires achieving specific distribution targets, such as controlling the probability distribution of consumption at individual nodes during peak demand periods. This paper addresses this limitation of the GMFG framework by extending it to accommodate terminal constraints on the probability measures associated with graphon mean field systems, enabling the steering of population distributions towards desired predefined outcomes. This extension significantly broadens the GMFG framework's applicability to real-world scenarios where precise control over probability distributions is essential, opening up new avenues for modeling and optimizing complex networked systems.

The study of terminal probability distributions in networks of nonlinear multi-agent systems is particularly relevant in the context of emerging technologies and societal challenges. For instance, in smart grid management, controlling the distribution of energy consumption across a network can help balance load and prevent blackouts. In traffic control systems, shaping the probability distribution of vehicles across a road network can optimize traffic flow and reduce congestion. Additionally, in epidemiology, understanding and controlling the distribution of infected individuals across a population network is crucial for effective disease management. These practical applications underscore the importance of developing fundamental methods for probability distribution control in large-scale networked systems.

The problem of assigning probability distributions to the state of single-agent stochastic systems has been the subject of multiple control theoretic investigations. Historically, the majority of the earlier research has focused on linear systems with quadratic costs, where the resulting probability distributions are Gaussian. Under these assumptions, the dynamics of the mean state and covariance state become decoupled, yielding to the decomposition of the input into a part for steering the mean process, and the other part for steering the covariance. For infinite time horizons, this problem has been extensively studied, with key contributions addressing assignable covariances and steady-state behaviors[17–20]. For finite time horizons, both continuous and discrete time methodologies have been developed to assign distributions effectively, as detailed in [21–30]. The integration of input constraints into these models is explored in [21] and, further, convex relaxations for linear systems subject to probabilistic constraints are investigated in [29, 30], which guarantee upper bounds on the probability of constraint violations. Additionally, model predictive control (MPC) approaches, which adaptively update control strategies based on evolving information, are discussed in [31–34].

The extension of distribution assignment techniques to single-agent nonlinear stochastic systems has been a focus of recent research. For nonlinear systems, the convenient properties of linear systems no longer apply: The problem is not separable into mean and covariance steering components, and the associated probability distributions are generally non-Gaussian, even when the noise processes are of Gaussian nature (i.e., Brownian noise and Wiener processes). Initial work has been conducted on feedback-linearizable systems[35], where the problem is transformed into a linear control problem within a new transformation of the state space. More advanced methods include iterative linearization[36] and differential dynamic programming approximations[37]. More recently, the problem has been reformulated in the context of the nonlinear Schrödinger bridge problem, as explored in [38, 39]. However, these results are limited to systems where the stochastic dynamics conform to gradient flow forms.

For more general single-agent optimal control problems subject to distribution constraints, the necessary optimality conditions are established in [40] for systems with: (i) Dynamics governed by general nonlinear Itô differential equations, (ii) costs in general nonlinear forms, and (iii) desired probability distributions of any form, not necessarily Gaussian. These conditions are established through an extension of Convex Duality Optimal Control (CDOC) in the presence of constraints on the associated probability measures, which identifies the value function as the supremizing function of an optimization problem over a class of Hamilton-Jacobi (HJ) problems where the optimization objective, in addition to the value function evaluation at the initial state, includes an extra term which is the integral of the product of the value function at the terminal time and the desired probability distribution.

Convex duality methods for optimal control was originally proposed by Vinter and Lewis[41, 42] and later extended by Fleming and Vermes for piecewise deterministic[43] and stochastic[44] processes. The fundamental idea of CDOC is the introduction of a weak formulation that embeds the original (strong) problem into a convex linear program over the space of Radon measures. Upon establishing the equivalence of the two problems, new necessary and sufficient optimality condition are obtained by invoking the Fenchel-Rockafellar duality theorem. This methodology, which is particularly useful in the characterization of optimal policies in certain classes of controls by investigating the extreme points of the set of HJ problems[45–48], has been applied to both continuous[49–51] and hybrid systems[52–54].

This paper builds on recent advancements in single-agent distributional control and covariance steering theory by extending the results to multi-agent stochastic systems with complex interaction structures of graphon mean field systems. Specifically, this work extends the distributionally constrained convex duality optimal control (DC-CDOC) established in [40, 55] for assigning probability measures to single agent nonlinear stochastic systems, to the case of a large network of nonlinear multi-agent systems within the framework of graphon mean field games theory developed in [1, 11]. By integrating these approaches, this paper formulates a decentralized control strategy for the network that ensures probability distributions across both overall populations and subpopulations are steered towards predetermined desired distributions. The established control strategies, which are optimal for the associated infinite population representation of the system over an infinite network, are ε-optimal for the original system with finite agents over a finite network where $\varepsilon \to 0$ as the number of agents and the number of nodes approach infinity[1].

The DC-CDOC method proposed in this paper extends the single-agent distributional control techniques to the multi-agent setting within the graphon mean field framework. This extension allows for the control of probability distributions in large-scale networked systems, which was not previously possible with existing graphon mean field game approaches. By reformulating the problem as a linear program over the space of Radon measures and leveraging convex duality relationships between measure spaces and continuous functions, the DC-CDOC method enables the steering of population distributions towards desired predefined outcomes. This approach overcomes the limitations of previous GMFG frameworks, which lacked the capability to impose constraints on associated probability distributions. The integration of DC-CDOC with graphon mean field theory provides a powerful tool for analyzing and controlling large-scale multi-agent systems with complex network structures, while maintaining analytical tractability even as the number of agents grows very large.

The structure of the paper is as follows. Section 2 introduces the network system model and its equivalent representation by the graphon dynamical system. Section 3 provides a reformulation of the graphon dynamical system and the associated optimal control problem as the evolution of measures and the associated linear program in the space of signed measures.

Section 4 presents the optimality conditions for the associated graphon mean field system. The results are illustrated via a numerical example in Section 5. Concluding remarks and future research directions are discussed in Section 6.

2 Problem Formulation

2.1 Finite Population over Finite Network

Consider a population of N agents whose interactions are represented by a set of weighted finite graphs G_k described by its set of nodes (or vertices) $\mathcal{V}_k = \{1, \cdots, M_k\}$ and weights $g_{ij}^k \in [0,1]$ for $(i,j) \in \mathcal{V}_k \times \mathcal{V}_k$. Each node $l \in \mathcal{V}_k$ is occupied by a set of agents which is called a cluster of the population and is denoted by $\mathcal{C}_l$. Hence, the number of clusters is M_k and the total number of agents is $N = \sum_{l=1}^{M_k} |\mathcal{C}_l|$. For convenience of notation, we denote by $\mathcal{C}(i)$ the cluster to which agent i belongs.

Following the construction of Graphon mean field games (GMFG) in [11], the dynamics of agent i at time $t \in [t_0, t_f]$ are represented by

$$dx_i = \left(\frac{1}{|\mathcal{C}(i)|} \sum_{j \in \mathcal{C}(i)} f_0(x_i, u_i, x_j) + \frac{1}{M_k} \sum_{l=1}^{M_k} g_{\mathcal{C}(i)\mathcal{C}_l}^k \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} f(x_i, u_i, x_j) \right) dt + \sigma\, dw_i, \tag{1}$$

where $x_i \in \mathbb{R}^{n_x}$ is the state, $u_i \in \mathbb{R}^{n_u}$ is the input, $w_i \in \mathbb{R}^{n_w}$ is a standard Wiener process, and where w_i and w_j are independent processes for all $1 \le i \neq j \le N$. All initial states are taken to be independent and have finite second moment.

Each cluster $\mathcal{C}_l$'s empirical distribution of its individual states at time t is denoted by

$$\overline{\mu}_l(t,x) = \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} \boldsymbol{\delta}_{x_j(t)}(x), \tag{2}$$

where $\boldsymbol{\delta}_x$ denotes the Dirac delta measure at x. Similarly, the total population's empirical distribution of all individual states is denoted by

$$\overline{\mu}(t,x) = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\delta}_{x_i(t)}(x) \equiv \frac{1}{N} \sum_{l=1}^{M_k} |\mathcal{C}_l| \cdot \overline{\mu}_l(t,x). \tag{3}$$

Let $\{\mathsf{p}_l^d\}_{l=1}^{M_k}$ denote a set of desired probability distributions. The objective of the control problem is to determine the set of input policies $[\![u_i]\!] := \{u_i(s); s \in [t_0, t_f]\}$, $i \in \mathcal{C}_l$ for each cluster such that $\overline{\mu}_l(t_f, x)$ matches p_l^d, i.e.,

$$\mathbb{E} \int_{B_x} \overline{\mu}_l(t_f, x)\, dx = \int_{B_x} \mathsf{p}_l^d(dx), \tag{4}$$

for all $l \in \{1, \cdots, M_k\}$ and for every Borel set $B_x \in \mathbb{R}^{n_x}$, while minimizing the cost

$$\begin{aligned} &J_i([\![u_i]\!], [\![u_{-i}]\!]) \\ =&\mathbb{E}\Bigg[\int_{t_0}^{t_f} \left(\frac{1}{|\mathcal{C}(i)|} \sum_{j \in \mathcal{C}(i)} \ell_0(x_i, u_i, x_j) + \frac{1}{M_k} \sum_{l=1}^{M_k} g_{\mathcal{C}(i)\mathcal{C}_l} \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} \ell(x_i, u_i, x_j) \right) dt \Bigg]. \end{aligned} \tag{5}$$

The team's cost is the average cost over all agents in the network, i.e.,

$$J_{G_k} = \frac{1}{N} \sum_{i=1}^{N} J_i([\![u_i]\!], [\![u_{-i}]\!]). \tag{6}$$

Hence, the original problem is to determine the set of control strategies $\{[\![u_i]\!]\}_{i=1}^N$ while each agent with the dynamics (1) minimizes the cost (5) under the restriction to the constraint $\mathsf{P}\big(x_i(t_f)\big) = \int_{B_x} \mathsf{p}_l^d(dx)$.

2.2 Infinite Population over Finite Network

In the asymptotic local population limit, i.e., in the case where $|\mathcal{C}_l^k| \to \infty$ for all $l \in \{1, \cdots, M_k\}$, the dynamics at time $t \in [t_0, t_f]$ of a generic agent α in the cluster $\mathcal{C}_p$, i.e., $\alpha \in \mathcal{C}_p$, are given by

$$dx_\alpha = \bigg(\int_{\mathbb{R}^{n_x}} f_0(x_\alpha, u_\alpha, z)\mu_p(t, dz) + \frac{1}{M_k} \sum_{l=1}^{M_k} g_{\mathcal{C}_p\mathcal{C}_l} \int_{\mathbb{R}^{n_x}} f(x_\alpha, u_\alpha, z)\mu_l(t, dz) \bigg) dt + \sigma\, dw_\alpha, \tag{7}$$

where μ_p is the local mean field generated by agents at vertex $p \in \{1, \cdots, M_k\}$ at time $t \in [t_0, t_f]$ defined as

$$\mu_p(t, x) := \lim_{|\mathcal{C}_p| \to \infty} \overline{\mu}_p(t, x) \tag{8}$$

and the total population's distribution is denoted by $\mu(t, x)$ and is related to local mean fields by

$$\mu(t, x) = \sum_{l=1}^{M_k} \lim_{|\mathcal{C}_l| \to \infty} \frac{|\mathcal{C}_l|}{N} \cdot \mu_l(t, x). \tag{9}$$

Similarly, the cost is given as

$$\begin{aligned} &J_\alpha([\![u_\alpha]\!], [\![u_{-\alpha}]\!]) \\ =&\mathbb{E}\bigg[\int_{t_0}^{t_f} \bigg(\int_{\mathbb{R}^{n_x}} \ell_0(x_\alpha, u_\alpha, z)\mu_\alpha(t, dz) + \frac{1}{M_k} \sum_{l=1}^{M_k} g_{\mathcal{C}_p\mathcal{C}_l} \int_{\mathbb{R}^{n_x}} \ell(x_\alpha, u_\alpha, z)\mu_l(t, dz) \bigg) dt \bigg]. \end{aligned} \tag{10}$$

The objective of the control problem then becomes the determination of the family of input policies $[\![u_\alpha]\!] := \{u_\alpha(s); s \in [t_0, t_f]\}$, $\alpha \in \mathcal{C}_l$ such that $\mu_l(t_f, x)$ matches p_l^d, i.e.,

$$\int_{B_x} \mu_l(t_f, x)\, dx = \int_{B_x} \mathsf{p}_l^d(dx), \tag{11}$$

for all $l \in \{1, \cdots, M_k\}$, and for every Borel set $B_x \in \mathbb{R}^{n_x}$, while minimizing the cost (10).

It shall be remarked that the transition from a finite population to an infinite population over a finite network is a crucial step in the approach. While this approximation introduces some errors, it significantly simplifies the analysis and computation of optimal strategies. As it is a characteristic of mean field control approaches, the solutions obtained for the infinite population case are ε-optimal for the original finite population problem, with ε approaching zero as the number of agents increases.

2.3 Infinite Population over Infinite Network

Consider a uniform partition $\{I_1, \cdots, I_{M_k}\}$ of the interval $[0,1]$ with $I_1 = [0, \frac{1}{M_k}]$ and $I_j = (\frac{j-1}{M_j}, \frac{j}{M_k}]$ for $j \in \{2, \cdots, M_k\}$. We associate each cluster $p \in \{1, \cdots, M_k\}$ with the partition I_p. We denote by $g^k_{\alpha\beta}$, $(\alpha, \beta) \in [0,1]^2$, the *step function graphon* corresponding to the adjacency matrix of the underlying graph $G_k = [g^k_{ij}]_{(i,j)\in\mathcal{V}_k\times\mathcal{V}_k}$ which is defined by

$$g^k_{\nu\varphi} := \sum_{i=1}^{M_k}\sum_{j=1}^{M_k} \mathbb{I}_{I_i}(\nu)\cdot\mathbb{I}_{I_j}(\varphi)\cdot g^k_{ij}, \quad (\nu,\varphi)\in[0,1]^2, \tag{12}$$

where $\mathbb{I}_I$ is the indicator function of the interval I, as illustrated for an example graph with $M_k = 40$ in Figure 1-(a), and with $M_k = 100$ in Figure 1-(b).

We denote by $g_{\alpha\beta}$ the graphon limit of the finite graph $G_k = [g^k_{ij}]$ as $M_k \to \infty$ in the sense that

$$\lim_{M_k\to\infty} \max_{i\in\{1,\cdots,M_k\}} \sum_{j=1}^{M_k} \left| \frac{g^k_{\mathcal{C}^k_i\mathcal{C}^k_j}}{M_k} - \int_{I_j} g_{I_{i^*},\beta}\, d\beta \right| = 0, \tag{13}$$

(see [11, (H11)]), where I_{i^*} is the midpoint of the subinterval $I_i \in \{I_1, \cdots, I_{M_k}\}$ of length $1/M_k$, as illustrated in Figure 1-(c) for the limiting behavior of the graphs in Figure 1-(a) and (b). We also denote by $g_{\alpha,\bullet} := [g_{\alpha\beta}]_{\beta\in[0,1]}$ the section of g at α.

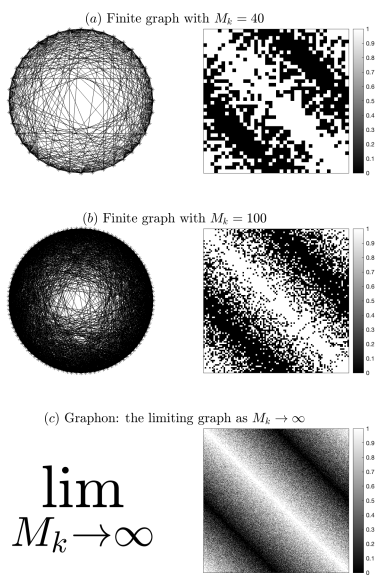


Figure 1 The graph (left) and the associated pixel representation (right) for a randomized nearest neighbors connectivity graph

Defining $\mu_G := \{\mu_\beta\}_{\beta\in[0,1]}$ as the ensemble of local mean fields, we introduce the shorthand notations

$$\widetilde{f}[x_\alpha, u_\alpha, \mu_G, g_{\alpha,\bullet}] := \int_{\mathbb{R}^{n_x}} f_0(x_\alpha, u_\alpha, z)\mu_\alpha(t, dz) + \int_0^1 \int_{\mathbb{R}^{n_x}} f(x_\alpha, u_\alpha, z)\ g_{\alpha\beta}\ \mu_\beta(t, dz)\, d\beta, \quad (14)$$

$$\widetilde{\ell}[x_\alpha, u_\alpha, \mu_G, g_{\alpha,\bullet}] := \int_{\mathbb{R}^{n_x}} \ell_0(x_\alpha, u_\alpha, z)\mu_\alpha(t, dz) + \int_0^1 \int_{\mathbb{R}^{n_x}} \ell(x_\alpha, u_\alpha, z)\ g_{\alpha\beta}\ \mu_\beta(t, dz)\, d\beta. \quad (15)$$

Hence, the local graphon dynamics is represented as

$$dx_\alpha = \widetilde{f}[x_\alpha, u_\alpha, \mu_G, g_{\alpha,\bullet}]dt + \sigma\, dw_\alpha \quad (16)$$

and the associated cost becomes

$$J_\alpha([\![u_\alpha]\!], [\![u_{-\alpha}]\!]) = \mathbb{E}\left[\int_{t_0}^{t_f} \widetilde{\ell}[x_\alpha, u_\alpha, \mu_G, g_{\alpha,\bullet}]dt\right]. \quad (17)$$

The objective of the control problem then becomes the determination of the family of input policies $[\![u_\alpha]\!] := \{u_\alpha(s); s \in [t_0, t_f]\}$, $\alpha \in [0, 1]$ such that each agent α minimizes the cost (17) under the restriction to the constraint that $\mu_\alpha(t_f, x)$ matches p_α^d, i.e.,

$$\int_{B_x} \mu_\alpha(t_f, x)\, dx = \mathsf{p}_\alpha^d(B_x), \quad (18)$$

for all $\alpha \in [0, 1]$, and for every Borel set $B_x \in \mathbb{R}^{n_x}$, while minimizing the cost (17). Thus, the value function of a representative agent $\alpha \in [0, 1]$ is defined as

$$V_\alpha\left(t_0, \rho_\alpha(t_0, \cdot)\right) = \inf_{[\![u_\alpha]\!]} \left\{\mathbb{E}\left[\int_{t_0}^{t_f} \widetilde{\ell}[x_\alpha, u_\alpha, \mu_G, g_{\alpha,\bullet}]dt\right] \text{ s.t.} \int_{B_x} \mu_\alpha(t_f, x)\, dx = \mathsf{p}_\alpha^d(B_x)\right\}. \quad (19)$$

It shall be remarked that the approximation of a finite network with an infinite network (graphon) is another key step in our methodology. This approximation allows us to capture the limiting behavior of large, complex networks while maintaining analytical tractability. The graphon representation provides a continuous approximation of discrete graph structures, enabling the application of powerful analytical techniques. While this introduces some approximation error, it allows us to study both dense and sparse network structures within a unified framework, providing insights that would be difficult to obtain from finite network analysis alone.

3 Convex Duality Optimal Control Formulation

In this section, we rewrite the case of infinite population over infinite network presented in Subsection 2.3 as a convex linear program in the space of signed measure and establish a graphon mean field version of the convex duality relations of the single agent counterpart in [40].

3.1 Strong Problem

We define the *input-state-time occupation measure* as

$$\mathfrak{m}_\alpha^{([\![u_\alpha]\!], [\![u_{-\alpha}]\!])}\left(B_t, B_x, B_u\right) := \mathbb{E}\left[\int_{B_t} \mathbb{I}_{B_x}\big(x_\alpha(s)\big) \cdot \mathbb{I}_{B_u}\big(u_\alpha(s)\big)\, ds\right], \quad (20)$$

for arbitrary Borel sets $B_t \subset [t_0, t_f]$, $B_x \subset \mathbb{R}^{n_x}$, $B_u \subset U$.

We denote by $\mathfrak{m}_G := \{\mathfrak{m}_\beta^{([\![u_\beta]\!],[\![u_{-\beta}]\!])}\}_{\beta\in[0,1]}$ an ensemble of local occupation measures. We also use the shorthand notation $\mathfrak{m}_G^\beta$ for its element corresponding to $\beta \in [0,1]$ which, evidently, $\mathfrak{m}_G^\alpha = \mathfrak{m}_\alpha^{([\![u_\alpha]\!],[\![u_{-\alpha}]\!])}$ for all $\alpha \in [0,1]$.

It is worth remarking that if $f_0(s,x,U) := \{l(s,x,u) : u \in \mathbb{R}^{n_u}\}$ and $f(s,x,U) := \{l(s,x,u) : u \in \mathbb{R}^{n_u}\}$ are convex for all $s \in [t_0,t_f]$, $x \in \mathbb{R}^{n_x}$, then under the same regularity assumptions as in [11], the Hölder continuity of the set of local mean fields $\mu_G := \{\mu_\alpha\}_{\alpha\in[0,1]}$ also holds for $\mathfrak{m}_G$.

In order to accommodate variations in an individual strategy while the remainder of the population maintain their current strategy (as is required for the study of Nash equilibrium as in [11] or for agent by agent optimization in teams as in here) we denote by $\mathfrak{m}_G^{\alpha^*}$ the unaltered strategy at the location $\alpha \in [0,1]$ (which is identified through the limiting value in $\mathfrak{m}_G$) whereas $\mathfrak{m}_\alpha^{([\![u_\alpha]\!],[\![u_{-\alpha}]\!])}$ is exclusively used for the current strategy of the agent at the location α.

It follows from the definition (20) that for every collection of admissible inputs $\{[\![u_\alpha]\!]\}_{\alpha\in[0,1]}$, measurable functions $\ell_0, \ell : [t_0,t_f) \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}$, with $\ell_0(s,x,\mathbb{R}^{n_u}) := \{l(s,x,u) : u \in \mathbb{R}^{n_u}\}$, and $\ell(s,x,\mathbb{R}^{n_u}) := \{l(s,x,u) : u \in \mathbb{R}^{n_u}\}$ convex for all $s \in [t_0,t_f]$, $x \in \mathbb{R}^{n_x}$, and given an ensemble of local occupation measures $\mathfrak{m}_G$, it is the case that

$$\begin{aligned}
&\mathbb{E}\Big[\int_{t_0}^{t_f}\Big(\int_{\mathbb{R}^{n_x}} \ell_0(x_\alpha,u_\alpha,z)\mu_\alpha(t,dz) + \int_0^1\int_{\mathbb{R}^{n_x}} \ell(x_\alpha,u_\alpha,z)\ g_{\alpha\beta}\ \mu_\beta(t,dz)\,d\beta\Big)dt\Big] \\
&= \int\limits_{[t_0,t_f]\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u}} \int_{\mathbb{R}^{n_x}\times\mathbb{R}^{n_u}} \ell_0(x,u,z)\mathfrak{m}_G^{\alpha^*}(t,dz,du')\,\mathfrak{m}_\alpha^{([\![u_\alpha]\!],[\![u_{-\alpha}]\!])}(dt,dx,du) \\
&\quad + \int\limits_{[t_0,t_f]\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u}} \int_{[0,1]\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u}} \ell(x,u,z)\mathfrak{m}_G^\beta(t,dz,du')d\beta\,\mathfrak{m}_\alpha^{([\![u_\alpha]\!],[\![u_{-\alpha}]\!])}(dt,dx,du) \\
&=: \big\langle \widetilde{\ell},\, \mathfrak{m}_\alpha^{([\![u_\alpha]\!],[\![u_{-\alpha}]\!])}\big\rangle_{\mathfrak{m}_G}.
\end{aligned} \tag{21}$$

It shall be emphasized that the bilinear operation $\langle \widetilde{\ell}, \mathfrak{m}_\alpha^{([\![u_\alpha]\!],[\![u_{-\alpha}]\!])}\rangle_{\mathfrak{m}_G}$ defined in (21) has, indeed, a linear dependence on $\mathfrak{m}_\alpha^{([\![u_\alpha]\!],[\![u_{-\alpha}]\!])}$ since $\mathfrak{m}_G^{\alpha^*}$ in the first set of integrations and $\mathfrak{m}_G^\beta$ in the second set of integrations are determined from $\mathfrak{m}_G$ which remain unchanged while $\mathfrak{m}_\alpha^{([\![u_\alpha]\!],[\![u_{-\alpha}]\!])}$ is permitted to change.

We also define the *terminal state occupation measure* as

$$\kappa_\alpha^{([\![u_\alpha]\!],[\![u_{-\alpha}]\!])}(B_x) := \mathsf{P}(x_\alpha(t_f) \in B_x), \tag{22}$$

for an arbitrary Borel set $B_x \subset \mathbb{R}^{n_x}$. We denote by $\mathcal{M}_S^{\alpha|G}$ the set of occupations measures corresponding to all $[\![u_\alpha]\!] \in \mathcal{U}$ with $[\![u_{-\alpha}]\!]$ given as part of $\mathfrak{m}_G$, i.e.,

$$\begin{aligned}
\mathcal{M}_S^{\alpha|G} := \Big\{&\mathfrak{m}_\alpha^{([\![u_\alpha]\!],[\![u_{-\alpha}]\!])} : [\![u_\alpha]\!] \in \mathcal{U}, \mathsf{P}\big(u_\beta \in B_u \,\big|\, t \in B_t, x_\beta \in B_x\big) \\
&= \int\limits_{B_t}\int\limits_{B_x}\int\limits_{B_u} \mathfrak{m}_G^\beta(dt,dx,du)\Big\}.
\end{aligned} \tag{23}$$

Thus, the graphon problem with the cost (10) is represented in terms of occupation measures in the form of the *strong problem*:

$$V_\alpha\left(t_0, \rho_\alpha(t_0, \cdot)\right) = \inf_{\mathfrak{m}_\alpha^{(\llbracket u_\alpha \rrbracket, \llbracket u_{-\alpha} \rrbracket)} \in \mathcal{M}_S^{\alpha|G}} \left\{ \left\langle \widetilde{\ell}, \mathfrak{m}_\alpha^{(\llbracket u_\alpha \rrbracket, \llbracket u_{-\alpha} \rrbracket)} \right\rangle \text{ s.t. } \kappa^{(\llbracket u_\alpha \rrbracket, \llbracket u_{-\alpha} \rrbracket)} = \mathsf{p}_\alpha^d \right\}. \tag{SP}$$

We refer to the reformulation (SP) as the *strong problem* due to the direct correspondence between (19) and (SP). We note that for every measurable function $\widetilde{\ell}$, the problem (SP) is an optimization problem with a linear objective defined over the space $\mathcal{M}_S^{\alpha|G}$. However, the identification of this space is not straightforward as it is associated with implementing all admissible inputs $\llbracket u_\alpha \rrbracket \in \mathcal{U}$ on the stochastic differential equation (16). To address this issue, we present in Subsection 3.2 a problem defined directly over the space of measures which tightly embeds our original problem.

3.2 Weak Problem

It follows from [40, Lemma 2] that for every twice continuously differentiable function $v \in C^2\left([t_0, t_f] \times \mathbb{R}^{n_x}\right)$, Dynkin's formula is expressed as

$$\left\langle \mathcal{A}_{\alpha|\mu_G} v, \mathfrak{m}_\alpha^{(\llbracket u_\alpha \rrbracket, \llbracket u_{-\alpha} \rrbracket)} \right\rangle_{\mathfrak{m}_G} = \left\langle v, \kappa^{(\llbracket u_\alpha \rrbracket, \llbracket u_{-\alpha} \rrbracket)} \right\rangle_{\mathfrak{m}_G} - \left\langle v, \rho(t_0, \cdot) \right\rangle_{\mathfrak{m}_G}, \tag{24}$$

where $\mathcal{A}_{\alpha|\mu_G}$ is the infinitesimal operator of the Markov process (16), written as

$$\begin{aligned} &\mathcal{A}^u_{\alpha|\mu_G} v(t,x) \\ =&\frac{\partial v(t,x)}{\partial t} + \frac{1}{2}\mathrm{tr}\left(\sigma^{\mathrm{T}}\sigma\, \frac{\partial^2 v(t,x)}{\partial x^2}\right) \\ &+ \left[\frac{\partial v(t,x)}{\partial x}\right]^{\mathrm{T}} \left(\int_{\mathbb{R}^{n_x}} f_0(x_\alpha, u_\alpha, z)\mu_\alpha(t,dz) + \int_0^1 \int_{\mathbb{R}^{n_x}} f(x_\alpha, u_\alpha, z)\, g_{\alpha\beta}\, \mu_\beta(t,dz)\, d\beta \right). \end{aligned} \tag{25}$$

Defining $\mathcal{A}^*_{\alpha|\mathfrak{m}_G}$ as the adjoint of (25) defined as the operator satisfying

$$\left\langle \mathcal{A}_{\alpha|\mu_G} v, \mathfrak{m} \right\rangle_{\mathfrak{m}_G} = \left\langle v, \mathcal{A}^*_{\alpha|\mathfrak{m}_G}\mathfrak{m} \right\rangle_{\mathfrak{m}_G}, \tag{26}$$

for every Borel measure $\mathfrak{m}$, and any twice continuously differentiable function $v \in C^2([t_0, t_f) \times \mathbb{R}^{n_x})$. Hence, for every $\mathfrak{m}_\alpha^{(\llbracket u_\alpha \rrbracket, \llbracket u_{-\alpha} \rrbracket)} \in \mathcal{M}_S^{\alpha|G}$, it follows from [40, Theorem 1] that

$$\mathcal{A}^*_{\alpha|\mathfrak{m}_G} \mathfrak{m}_\alpha^{(\llbracket u_\alpha \rrbracket, \llbracket u_{-\alpha} \rrbracket)} = \mathsf{p}_\alpha^d - \rho(t_0, \cdot). \tag{27}$$

Accordingly, we define the weak problem in the space of signed measures, and the associated weak value function of agent α as

$$W_\alpha\left(t_0, \rho_\alpha(t_0, \cdot)\right) = \inf_{\mathfrak{m} \in \mathfrak{M}_\pm} \left\{ \left\langle \widetilde{\ell}, \mathfrak{m} \right\rangle, \text{ s.t. } \mathfrak{m} \in \mathcal{M}_{PB} \cap \mathcal{M}_{\mathcal{A}} \right\}, \tag{28}$$

where

$$\mathcal{M}_{PB} := \left\{ \mathfrak{m} \in \mathfrak{M}_+\left([t_0, t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}\right) : \|\mathfrak{m}\| \le (t_f - t_0) \right\}, \tag{29}$$

$$\mathcal{M}_{\mathcal{A}} := \left\{ \mathfrak{m} \in \mathfrak{M}_\pm\left([t_0, t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}\right) : \mathcal{A}^*_{\alpha|\mathfrak{m}_G}\mathfrak{m} = \mathsf{p}_\alpha^d - \rho(t_0, \cdot) \right\}. \tag{30}$$

Over the compact Hausdorff space $[t_0, t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}$, the Banach space of continuous functions $C\left([t_0, t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}\right)$ equipped with the sup-norm has a topological dual $C^*([t_0, t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u})$ that is isometrically isomorphic to $\mathfrak{M}_\pm\left([t_0, t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}\right)$ equipped with the norm $\|\mathfrak{m}\| := \int d\mathfrak{m}^+ + \int d\mathfrak{m}^-$. The norm topology of C and the weak dual topology of $\mathfrak{M}_\pm$ are compatible with the pairing defined by the bilinear form $\langle c, \mathfrak{m}\rangle$ for all $c \in C\left([t_0, t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}\right)$, and $\mathfrak{m} \in \mathfrak{M}_\pm\left([t_0, t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}\right)$.

Endowing the space of continuous functions with the topology of the sup-norm and endowing the space of signed measures, $\mathfrak{M}_\pm$, with a weak dual topology, based upon the dual relationship between the space of measures and that of the continuous functions, we will argue below that it follows that $\mathcal{M}_{PB} \cap \mathcal{M}_{\mathcal{A}}$ is w*-compact and hence, the infimum in (28) is achieved and is equal to the minimum. Thus, we directly define the *weak problem* as

$$W_\alpha\left(t_0, \rho_\alpha(t_0, \cdot)\right) = \min_{\mathfrak{m} \in \mathfrak{M}_\pm} \left\{ \left\langle \widetilde{\ell}, \mathfrak{m} \right\rangle, \text{ s.t. } \mathfrak{m} \in \mathcal{M}_{PB} \cap \mathcal{M}_{\mathcal{A}} \right\}. \tag{WP}$$

3.3 Fenchel Normal Form

Using the notion of weak value function, we reformulate the convexly constrained linear program as an unconstrained convex problem by introducing the functionals h_1 and h_2 : $\mathfrak{M}_\pm\left([t_0, t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}\right) \to \overline{\mathbb{R}}$ defined by

$$h_1(\mathfrak{m}) := \begin{cases} \left\langle \widetilde{\ell}, \mathfrak{m} \right\rangle, & \text{if } \mathfrak{m} \in \mathcal{M}_{PB}, \\ +\infty, & \text{otherwise}, \end{cases} \tag{31}$$

$$h_2(\mathfrak{m}) := \begin{cases} 0, & \text{if } \mathfrak{m} \in \mathcal{M}_{\mathcal{A}}, \\ -\infty, & \text{otherwise}. \end{cases} \tag{32}$$

Both h_1 and $-h_2$ are convex and lower semi-continuous[44] and, hence,

$$W_\alpha\big(t_0, \rho_\alpha(t_0, \cdot)\big) = \min_{\mathfrak{m} \in \mathfrak{M}_\pm([t_0, t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u})} \left\{ h_1(\mathfrak{m}) - h_2(\mathfrak{m}) \right\}. \tag{33}$$

3.4 Legendre-Fenchel Transform

The real-valued functional h_1 is convex and its convex conjugate (Legendre-Fenchel transform) is defined by

$$h_1^*(c) := \sup_{\mathfrak{m} \in \mathfrak{M}_\pm([t_0, t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u})} \left\{ \left\langle c, \mathfrak{m} \right\rangle - h_1(\mathfrak{m}) \right\}. \tag{34}$$

Lemma 3.1

$$h_1^*(c) = (t_f - t_0) \cdot \|(c - \widetilde{\ell})^+\|, \tag{35}$$

where $(f)^+$ *denotes the positive part of the function* f, *i.e.,* $f^+(x) = \max\{0, f(x)\}$.

Proof The proof follows a similar structure to the approach in [44, Lemma 4.1], with necessary adaptations specific to the current framework. ∎

For the concave functional h_2 the Legendre-Fenchel transform is defined as

$$h_2^*(c) := \inf_{\mathfrak{m}\in\mathfrak{M}_\pm([t_0,t_f]\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u})} \Big\{\big\langle c,\mathfrak{m}\big\rangle - h_2\big(\mathfrak{m}\big)\Big\}. \tag{36}$$

Lemma 3.2

$$h_2^*(c) = \begin{cases} \lim\limits_{k\to\infty}\Big(V_\alpha\big(t_0,\rho_\alpha(t_0,\cdot)\big) - \big\langle v_k,\mathsf{p}_d\big\rangle\Big), & \textit{if } c = -\lim\limits_{k\to\infty}\mathcal{A}v_k, \\ -\infty, & \textit{otherwise.} \end{cases} \tag{37}$$

Proof The proof is a modification of [44, Lemma 4.2], with necessary adaptations specific to the current framework. ∎

3.5 The Hamilton-Jacobi Problem

Theorem 3.3

$$W_\alpha\big(t_0,\rho_\alpha(t_0,\cdot)\big) = \sup_{v\in C^2([t_0,t_f]\times\mathbb{R}^{n_x})}\Big\{\int_{\mathbb{R}^{n_x}} v_\alpha(t_0,x)\rho_\alpha(t_0,dx) - \int_{\mathbb{R}^{n_x}} v(t_f,x)\mathsf{p}_d(dx), \quad \text{s.t. } \mathcal{A}_{\alpha|\mu_G}v+\ell\ge 0\Big\}. \tag{38}$$

Proof Applying the Rockafellar duality theorem[44] to $C^*([t_0,t_f]\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u}) = \mathfrak{M}_\pm([t_0, t_f]\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u})$, we obtain

$$\min_{\mathfrak{m}\in\mathfrak{M}_\pm([t_0,t_f]\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u})}\Big\{h_1(\mathfrak{m}) - h_2(\mathfrak{m})\Big\} = \sup_{c\in C([t_0,t_f]\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u})}\Big\{h_2^*(c) - h_1^*(c)\Big\}, \tag{39}$$

whenever the set $\{c : h_2^*(c) > -\infty\}$ contains a continuity point of $h_1^*(c)$ that is finite. Since h_1^* is continuous and finite on whole $C\left([t_0,t_f]\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u}\right)$ and h_2^* is not identically $-\infty$ we deduce that (39) holds. The substitution of (39) into (33) yields

$$\begin{aligned} W_\alpha\big(t_0,\rho_\alpha(t_0,\cdot)\big) &= \sup_{c\in C([t_0,t_f]\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u})}\Big\{h_2^*(c) - h_1^*(c)\Big\} \\ &\overset{(35)}{\underset{(37)}{=}} \sup_{c\in C}\Big\{\lim_{k\to\infty}\Big(V_\alpha\big(t_0,\rho_\alpha(t_0,\cdot)\big) - \big\langle v_k,\mathsf{p}_d\big\rangle\Big) - (t_f-t_0)\|(c-\widetilde{\ell})^+\| \\ &\qquad \text{s.t. } c = -\lim_{k\to\infty}\mathcal{A}_{\alpha|\mu_G}v_k\Big\}. \end{aligned} \tag{40}$$

Using the fact that $\{\mathcal{A}_{\alpha|\mu_G}v : v\in C^2([t_0,t_f]\times\mathbb{R}^{n_x})\}$ is dense in $\{c\in C([t_0,t_f]\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u}) : h_2^*(c) > -\infty\}$, we obtain

$$\begin{aligned} &W_\alpha\big(t_0,\rho_\alpha(t_0,\cdot)\big) \\ &= \sup_{v\in C^2}\Big\{\Big(\big\langle v,\rho_\alpha(t_0,\cdot)\big)\big\rangle - \big\langle v,\mathsf{p}_d\big\rangle\Big) - (t_f-t_0)\|(c-\widetilde{\ell})^+\| \text{ s.t. } c = -\mathcal{A}_{\alpha|\mu_G}v\Big\}. \end{aligned} \tag{41}$$

To conclude the proof it suffices to show that for every $v\in C^2([t_0,t_f]\times\mathbb{R}^{n_x})$ there exists a $\widehat{v}\in C^2([t_0,t_f]\times\mathbb{R}^{n_x})$ such that $\mathcal{A}_{\alpha|\mu_G}\widehat{v}+\widetilde{\ell}\ge 0$ and $\widehat{v}(t_0,\rho_\alpha(t_0,\cdot))\ge v(t_0,\rho_\alpha(t_0,\cdot))$; this is possible by defining $\widehat{v} := v - (t_f - t_0)\|(\mathcal{A}_{\alpha|\mu_G} v + \widetilde{\ell})^-\|$, which yields that

$$\begin{aligned}\mathcal{A}_{\alpha|\mu_G}\widehat{v} + \widetilde{\ell} \equiv\ & \mathcal{A}_{\alpha|\mu_G} v + \widetilde{\ell} + \|(\mathcal{A}_{\alpha|\mu_G} v + \widetilde{\ell})^-\| \\ \geq\ & \mathcal{A}_{\alpha|\mu_G} v + \widetilde{\ell} + \sup_{(s,x,u)\in[t_0,t_f]\times X\times U} |(\mathcal{A}^u_{\alpha|\mu_G} v(s,x) + \widetilde{\ell}[x_\alpha, u_\alpha, \mu_G, g_{\alpha,\bullet}])^-| \\ \geq\ & 0.\end{aligned} \tag{42}$$

The proof is completed. ∎

3.6 Equivalence of the Weak and Strong Problems

It follows from the definitions (WP) and (SP) of the weak and strong value functions that

$$W_\alpha\big(t_0, \rho_\alpha(t_0,\cdot)\big) = \min_{\mathfrak{m}\in\mathcal{M}_W} \big\langle \widetilde{\ell}, \mathfrak{m}\big\rangle \leq V\big(t_0, \rho_\alpha(t_0,\cdot)\big) = \inf_{\mathfrak{m}^{[\![u]\!]}\in\mathcal{M}_S} \Big\{\big\langle \widetilde{\ell}, \mathfrak{m}^{[\![u]\!]}\big\rangle\Big\}. \tag{43}$$

Since $\mathcal{M}_S \subset \mathcal{M}_W := \mathcal{M}_{PB} \cap \mathcal{M}_{\mathcal{A}}$. In order to show the equivalence of the weak and the strong problems, we need to show that strict inequality cannot hold and hence, the weak and the strong value functions coincide.

Theorem 3.4 *The weak and the strong value functions are equal, i.e.,*

$$W_\alpha\big(t_0, \rho_\alpha(t_0,\cdot)\big) = V_\alpha\big(t_0, \rho_\alpha(t_0,\cdot)\big). \tag{44}$$

Proof Let's assume that this is not true, i.e., there exist $(\mathfrak{m}_0, \kappa_0) \in \mathcal{M}_{PB} \cap \mathcal{M}_{\mathcal{A}} \setminus \mathcal{M}_S$ such that

$$W_\alpha\big(t_0, \rho_\alpha(t_0,\cdot)\big) = \big\langle \widetilde{\ell}, \mathfrak{m}_0\big\rangle < V_\alpha\big(t_0, \rho_\alpha(t_0,\cdot)\big) = \inf_{\mathfrak{m}^{[\![u]\!]}\in\mathcal{M}_S} \Big\{\big\langle \widetilde{\ell}, \mathfrak{m}^{[\![u]\!]}\big\rangle\Big\}. \tag{45}$$

This means that the w*-continuous linear functional $\big\langle \widetilde{\ell}, \mathfrak{m}\big\rangle$ separates an element $\mathfrak{m}_0 \in \mathcal{M}_{PB} \cap \mathcal{M}_{\mathcal{A}}$ from the w* convex closure $\overline{\mathrm{cov}\mathcal{M}_S}$ of $\mathcal{M}_S$. Then by [44, Theorem 3], for every $\varepsilon > 0$, there exists $V^{(\varepsilon)}$ whose partial derivatives $V^{(\varepsilon)}_t$, $V^{(\varepsilon)}_{x_i}$, $V^{(\varepsilon)}_{x_i x_j}$ are defined almost everywhere, are essentially bounded and, further,

$$\|V - V^{(\varepsilon)}\| \leq \varepsilon, \quad \mathcal{A}^u_{\alpha|\mu_G} V^{(\varepsilon)}(s,x) + \widetilde{\ell}[x_\alpha, u_\alpha, \mu_G, g_{\alpha,\bullet}] \geq 0, \tag{46}$$

for all $(s,x,u) \in [t_0,t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}$. Since $V^{(\varepsilon)}$ is not necessarily in $C^2([t_0,t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u})$, in order to apply Dynkin's formula (24), we also need to invoke [44, Lemma 5.1] that for every $\delta > 0$, there exists $V^{(\varepsilon,\delta)} \in C^2([t_0,t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u})$ for which

$$\|V^{(\varepsilon,\delta)} - V^{(\varepsilon)}\| < \delta, \quad \text{s.t. } \|\mathcal{A}_{\alpha|\mu_G} V^{(\varepsilon,\delta)}\| \leq \|\mathcal{A}_{\alpha|\mu_G} V^{(\varepsilon)}\| + \delta, \tag{47}$$

$$\mathcal{A}_{\alpha|\mu_G} V^{(\varepsilon,\delta)} + \widetilde{\ell} \geq -\delta, \quad \text{on } [t_0+\delta, t_f-\delta] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}. \tag{48}$$

Then by (24),

$$\begin{aligned} V^{(\varepsilon,\delta)}\big(t_0, \rho_\alpha(t_0,\cdot)\big) - \big\langle V^{(\varepsilon,\delta)}_T, \mathsf{p}_d\big\rangle =& -\big\langle \mathcal{A}_{\alpha|\mu_G} V^{(\varepsilon,\delta)}, \mathfrak{m}_0\big\rangle \\ \leq& \int_{[\delta,T-\delta)\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u}} \widetilde{\ell}\, d\mathfrak{m}_0 + \delta \int_{[\delta,T-\delta)\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u}} d\mathfrak{m}_0 \\ &+ \|\mathcal{A}_{\alpha|\mu_G} V^{(\varepsilon,\delta)}\| \int_{\big\{[0,\delta)\cup[T-\delta,T)\big\}\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u}} d\mathfrak{m}_0 \end{aligned} \tag{49}$$

and hence,

$$V^{(\varepsilon,\delta)}\big(t_0,\rho_\alpha(t_0,\cdot)\big) - \big\langle V_T^{(\varepsilon,\delta)}, \mathsf{p}_d\big\rangle \le \big\langle \widetilde{\ell}, \mathfrak{m}_0\big\rangle + 2\cdot\delta\cdot(t_f-t_0)\big(1+\|\mathcal{A}_{\alpha|\mu_G}V^{(\varepsilon,\delta)}\|\big). \tag{50}$$

Employing $\|V - V^{(\varepsilon,\delta)}\| < \varepsilon + \delta$ from (46) and (47), and choosing first ε then δ sufficiently small, we arrive at

$$V\big(t_0,\rho_\alpha(t_0,\cdot)\big) - \big\langle V_T^{(\varepsilon,\delta)}, \mathsf{p}_d\big\rangle \le \big\langle \widetilde{\ell}, \mathfrak{m}_0\big\rangle, \tag{51}$$

that is in contradiction with the hypothesis (45). Therefore, the equivalence (44) holds true. ∎

4 Graphon Mean Field Optimality Conditions

Assuming that all measures μ_α, $\alpha \in [0,1]$, possess density functions denoted by ρ_α, we invoke the ε-Nash equilibrium conditions of [11] for graphon mean field games and employ the results of the convex duality based formulation for probability assignment established in Section 3, in order to obtain the following identification of optimal solutions for assigning probability measures to large scale network of nonlinear stochastic agents.

Theorem 4.1 (Main Result) *The optimal solution for the problem of assigning probability distributions to the infinite population over infinite network of Subsection* 2.3 *is identified from the following relations:*

$$\begin{aligned} V_\alpha(t_0,\rho_\alpha(t_0,\cdot)) = \sup_{v_\alpha\in C^2([t_0,t_f]\times\mathbb{R}^{n_x})} \Bigg\{ &\int_{\mathbb{R}^{n_x}} v_\alpha(t_0,x)\rho_\alpha(t_0,dx) - \int_{\mathbb{R}^{n_x}} v_\alpha(t_f,x)\mathsf{p}_\alpha^d(dx), \\ \text{s.t. } &\frac{\partial v_\alpha(t,x)}{\partial t} + \left[\frac{\partial v_\alpha(t,x)}{\partial x}\right]^{\mathrm{T}} \widetilde{f}[x,u,\mu_G;g_\alpha] + \frac{1}{2}\mathrm{tr}\Big(\sigma^{\mathrm{T}}\sigma\, \frac{\partial^2 v_\alpha(t,x)}{\partial x^2}\Big) \\ &+ \widetilde{\ell}(x,u,\mu_G;g_\alpha] \ge 0, \text{for all } (t,x,u)\in[t_0,t_f]\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_u} \Bigg\}, \end{aligned} \tag{52}$$

together with

$$\frac{\partial \rho_\alpha(t,x)}{\partial t} = -\frac{\partial\{\widetilde{f}[x,u_\alpha^*(t,x),\mu_G;g_\alpha]\rho_\alpha(t,x)\}}{\partial x} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}[\sigma\,\sigma^{\mathrm{T}}]_{ij}\frac{\partial^2\rho_\alpha(t,x)}{\partial x_i\partial_j}, \tag{53}$$

with $u_\alpha^*(t,x)$ *determined from*

$$u_\alpha^*(t,x) \in \operatorname*{arg\,inf}_{u\in\mathbb{R}^{n_u}}\left\{\widetilde{\ell}(x,u,\mu_G;g_\alpha] + \left[\frac{\partial v_\alpha(t,x)}{\partial x}\right]^{\mathrm{T}}\widetilde{f}[x,u,\mu_G;g_\alpha]\right\}. \tag{54}$$

Proof The result is obtained from the ε-Nash optimality conditions of [11] for graphon mean field games with the substitution of (44) from Theorem 3.4 into (38) from Theorem 3.3. The proof is completed. ∎

5 Numerical Illustration

Consider a population of 2000 agents in $M_k = 40$ clusters with $|\mathcal{C}_l| = 50$, for all $l \in \{1,\cdots,M_k\}$ whose graph G_k is a realization of a randomized nearest neighbors connectivity graph as displayed in Figure 1-(a).

Agents dynamics are represented as

$$dx_i = \left(\begin{bmatrix} 0 & 1 \\ 2 & -3 \end{bmatrix} x_i + \frac{1}{40} \sum_{l=1}^{40} g^k_{\mathcal{C}(i)\mathcal{C}_l} \begin{bmatrix} 0 & 1 \\ 2 & -3 \end{bmatrix} \overline{\mu}_l(t,x) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_s \right) ds + \begin{bmatrix} 0 \\ 1 \end{bmatrix} dw_s, \tag{55}$$

and their costs are given as

$$J_i\big(\llbracket u_i \rrbracket, \llbracket u_{-i} \rrbracket\big) := \mathbb{E}\left[\int_{t_0}^{t_f} \frac{1}{2} u_i^2 ds + \frac{1}{40} \sum_{l=1}^{40} g^k_{\mathcal{C}(i)\mathcal{C}_l} \frac{1}{2} \left\| x_i - \overline{\mu}_l(t,x) \right\|^2 \right]. \tag{56}$$

Due to the linearity of the system and the Gaussian form of the desired distribution, the graphon mean field optimal solutions of Section 4 can be solved by the procedure proposed in [55, Subsection 9.5.1] combined with LQG graphon mean field solutions of [56] as follows.

Step 1 Set the iteration counter to $k = 0$, and initiate the algorithm with an arbitrary terminal cost function $L^k_\alpha(x) = \frac{1}{2} x^{\mathrm{T}} H^k_\alpha x + (s^k_\alpha)^{\mathrm{T}} x + \delta^k_\alpha$.

Step 2 Solve the Riccati equations of [56] for the graphon mean field problem with $v^k_\alpha(t_f, x) = L^k_\alpha(x)$.

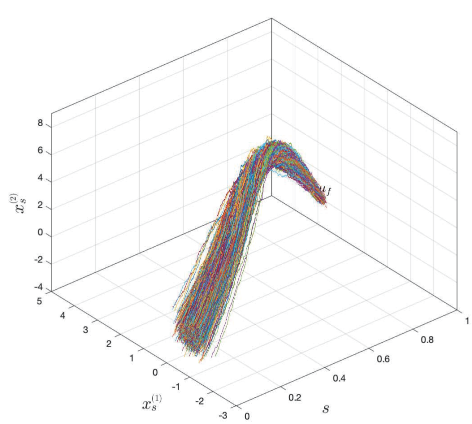
Step 3 Evaluate $\int_{\mathbb{R}^{n_x}} v^k_\alpha(t_0, x)\rho(t_0, dx) - \int_{\mathbb{R}^{n_x}} v^k_\alpha(t_f, x)\mathsf{p}_d(dx)$ from the solution to the Riccati equations.

Step 4 Update μ^k_G to the corresponding ensemble of mean fields.

Step 5 Update $L^{k+1}(x)$ using an ascent direction for the cost

$$\int_{\mathbb{R}^{n_x}} v^k_\alpha(t_0, x)\rho(t_0, dx) - \int_{\mathbb{R}^{n_x}} v^k_\alpha(t_f, x)\mathsf{p}_d(dx).$$

With the consideration of the time horizon as $[t_0, t_f] = [0, 1]$ and with the initial and desired probability distributions of all clusters given, respectively, as $\rho_l(t_0, \cdot) \sim \mathcal{N}\left(\left[\begin{smallmatrix} -1 \\ -1 \end{smallmatrix}\right], \left[\begin{smallmatrix} 1/4 & 0 \\ 0 & 1/2 \end{smallmatrix}\right] \right)$, and $\mathsf{p}^d_l \sim \mathcal{N}\left(\left[\begin{smallmatrix} 3 \\ -2 \end{smallmatrix}\right], \left[\begin{smallmatrix} 1/10 & 0 \\ 0 & 1/10 \end{smallmatrix}\right] \right)$, for $l \in \{1, \cdots, M_k\}$, a sample realization of the trajectories of the agents as well as their input process are displayed in Figure 2. The evolution of the empirical distribution of the population is displayed in increments of 0.1 seconds in Figure 3.



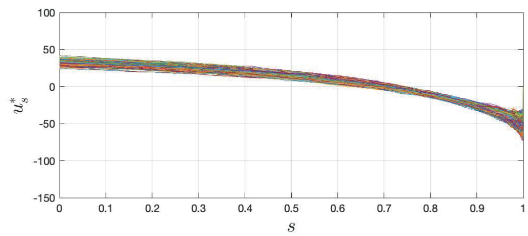


Figure 2 Sample paths for the evolution of the states (top) and their components (the two middle) as well as the corresponding input processes

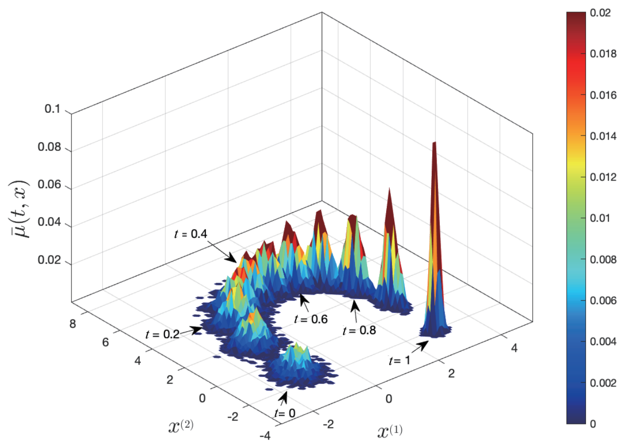


Figure 3 The evolution of the empirical distribution of the population displayed in increments of 0.1 seconds

6 Concluding Remarks

This article introduces an approach to approximately control the probability distributions associated with networks of nonlinear stochastic systems by leveraging the inherent limit described by graphon mean field systems. The key advantage of this approach, compared to seeking exact optimality through consideration of all agent interactions, is its reduced computational demand, requiring only the study of a parameterized family of representative agents.

The proposed approach extends the graphon mean field games framework by incorporating terminal constraints on the associated probability measures, making it possible to steer population distributions to desired values, hence broadening the applicability of GMFG to scenarios where specific distributional outcomes are essential.

It shall be remarked that due to the coupling of the value function determination with the identification of the associated measures, the associated numerical schemes often require iterations over the two segments of the solution methodology.

It is important to note that our approach to the graphon mean field distribution control problem is framed within a team setting, with the satisfaction of certain probability distributions is considered a collective goal. This choice was made because the act of steering population distributions towards desired outcomes inherently requires coordination among agents. However, we acknowledge that this formulation yields local ε-optimality, similar to the person-by-person optimal (PBPO) solutions typically found in mean field teams (MFT) and Graphon mean field teams (GMFT) literature. The global optimality of PBPO solutions in general non-linear, non-Gaussian setups remains an open question, except for some simple instances of linear-quadratic models. Investigating whether the current results can be extended or adapted to a game-theoretic framework, where agents may have conflicting objectives, could provide valuable insights into the broader applicability of graphon mean field distribution control techniques.

The studied example in this article demonstrates the ability of the proposed method to steer probability distributions to desired values for a network of linear systems interacting over a graph with a known limiting graphon. However, identifying the graphon associated with a given graph remains challenging. A plausible empirical approach involves fitting two-dimensional Fourier series to the step function representation of the adjacency matrix. Such parametric modeling could resemble techniques used in statistics and system identification. Additionally, due to the compactness of graphon operators, representations or approximations through simple spectral decomposition are possible[15].

Future research directions include the development of computational algorithms for numerical solutions of graphon mean field distribution control problems in general nonlinear cases, as well as the accommodation of hybrid systems features, particularly controlled and autonomous switchings with exact equality and almost surely equality constraints as switching manifolds as those expressed in [57].

In summary, this work advances theoretical foundations in network control while paving the way for practical applications across diverse fields. By addressing the outlined challenges, future research can further develop graphon mean field distribution control as a tool for managing complex, large-scale networked systems.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Caines P E and Huang M Y, Graphon mean field games and the GMFG equations: ε-Nash equilibria, *Proceedings of the* 58*th IEEE Conference on Decision and Control* (*CDC*), 2019, 286–292.

[2] Huang M Y, Malhamé R P, and Caines P E, Large population stochastic dynamic games:

Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle, 2006, DOI: 10.4310/CIS.2006.v6.n3.a5.

[3] Huang M Y, Caines P E, and Malhamé R P, Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε-Nash equilibria, *IEEE Transactions on Automatic Control*, 2007, **52**(9): 1560–1571.

[4] Lasry J M and Lions P L, Jeux à champ moyen, I-Le cas stationnaire, *Comptes Rendus Mathématique*, 2006, **343**(9): 619–625.

[5] Lasry J M and Lions P L, Jeux à champ moyen, II-Horizon fini et contrôle optimal, *Comptes Rendus. Mathématique*, 2006, **343**(10): 679–684.

[6] Lasry J M and Lions P L, Mean field games, *Japanese Journal of Mathematics*, 2007, **2**(1): 229–260.

[7] Guéant O, Lasry J M, and Lions P L, *Mean Field Games and Applications*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, 205–266.

[8] Carmona R and Delarue F, Mean field forward-backward stochastic differential equations, 2013, arXiv: 1211.4186.

[9] Achdou Y and Capuzzo-Dolcetta I, Mean field games: Numerical methods, *SIAM Journal on Numerical Analysis*, 2010, **48**(3): 1136–1162.

[10] Lovász L, *Large Networks and Graph Limits*, Volume 60, American Mathematical Soc., Providence, 2012.

[11] Caines P E and Huang M Y, Graphon mean field games and their equations, *SIAM Journal on Control and Optimization*, 2021, **59**(6): 4373–4399.

[12] Caines P E, Ho D, Huang M Y, et al., On the graphon mean field game equations: Individual agent affine dynamics and mean field dependent performance functions, *ESAIM: Control, Optimisation and Calculus of Variations*, 2022, **28**: 24, DOI: 10.1051/cocv/2022020.

[13] Carmona R, Cooney D B, Graves C V, et al., Stochastic graphon games: I. The static case, *Mathematics of Operations Research*, 2022, **47**(1): 750–778.

[14] Delarue F, Mean field games: A toy model on an Erdös-Renyi graph, *ESAIM: Proceedings and Surveys*, 2017, **60**: 1–26.

[15] Gao S and Caines P E, Graphon control of large-scale networks of linear systems, *IEEE Transactions on Automatic Control*, 2020, **65**(10): 4090–4105.

[16] Lacker D and Soret A, A case study on stochastic games on large graphs in mean field and sparse regimes, *Mathematics of Operations Research*, 2022, **47**(2): 1530–1565.

[17] Grigoriadis K M and Skelton R E, Minimum-energy covariance controllers, *Automatica*, 1997, **33**(4): 569–578.

[18] Hotz A and Skelton R E, Covariance control theory, *International Journal of Control*, 1987, **46**(1): 13–32.

[19] Xu J H and Skelton R E, An improved covariance assignment theory for discrete systems, *IEEE Transactions on Automatic Control*, 1992, **37**(10): 1588–1591.

[20] Yasuda K, Skelton R E, and Grigoriadis K M, Covariance controllers: A new parametrization of the class of all stabilizing controllers, *Automatica*, 1993, **29**(3): 785–788.

[21] Bakolas E, Finite-Horizon covariance control for discrete-time stochastic linear systems subject to input constraints, *Automatica*, 2018, **91**: 61–68.

[22] Bakolas E, Optimal covariance control for discrete-time stochastic linear systems subject to

constraints, *Proceedings of the 55th IEEE Conference on Decision and Control*, 2016, 1153–1158.

[23] Beghi A, On the relative entropy of discrete-time Markov processes with given end-point densities, *IEEE Transactions on Information Theory*, 1996, **42**(5): 1529–1535.

[24] Chen Y X, Georgiou T T, and Pavon M, Optimal steering of a linear stochastic system to a final probability distribution, Part I, *IEEE Transactions on Automatic Control*, 2015, **61**(5): 1158–1169.

[25] Chen Y X, Georgiou T T, and Pavon M, Optimal steering of a linear stochastic system to a final probability distribution, Part II, *IEEE Transactions on Automatic Control*, 2015, **61**(5): 1170–1180.

[26] Chen Y X, Georgiou T T, and Pavon M, Stochastic control liaisons: Richard Sinkhorn meets Gaspard Monge on a Schrodinger bridge, *Siam Review*, 2021, **63**(2): 249–313.

[27] Goldshtein M and Tsiotras P, Finite-horizon covariance control of linear time-varying systems, *Proceedings of the 56th IEEE Conference on Decision and Control*, 2017, 3606–3611.

[28] Halder A and Wendel E, Finite horizon linear quadratic Gaussian density regulator with wasserstein terminal cost, *Proceedings of the American Control Conference*, 2016, 7249–7254.

[29] Okamoto K and Tsiotras P, Optimal stochastic vehicle path planning using covariance steering, *IEEE Robotics and Automation Letters*, 2019, **4**(3): 2276–2281.

[30] Okamoto K, Goldshtein M, and Tsiotras P, Optimal covariance control for stochastic systems under chance constraints, *IEEE Control Systems Letters*, 2018, **2**(2): 266–271.

[31] Liu C C, Gray A, Lee C, et al., Nonlinear stochastic predictive control with unscented transformation for semi-autonomous vehicles, 2014 *American Control Conference, IEEE*, 2014, 5574–5579.

[32] Mesbah A, Streif S, Findeisen R, et al., Stochastic nonlinear model predictive control with probabilistic constraints, *Proceedings of the American Control Conference*, 2014, 2413–2419.

[33] Okamoto K and Tsiotras P, Stochastic model predictive control for constrained linear systems using optimal covariance steering, 2019, http://arxiv.org/abs/1905.13296.

[34] Sehr M A and Bitmead R R, Particle model predictive control: Tractable stochastic nonlinear output-feedback MPC, *IFAC-PapersOnLine*, 2017, **50**(1): 15361–15366.

[35] Caluya K F and Halder A, Finite horizon density control for static state feedback linearizable systems, 2019, arXiv: 1904.02272.

[36] Ridderhof J, Okamoto K, and Tsiotras P, Nonlinear uncertainty control with iterative covariance steering, 2019 *IEEE 58th Conference on Decision and Control (CDC)*, 2019, 3484–3490.

[37] Yi Z J, Cao Z F, Theodorou E, et al., Nonlinear covariance control via differential dynamic programming, 2020 *American Control Conference (ACC), IEEE,* 2020, 3571–3576.

[38] Caluya K F and Halder A, Wasserstein proximal algorithms for the Schrödinger bridge problem: Density control with nonlinear drift, *IEEE Transactions on Automatic Control*, 2021, **67**(3): 1163–1178.

[39] Nodozi I and Halder A, Schrödinger meets Kuramoto via Feynman-Kac: Minimum effort distribution steering for noisy nonuniform kuramoto oscillators, 2022 *IEEE 61st Conference on Decision and Control (CDC)*, 2022, 2953–2960.

[40] Pakniyat A, A convex duality approach for assigning probability distributions to the state of nonlinear stochastic systems, *IEEE Control Systems Letters*, 2022, **6**: 3080–3085.

[41] Vinter R B and Lewis R M, The equivalence of strong and weak formulations for certain problems in optimal control, *SIAM Journal on Control and Optimization*, 1978, **16**(4): 546–570.

[42] Vinter R B and Lewis R M, A necessary and sufficient condition for optimality of dynamic programming type, making no a priori assumptions on the controls, *SIAM Journal on Control and Optimization*, 1978, **16**(4): 571–583.

[43] Vermes D, Optimal control of piecewise deterministic Markov processes, *Stochastics: An International Journal of Probability and Stochastic Processes*, 1985, **14**(3): 165–207.

[44] Fleming W H and Vermes D, Convex duality approach to the optimal control of diffusions, *SIAM Journal on Control and Optimization*, 1989, **27**(5): 1136–1155.

[45] Bhatt A G and Borkar V S, Occupation measures for controlled Markov processes: Characterization and optimality, *The Annals of Probability*, 1996, 1531–1562.

[46] Cho M and Stockbridge R, Linear programming formulation for optimal stopping problems, *SIAM Journal on Control and Optimization*, 2002, **40**(6): 1965–1982.

[47] Kurtz T and Stockbridge R, Existence of Markov controls and characterization of optimal Markov controls, *SIAM Journal on Control and Optimization*, 1998, **36**(2): 609–653.

[48] Zhu H, Convex duality for finite-fuel problems in singular stochastic control, *Journal of Optimization Theory and Applications*, 1992, **75**(1): 155–181.

[49] Chen Y X and Ames A D, Duality between density function and value function with applications in constrained optimal control and Markov decision process, 2019, arXiv: 1902.09583.

[50] Lasserre J B, Henrion D, Prieur C, et al., Nonlinear optimal control via occupation measures and LMI-relaxations, *SIAM Journal on Control and Optimization*, 2008, **47**(4): 1643–1666.

[51] Pauwels E, Henrion D, and Lasserre J B, Linear conic optimization for inverse optimal control, *SIAM Journal on Control and Optimization*, 2016, **54**(3): 1798–1825.

[52] Majumdar A, Vasudevan R, Tobenkin M M, et al., Convex optimization of nonlinear feedback controllers via occupation measures, *The International Journal of Robotics Research*, 2014, **33**(9): 1209–1230.

[53] Vasudevan R, Gonzalez H, Bajcsy R, et al., Consistent approximations for the optimal control of constrained switched systems — Part 1: A conceptual algorithm, *SIAM Journal on Control and Optimization*, 2013, **51**(6): 4463–4483.

[54] Zhao P C, Mohan S, and Vasudevan R, Optimal control of polynomial Hybrid systems via convex relaxations, *IEEE Transactions on Automatic Control*, 2019, **65**(5): 2062–2077.

[55] Pakniyat A, Theoretical guarantees for satisfaction of terminal state constraints for nonlinear stochastic systems, *IUTAM Symposium on Optimal Guidance and Control for Autonomous Systems*, Springer, 2023, 135–163.

[56] Gao S, Caines P E, and Huang M Y, LQG graphon mean field games: Analysis via graphon-invariant subspaces, *IEEE Transactions on Automatic Control*, 2023, **68**(12): 7482–7497.

[57] Pakniyat A and Caines P E, On the stochastic minimum principle for hybrid systems, *Proceedings of the* 55*th IEEE Conference on Decision and Control*, Las Vegas, NV, USA, 2016, 1139–1144.