Deep Control with Risk-Sensitive Linear Quadratic Models: A Gauge Transformation

Jalal Arabneydi, Member, IEEE, and Amir G. Aghdam, Senior Member, IEEE

Abstract—Motivated from invariance principle and recent developments in artificial intelligence, we introduce a risk-sensitive linear quadratic control system whose solution resembles a deep neural network; hence, it is called deep control. In particular, we consider a population of decision makers that are partitioned into multiple sub-populations wherein the decision makers in each sub-population are coupled in both dynamics and cost function through a number of linear regressions of the states and actions of the decision makers. Two non-classical information structures are considered: deep-state sharing and partial deep-state sharing, where deep state refers to the linear regression of the states of the decision makers in each sub-population. For deep-state sharing structure, a closed-form representation of the globally optimal strategy is obtained in terms of a deep Riccati equation, whose dimension is independent of the number of decision makers in each sub-population, i.e., the optimal solution is scalable. In addition, two sub-optimal sequential strategies under partial deep-state sharing information structure are proposed by introducing two Kalman-like filters, one based on the finite-population model and the other one based on the infinite-population model. It is shown that the prices of information associated with the above sub-optimal solutions converge to zero as the number of decision makers goes to infinity. Two numerical examples of a supply-chain management system are presented to demonstrate the efficacy of the obtained results.

I. INTRODUCTION

Recently, there has been a surge of interest in complex networked systems such as internet of things, social networks and smart grids; to name only a few. Such systems often consist of a large number of interconnected decision makers wherein a single decision maker (microscopic entity) has the potential to alter the behavior of the entire (macroscopic) system, a phenomenon known as butterfly effect. To avoid chaotic situations in such applications, a social welfare cost function is often defined as a common cost for all decision makers in order to enforce the desired behavior of the system. When it comes to practice, it is also important to take into account practical limitations such as the privacy of decision makers, limited computational power of processors and restricted capacity of communications. Under these limitations, the above optimization problem becomes a gigantic non-convex Bayesian optimization problem, where its exploration space grows exponentially with the control horizon and number of decision makers, which is intractable. As a result, it is important to find a systematic approach to upgrade the classical results in such a way that they can be efficiently applied to modern control systems containing many interconnected decision makers.

Inspired by new architectural developments in deep neural networks that have shown a remarkable performance in analyzing big data [1], [2], we strive to provide an analogous framework in optimal control theory in order to solve problems with a large number of decision makers. Recently, the authors have introduced deep team (alternatively called deep control) in [3] to study a class of big Markov decision processes with discrete state and action spaces, which shares some resemblances with the convolutional neural networks. In this paper, we propose a deep control system with continuous state and action spaces, whose solution is similar to a deep neural network with multiple layers of weighted sums and products. In particular, we consider a system consisting of multiple sub-populations of decision makers that are coupled in dynamics and cost through the weighted average of the states and actions of decision makers. Two non-classical information structures are considered: deep-state sharing and partial deep-state sharing, where deep state refers to the weighted average of states. In the former information structure, every decision maker observes its local state and the joint deep state while in the latter structure, the deep states of a subset (possibly all) of sub-populations are not observed. By using a gauge transformation and proposing a carefully constructed ansatz, we find a low-dimensional solution for the Hamilton-Jacobi-Bellman (HJB) equation in terms of a deep Riccati equation (that consists of several local Riccati equations and one global Riccati equation). The salient feature of the deep Riccati equation is the fact that its dimension is independent of the number of decision makers in each sub-population; hence, the optimal solution is scalable. It is to be noted that the solution itself may depend on the number of decision makers whereas the complexity of finding the solution does not. In addition, we propose two Kalman-like filters to compute two sub-optimal solutions, one based on the finite-population model and one based on the infinite-population one. Furthermore, we show that the prices of information and robustness converge to zero, as the number of decision makers tends to infinity. The main results are also extended to infinite-horizon cost function.

This paper is a generalized version of the previously published results for the risk-neutral model in [4]–[8]. In contrast to the risk-neutral case wherein certainty equivalence principle
The remainder of the paper is organized as follows. In Section II, the problem of deep linear quadratic control is formulated. In Section III, optimal solution is computed for deep-state sharing information structure and two sub-optimal solutions are proposed for partial deep-state sharing structure, whose performances converge to the optimal one as the number of decision makers goes to infinity. In Section IV, the proof is present in Appendix A. Preliminaries on invariance principle from invariant theory in mathematics and invariance mechanics in physics, that play a key role in describing natural processes, we study an equivariant linear quadratic system, where features have linear dynamics and quadratic cost functions.

Consider a simple linear quadratic system consisting of agents (decision makers) with the following dynamics at time 

\[ \dot{x} = A_t x_t + B_t u_t, \]

where \( x_t = \text{vec}(x_1^T, \ldots, x_n^T) \in \mathbb{R}^n \) is the joint state and \( u_t = \text{vec}(u_1^T, \ldots, u_n^T) \in \mathbb{R}^n \) is the joint action. The cost is defined as 

\[ \|x_T\|_{Q_T} + \int_0^T (\|x_t\|_{Q_t} + \|u_t\|_{R_t}) dt, \]

where \( T \in (0, \infty) \). Denote by \( P = P(1) \circ P(2) \circ \cdots \circ P(Y) \in \mathbb{R}^{n \times n} \) the composition (concatenation) of different transformations (filters) \( P(y) \), \( y \in \mathbb{N}_Y \). Suppose \( P \) is a real-valued non-singular square matrix, where \( v_m \) and \( \lambda_m \), \( m \in \mathbb{N}_n \), denote the \( m \)-th eigenvector and \( m \)-th eigenvalue of \( P \), respectively.

**Definition 1 (Equivariant linear quadratic system).** A linear quadratic system is said to be equivariant to transformation \( P \) if the following conditions hold at any time \( t \in [0, \infty) \):

- **Equivariant dynamics:** \( P \dot{x} = A_t P x_t + B_t P u_t \),
- **Equivariant cost:** instantaneous cost of the transformed system is proportional to that of the original system along each eigenvector such that

\[ \|P x_t\|_{Q_t} + \|P u_t\|_{R_t} = \sum_{m=1}^{\infty} c_m (\|\text{vec}(x_t,v_m)\|_{Q_t} + \|\text{vec}(u_t,v_m)\|_{R_t}), \]

where \( c_m \) is a positive constant, i.e., \( c_m = \langle \lambda_m, \lambda_m \rangle \).

In what follows, we propose a class of equivariant linear quadratic system for symmetric transformations.

**Proposition 1.** For any \( a_t, b_t, q_t, r_t, \bar{a}_t^m, \bar{b}_t^m, \bar{q}_t^m, \bar{r}_t^m \in \mathbb{R} \), \( m \in \mathbb{N}_n \), \( t \in [0, T] \), the following linear quadratic system is equivariant to the real-valued symmetric transformation \( P \),

\[ \ddot{x}^m_t = a_t x_t + b_t u_t + \sum_{m=1}^{\infty} \alpha_t^i m \bar{x}^m_t \dot{x}_t + \sum_{m=1}^{\infty} \alpha_t^i m \bar{u}^m_t \dot{u}_t, \]

with the cost function:

\[ \sum_{m=1}^{\infty} (\|\bar{x}^m_t\|_{q_t} + \int_0^T (\|\bar{x}^m_t\|_{n q_t} + \|\bar{u}_t\|_{n r_t}) dt) + \sum_{i=1}^{\infty} (\|\bar{x}^i_t\|_{q_i} + \int_0^T (\|\bar{x}^i_t\|_{n q_i} + \|\bar{u}_t\|_{n r_i}) dt), \]

where \( \alpha_t^i m \in \mathbb{R} \) is the \( i \)-th element of vector \( \sqrt{n} v^m \), and \( \bar{x}^m_t := \frac{1}{n} \sum_{i=1}^{\infty} \alpha_t^i m x^i_t \) and \( \bar{a}^m_t := \frac{1}{n} \sum_{i=1}^{\infty} \alpha_t^i m a^i_t \).

**Proof.** The proof is present in Appendix A. □

According to Proposition 1, each eigenvector represents a specific feature of transformation \( P \), where, for example, \( \bar{x}^m_t \) denotes the aggregate state of agents associated with the feature identified by the eigenvector \( v_m \), \( m \in \mathbb{N}_n \). In real-world applications, it is often practical to restrict attention to only a few number of dominant features corresponding to the largest eigenvalues. For the case of arbitrary permutation matrix \( P \) (which is generally not a symmetric matrix, i.e., it admits complex eigenvalues), we show that all features simplify the analysis, in the risk-sensitive case the principle does not hold [9]. To this end, we develop a more general approach in this paper to take the uncertainty into account.
become equally important, which results in an aggregate feature represented by the empirical (unweighted) average.

**Proposition 2.** For any $a_i, b_i, q_i, r_i, \bar{a}_i, \bar{b}_i, \bar{q}_i, \bar{r}_i \in \mathbb{R}$, $t \in [0,T]$, the following linear quadratic system is equivalent to every permutation $P$:

$$
\ddot{x}_i = a_i x_i + b_i u_i + \bar{a}_i \bar{x}_i + \bar{b}_i \bar{u}_i,
$$

with the cost function:

$$
\mathbb{E} \left[ \int_0^T \left( \sum_{i=1}^n \langle x_i^T, \dot{x}_i \rangle + \sum_{i=1}^n \langle \bar{x}_i, \dot{\bar{x}}_i \rangle \right) dt \right] + \frac{1}{n} \sum_{i=1}^n \langle x_i^T, \bar{x}_i \rangle + \frac{1}{n} \sum_{i=1}^n \langle \bar{x}_i^T, x_i \rangle dt,
$$

where $x_i := \frac{1}{n} \sum_{i=1}^n x_i^T$ and $\bar{u}_i := \frac{1}{n} \sum_{i=1}^n u_i^T$.

**Proof.** The proof is presented in Appendix B.

**Definition 2 (Partially equivariant system).** A system is said to be partially equivariant if it can be partitioned into $K$ distinct sub-populations, where dynamics and cost of agents in each sub-population are coupled through the features of all sub-populations. For the special case of $K = 1$, partially equivariant system reduces to an equivalent one.

From Propositions 1 and 2, we can deduce that agents in equivariant linear quadratic systems are coupled through a number of linear regressions of the states and actions of agents. In what follows, we consider a partially equivariant system consisting of multiple sub-populations with multivariate parameters. For ease of display, we present our main results for the case in which there is only one weighted average (linear regression) per sub-population, because adding orthonormal linear regressions in each sub-population does not add much complexity to our analysis, as demonstrated in Subsection IV-B.

**B. Model**

Consider a stochastic dynamic control system consisting of $K \in \mathbb{N}$ sub-populations wherein each sub-population $k \in \mathbb{N}_K$ contains $n_k \in \mathbb{N}$ agents. Let $x_i^k \in \mathbb{R}^{e_k}$ and $u_i^k \in \mathbb{R}^{d_k}$, respectively, denote the state and action of agent $i$ in sub-population $k \in \mathbb{N}_K$, where $e_i, d_i \in \mathbb{N}$. Let $\alpha^{i,k} \in \mathbb{R}$ be the impact factor of agent $i$ among its peers in sub-population $k$ so that the aggregate state and aggregate action of sub-population $k$ are described by:

$$
\bar{x}_i^k := \frac{1}{n_k} \sum_{i=1}^{n_k} \alpha^{i,k} x_i^k, \quad \bar{u}_i^k := \frac{1}{n_k} \sum_{i=1}^{n_k} \alpha^{i,k} u_i^k.
$$

Since weighted average is a linear regression, it can be used to approximate various types of interactions between agents, e.g., non-linear functions and complex networks. For ease of reference, we refer to $\bar{x}_i^k$ as the deep state of sub-population $k$ at time $t \in [0, \infty)$ in the sequel.

Define augmented vectors $\bar{x}_k := \text{vec}(\bar{x}_1^k, \ldots, \bar{x}_K^k)$ and $\bar{u}_k := \text{vec}(\bar{u}_1^k, \ldots, \bar{u}_K^k)$. The dynamics of agent $i$ in sub-population $k$ is affected by other agents through the deep states as follows:

$$
d(\bar{x}_i^k) = (A_i^k \bar{x}_i^k + B_i^k u_i^k + \alpha^{i,k} \bar{A}_i \bar{x}_i + \alpha^{i,k} \bar{B}_i \bar{u}_i) dt + C_i^k d\bar{w}_i^k,
$$

where $A_i^k, B_i^k, \bar{A}_i, \bar{B}_i$ and $C_i^k$ are time-varying matrices of appropriate dimensions, and $\{\bar{w}_i^k \in \mathbb{R}^{e_k}, t \in [0,\infty)\}$ is an $e_k$-dimensional standard Brownian motion. Without loss of generality, we consider the following normalization:

$$
\frac{1}{n_k} \sum_{i=1}^{n_k} (\alpha^{i,k})^2 = 1.
$$

Let $r_i^k \in \mathbb{R}^{e_k}$ be the desired operating point of agent $i$ in sub-population $k$ at time $t \in [0, \infty)$. It is assumed that the first derivative of the signal $r_i^k$ is a continuous function in time. The cost of agent $i$ in sub-population $k \in \mathbb{N}_K$ is defined by:

$$
J^k_i = \mathbb{E} \left[ \int_0^T \left( \langle \bar{x}_i^k - r_i^k, \bar{x}_i^k \rangle + \langle \bar{u}_i^k, \bar{u}_i^k \rangle \right) dt \right]
$$

where symmetric matrices $Q_i^k, R_i^k, \bar{Q}_i^k$ and $\bar{R}_i^k$ have appropriate dimensions, $t \in [0,T]$, where $T$ is the limit of the finite horizon. In addition, let $\mu_k \in [0,\infty)$ denote the impact factor of sub-population $k \in \mathbb{N}_K$ among all sub-populations. As an example, the impact factor of sub-population $k$ may be defined based on its size in a population or its topological configuration in a network. In the sequel, we refer to $\alpha^{i,k}$ as the microscopic impact factor (that determines the influence of an individual agent $i$ in sub-population $k$) and to $\mu_k$ as the macroscopic impact factor (that indicates the influence of sub-population $k$ in the population).

For any risk parameter $\theta \in [0,\infty)$, define the following risk-sensitive cost function:

$$
\gamma_T(\theta) = \frac{1}{\theta} \log \mathbb{E}[\exp(\theta J_T)],
$$

where

$$
J_T = \sum_{k=1}^K \mu_k^\theta \frac{1}{n_k} \sum_{i=1}^{n_k} J^k_i.
$$

By twice using the Taylor series expansion and keeping the first two dominant terms, it is concluded that for small $\theta$, $\gamma_T(\theta) \approx \mathbb{E}[J_T] + \frac{\theta}{2} \operatorname{var}(J_T)$ [12, 13]. An immediate consequence of this approximation is that the risk-sensitive cost function converges to the risk-neutral one as $\theta$ goes to zero, i.e. $\lim_{\theta \to 0} \gamma_T(\theta) = \mathbb{E}[J_T]$. It is to be noted that the risk-sensitive strategy takes into account not only the performance but also the robustness.

Denote by $x_t$, $u_t$ and $w_t$, the augmented state, action and noise of all agents at time $t \in [0, T]$, respectively. Let $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$ be a filtered probability space, where $\mathcal{F}_t$ is an increasing sigma-algebra generated by random variables $\{x_0, w_0, t \in [0, T]\}$. In addition, $\{x_0, w_0, t \in [0, T]\}$ are assumed to be mutually independent across agents and time horizon. We denote $\mu^{x,k} = \mathbb{E}(x^k_0)$, $\sigma^{x,k} = \mathbb{E}((x^k_0 - \mu^{x,k})^2)$ and $\sigma^{w,k} = C^k(\sigma^k)^T$ for every agent $i$ in sub-population $k \in \mathbb{N}_K$ at time $t \in [0, \infty)$. To have a well-posed problem, we assume that $\mu^{x,k}, \sigma^{x,k}, \sigma^{w,k}$ and $r_i^k$.
\( t \in [0, T] \), are uniformly bounded, and the set of admissible control actions are adapted to the filtration \( \mathcal{F}_t \) and square integrable for all agents.

Two special classes of sub-populations are described below.

**Definition 3 (Exchangeable class).** A sub-population \( k \in \mathbb{N}_K \) is said to be exchangeable when agents are equally important, i.e., \( \alpha^{i,k} = 1 \) for all agents in the sub-population \( k \).

**Definition 4 (Soft-constraint class).** A sub-population \( k \in \mathbb{N}_K \) has soft-constraint formulation when matrices \( \bar{A}^k_i \) and \( \bar{B}^k_i \) in the dynamics are zero.

Define the following matrices at any time \( t \in [0, T] \):

\[
\sigma_t := \text{diag}(\mu^1/n^1, \ldots, \mu^K/n^K), \\
C_t := \text{diag}(C^1_t, \ldots, C^K_t), \\
A_t := \text{diag}(A^1_t, \ldots, A^K_t) + [\bar{A}^1_t, \ldots, \bar{A}^K_t]_\tau, \\
B_t := \text{diag}(B^1_t, \ldots, B^K_t) + [\bar{B}^1_t, \ldots, \bar{B}^K_t]_\tau, \\
Q_t := \text{diag}(Q^1_t, \ldots, Q^K_t) + \sum_{k=1}^K \mu^k \bar{Q}^k_t, \\
R_t := \text{diag}(R^1_t, \ldots, R^K_t) + \sum_{k=1}^K \mu^k \bar{R}^k_t.
\]

(6)

To derive our main results, we make the following standard assumptions on the model.

**Assumption 1.** For any \( t \in [0, T] \), matrices \( Q^k_t, R^k_t, k \in \mathbb{N}_K \), and \( Q_t \) are positive semi-definite and matrices \( R^k_t \), \( k \in \mathbb{N}_K \), and \( R_t \) are positive definite.

**Assumption 2.** For any \( t \in [0, T] \) and \( k \in \mathbb{N}_K \), matrices \( B^k_t(R^k_t)^{-1}B^k_t - 2\theta(\mu^k/n^k)\sigma^w_{t,k} \) and \( B_tR_t^{-1}B_t^\tau - 2\theta \sigma_t \) are positive definite.

**C. Admissible strategy**

In this paper, we consider two non-classical information structures. The first one is called deep-state sharing (DSS), where each agent \( i \) of sub-population \( k \in \mathbb{N}_K \) observes its local state \( x^i_t \) as well as the joint deep state \( \bar{x}_k \), i.e.,

\[
u^i_t = g^i_t(x^i_t, \bar{x}_t).
\]

(DSS)

where \( g^i_t \) is called the control law of agent \( i \) at time \( t \in [0, T] \). In practice, deep state can be shared among agents in various ways. For example, in a stock market, the total amount of shares and trades are often announced publicly to be accessible to buyers, seller and traders. Another example is a swarm of robots wherein the deep state can be computed by local communications among agents using consensus algorithms [14], [15], on noting that the control process is often slower than the communication process.

The second information structure is called partial deep-state sharing (PDSS), where each agent \( i \) observes its local state as well as a subset \( S \subseteq \mathbb{N}_K \) of the deep states, i.e.,

\[
u^i_t = g^i_t(x^i_t, (\bar{x}^k_t)_{k \in S}).
\]

(PDSS)

When a sub-population is large, collecting and sharing its deep state among agents may not be feasible. In this case, PDSS structure is desirable as it does not contain the deep state of large sub-populations. It is to be noted that deep-state sharing and no-sharing information structures are special cases of PDSS, where \( S = \mathbb{N}_K \) and \( S = \emptyset \), respectively. In addition, note that the privacy of each agent is respected in the above structures because its local state is not shared with others. The problems investigated in this paper are defined below.

**Problem 1.** For deep-state sharing information structure, find an optimal strategy that minimizes the cost function \( [4] \).

Let \( n^* \) be the smallest sub-population whose deep state is not shared, i.e., \( n^* := \min_{k \in S} n^k \).

**Problem 2.** For partial deep-state sharing information structure, find a sub-optimal strategy whose performance converges to the optimal performance under DSS structure, as \( n^* \rightarrow \infty \).

**D. Main challenges and contributions**

There are two main challenges. The first one is the **curse of dimensionality**, where the matrices in Problems [1] and [2] are fully dense, i.e., the corresponding (centralized) Riccati equation is intractable for a large number of decision makers. The second challenge is the **imperfection of the information structure** wherein dynamic programming decomposition is not applicable, because agents cannot find a low-dimensional sufficient statistic for the sequential decomposition. The main contributions of this article are outlined below.

1) We obtain a closed-form tractable representation of the optimal solution for any arbitrary number of agents under DSS structure, where the feedback gains are computed by a deep Riccati equation whose dimension is independent of the number of agents in each sub-population (Theorem [1]).

2) We propose two sub-optimal solutions under PDSS structure by introducing two Kalman-like filters, one based on the finite-population model and one based on the infinite-population model (Theorem [2] and Corollary [2]). In addition, we extend our main results to the cases of infinite-horizon cost function (Theorem [3]) and multiple impact factors in a sub-population (Subsection [IV-B]).

3) We formally define **LQ deep control** and establish a bridge between optimal control theory and deep neural networks (Section [VI]). For example, we show analytically that the structure of the proposed neural network is independent of the number of orthonormal features, substantiating the fact that deep neural networks are structurally rich enough to learn complex systems.

**III. MAIN RESULTS**

In invariant physics, a gauge transformation, upon existence, is a powerful tool for the analysis of invariant systems. In simple words, a gauge transformation manipulates the degrees of freedom of an invariant system without changing its structure [16]. In this paper, we use a gauge transformation introduced in [4] Appendix A.2 for optimal control problems and in [17] for dynamic games. Define the following variables
for any agent $i \in \mathbb{N}_{n,k}$ of sub-population $k \in \mathbb{N}_K$ at time $t \in [0,T]$:  
\begin{align}
\Delta x_i^t &= x_i^t - \alpha^{i,k} \bar{x}_k^t, \quad \Delta u_i^t := u_i^t - \alpha^{i,k} \bar{u}_k^t, \\
\Delta w_i^t := w_i^t - \alpha^{i,k} \bar{w}_k^t, \quad \Delta r_i^t := r_i^t - \alpha^{i,k} \bar{r}_k^t, 
\end{align}

(7)

where
\[
\bar{w}_i^t := \frac{1}{n^k} \sum_{i=1}^{n^k} \alpha^{i,k} w_i^t, \quad \bar{r}_i^t := \frac{1}{n^k} \sum_{i=1}^{n^k} \alpha^{i,k} r_i^t.
\]

From equations (1), (2), (6) and (7), the dynamics of the transformed variables is described by:
\begin{align*}
d(\Delta x_i^t) &= (A_k^t \Delta x_i^t + B_k^t \Delta u_i^t)dt + C_i^t d(\Delta w_i^t), \\
d(\bar{x}_k^t) &= (A_i \bar{x}_i + B_i \bar{u}_i)dt + C_i d(\bar{w}_i),
\end{align*}

where $\bar{w}_i := \text{vec}(\bar{w}_i^1, \ldots, \bar{w}_i^K)$. The main feature of the gauge transformation (7) is that it induces the following orthogonal relations at each sub-population $k \in \mathbb{N}_K$, i.e.,
\[
\begin{align*}
\sum_{i=1}^{n^k} \alpha^{i,k}(\Delta x_i^t - \Delta r_i^t)^T Q_k^t (\bar{x}_k^t - \bar{r}_k^t) &= 0, \\
\sum_{i=1}^{n^k} \alpha^{i,k}(\Delta u_i^t)^T R_k^t \bar{u}_k^t &= 0.
\end{align*}
\]

Let $r_i := \text{vec}(r_i^1, \ldots, r_i^K)$. By using the above relations, one can express the cost function (5) from time $t$ to $T$ as follows:
\begin{align}
J_{t:T} := L_T &+ \sum_{k=1}^{K} \frac{\mu_k}{n^k} \sum_{i=1}^{n^k} \Delta L_i^T \\
&+ \int_{t}^{T} \left( L_r + \sum_{k=1}^{K} \frac{\mu_k}{n^k} \sum_{i=1}^{n^k} \Delta L_i^r \right) dt, 
\end{align}

(8)

where
\[
\begin{align*}
\Delta L_i^T := ||\Delta x_i^T - \Delta r_i^T||_{Q_k^T}, \\
L_T := ||\bar{x}_T - F_T ||_{\text{diag}(\mu_1 Q_1^T, \ldots, \mu_K Q_K^T)} + ||\bar{x}_T||_{\sum_{k=1}^{K} \mu_k Q_k^T}, \\
\Delta L_i^r := ||\Delta x_i^T - \Delta r_i^T||_{Q_k^T} + ||\Delta u_i^T||_{R_k^T}, \\
L_r := ||\bar{x}_i - \bar{r}_i||_{\text{diag}(\mu_1 Q_1^T, \ldots, \mu_K Q_K^T)} + ||\bar{x}_i||_{\sum_{k=1}^{K} \mu_k Q_k^T} \\
&+ ||\bar{u}_i||_{R_i}.
\end{align*}
\]

At any time $t \in [0,T]$, let $y_t := \{(\Delta x_i^t)_{i=1}^{n^k})_{k=1}^{K}, \bar{x}_k^t\}$ and $v_t := \{(\Delta u_i^t)_{i=1}^{n^k})_{k=1}^{K}, \bar{u}_i^t\}$ denote the centralized state and action of the transformed system, respectively. Suppose that $y_t$ is known to all agents. It will be shown later that the optimal centralized solution can be implemented under DSS structure. Define a real-valued function $\psi_t$ at time $t \in [0,T]$ as follows:
\[
\psi_t(y) := \mathbb{E}[e^{\gamma T \Delta r_T} | y_t = y],
\]

(9)

where $\gamma_T(\theta) = \frac{1}{2} \mathbb{E}[-\log \psi_t(y_0)]$ according to (4), (8) and (9). Since $\frac{1}{2} \mathbb{E}[-\log (\cdot)]$ is a strictly increasing function, any strategy that minimizes $\psi$ will also minimize $\gamma$. From (9) the Hamilton-Jacobi-Bellman equation can be written as follows:
\[
- \frac{d}{dt} \psi_t(y_t) = \inf_{\nu_t} \left[ \theta \psi_t(y_t)(L_t + \sum_{k=1}^{K} \frac{\mu_k}{n^k} \sum_{i=1}^{n^k} \Delta L_i^r) + \sum_{k=1}^{K} \sum_{i=1}^{n^k} \nabla \Delta x_i^t \psi_t(y_t)^T (A_i^t \Delta x_i^t + B_i^t \Delta u_i^t) \\
+ \frac{\mu_k}{n^k} \sum_{i=1}^{n^k} \nabla \Delta x_i^t \psi_t(y_t) R_i^t (\bar{x}_i^t - \bar{r}_i^t) \\
+ \frac{1}{2} \text{Tr}(\sum_{k=1}^{K} \sum_{i=1}^{n^k} \nabla \Delta x_i^t \psi_t(y_t) \sigma_{t,w,k} (\frac{d}{dt}(\mathbb{E}[(\Delta u_i^t)^T \Delta u_i^t])) \\
+ 2 \sum_{k=1}^{K} \sum_{i=1}^{n^k} \nabla \Delta x_i^t \psi_t(y_t) \sigma_{t,w,k} (\frac{d}{dt}(\mathbb{E}[(\Delta u_i^t)^T \bar{u}_i^T]))) \\
+ \frac{1}{2} \text{Tr}(\sum_{k=1}^{K} \sum_{i=1}^{n^k} \nabla \Delta x_i^t \psi_t(y_t) \sigma_{t,w,k} (\frac{d}{dt}(\mathbb{E}[(\bar{u}_i^T)^T \bar{u}_i^T]))) \right],
\]

(10)

where the cross terms associated with $\mathbb{E}[(\bar{u}_i^T)^T \bar{u}_i^T] = 0, k \neq k'$ do not appear in the above equation.

For every sub-population $k \in \mathbb{N}_K$, define a local scale-free HJB equation similar to a linear exponential quadratic problem with the dynamics characterized by matrices $A_i(k)$ and $B_i(k)$, a zero-mean noise with the covariance matrix $(\mu_k/n^k) \sigma_{w,k}$, and a tracking signal $\Delta r_T$, whose solution is given by:
\[
\begin{align*}
- \tilde{P}_k^T &+ \tilde{P}_k^t A_k^T + A_k^T \tilde{P}_k^t - \tilde{P}_k^t (B_k^T (R_k^t)^{-1} B_k^k + 20 (\mu_k/n^k) \sigma_{w,k}) P_k^T + s_k^T \\
- \tilde{v}_k &+ (\mu_k/n^k) \theta_k^T \text{Tr}(P_k^t \sigma_{w,k}^T), \\
- \tilde{d}_k^T &+ (A_k^T - B_k R_k^{-1} B_k^T)^T P_k^T \\
+ 20 (\mu_k/n^k) \sigma_{w,k}^T)^T \tilde{d}_k^T &- Q_k^T (r_i^t - \alpha^{i,k} r_i^k), \\
\tilde{J}_0^T &- \tilde{J}_0^T = s_k^T \theta (\Delta r_0^T r_0^T + f_0^T (\Delta r_0^T)^C dt) \\
&+ (\theta^T)^C \tilde{d}_k^T, \\
\Delta u_k^T &- (R_k^t)^{-1} B_k^T P_k^T \Delta x_i^t - (R_k^t)^{-1} (B_k^T)^T \phi_i^T.
\end{align*}
\]

(11)

with boundary conditions $P_k^T = Q_k^T, s_k^T = 1$ and $\tilde{d}_k^T = Q_k^T \Delta r_T$. In addition, define a global HJB equation similar to a linear exponential quadratic problem with the dynamics characterized by matrices $A_i$ and $B_i$, a zero-mean noise with the covariance matrix $\sigma_T$, and a tracking signal $\bar{r}_i$, whose solution is expressed by:
\[
\begin{align*}
- \tilde{P}_k &+ P_k A_i + A_i^T P_k - P_k (B_i R_i^{-1} B_i^T + 2 \theta \sigma_T) P_k, \\
- \tilde{s}_i &+ \text{Tr}(P_i \sigma_i), \\
- \tilde{v}_i &+ (A_i - B_i R_i^{-1} B_i^T)^T P_i, \\
- \text{diag}(\mu_1 Q_1^T, \ldots, \mu_K Q_K^T) \tilde{r}_i, \\
\psi_0(x_0) &- \text{diag}(\mu_1 Q_1^T, \ldots, \mu_K Q_K^T) \tilde{r}_i, \\
\psi_0(x_T) &- \text{diag}(\mu_1 Q_1^T, \ldots, \mu_K Q_K^T) \tilde{r}_i, \\
\Delta u_i^T &- (R_i^T)^{-1} B_i^T P_i \Delta x_i^t - (R_i^T)^{-1} (B_i^T)^T \phi_i.
\end{align*}
\]

(12)

with boundary conditions $P_T = Q_T, s_T = 1$ and $\tilde{d}_T = \text{diag}(\mu_1 Q_1^T, \ldots, \mu_K Q_K^T) F_T$. 

Theorem 1. Let Assumptions 1 and 2 hold. The optimal solution of Problem 7 is described as follows. For every agent $i \in \mathbb{N}_K$ of sub-population $k \in \mathbb{N}_K$ at any time $t \in [0, T]$:

$$u_{i,k}^* = F_t^{\theta,k}(x_t - \alpha^{i,k}x_t) + \alpha^{i,k}F_t^{\theta,k}x_t + q_{t}^{\theta,i,k} + \alpha^{i,k}\dot{q}_{t}^{\theta,k},$$

(13)

where the gains $F_t^{\theta,k}$, $k \in \mathbb{N}_K$, and the correction terms $q_{t}^{\theta,i,k}$, $i \in \mathbb{N}_i$, are computed from the solution of (11), i.e.,

$$F_t^{\theta,k} := -(P_t^{\theta,k})^{-1}(B_t^{\theta,k})^T P_t^{\theta,k} q_{t}^{\theta,i,k} := -(P_t^{\theta,k})^{-1}(B_t^{\theta,k})^T \phi_i,$n

and similarly, $F_t^{\theta,k}$ and $q_{t}^{\theta,i,k}$ are obtained from the solution of (12), i.e.,

$$F_t^{\theta,k} : = \begin{bmatrix} F_t^{\theta,1} \\ \vdots \\ F_t^{\theta,K} \end{bmatrix} = \begin{bmatrix} -R_{t}^{-1}B_{t}\phi_i \\ \vdots \\ -R_{t}^{-1}B_{t}\phi_i \end{bmatrix}: = -R_{t}^{-1}B_{t}\phi_i.$$

Proof. The proof follows in three steps. In the first step, we define an ansatz for the HJB equation (10) as follows:

$$\psi_i(y_t) = \sum_{k=1}^{K} \eta_k \iota_{k}^{\theta,\phi}(\|x_t\|_1 + h_t),$$

(14)

where $\eta_k$, $\iota_k$, and $h_t$ are scalars, and $P_t^{\theta,k}$ and $P_t$ are symmetric matrices with appropriate dimensions, $k \in \mathbb{N}_K$.

In the second step, we establish two fundamental properties of the above ansatz which lead to finding a low-dimensional solution. In particular, the following relationship holds for any sub-population $k \in \mathbb{N}_K$ at any time $t \in [0, T]$:

$$\text{Tr} \left( \sum_{i=1}^{n_k} \nabla \Delta x_t \psi_i(y_t) \sigma_t^{w,k} \left( \frac{d}{dt} \mathbb{E}[(\Delta u_t^{i})^T \Delta u_t] \right) \right) = \text{Tr} \left( \sum_{i=1}^{n_k} (P_t^{\theta,k})^T \Delta x_t \psi_i(y_t) \sigma_t^{w,k} \right) \sum_{i=1}^{n_k} \Delta x_t \psi_i(y_t) \sigma_t^{w,k} \left( \frac{d}{dt} \mathbb{E}[(\Delta u_t^{i})^T \Delta u_t] \right)$$

$$= \text{Tr} \left( (P_t^{\theta,k})^T P_t^{\theta,k} \psi_i(y_t) \sigma_t^{w,k} \right) \sum_{i=1}^{n_k} \Delta x_t \psi_i(y_t) \sigma_t^{w,k} \left( \frac{d}{dt} \mathbb{E}[(\Delta u_t^{i})^T \Delta u_t] \right)$$

(15)

where (a) follows from the linear dependence introduced by the gauge transformation $\sum_{j=1}^{n} \delta^{i,k} \Delta x_j = 0$; and the fact that $\mathbb{E}[(\Delta u_t^{i})^T \Delta u_t] = (1 - \alpha^{i,k}_t)^2 t$ and $\mathbb{E}[(\Delta u_t^{i})^T \Delta u_t] = -\alpha^{i,k}_t \alpha^{i,k}_t t$, and (b) follows from equation (14). In addition, for every $k \in \mathbb{N}_K$ and $t \in [0, T]$, one has:

$$\text{Tr} \left( \sum_{i=1}^{n_k} \nabla \Delta x_t \psi_i(y_t) \sigma_t^{w,k} \left( \frac{d}{dt} \mathbb{E}[(\Delta u_t^{i})^T \Delta u_t] \right) \right) = 0,$$

(15)

because $\mathbb{E}[(\Delta u_t^{i})^T \Delta u_t] = 0$. According to P1 and P2, equation (10) can be expressed by HJBs given by (11) and (12).

In the third step, we translate the solution back to the original variables, i.e., the optimal solution of agent $i$ in sub-population $k \in \mathbb{N}_k$ at time $t \in [0, T]$ can be expressed as:

$$u_{i,k}^* = F_t^{\theta,k}x_t + q_{t}^{\theta,i,k} + \alpha^{i,k}F_t^{\theta,k}x_t + \alpha^{i,k}q_{t}^{\theta,i,k}.$$

The proof is complete.

Definition 5 (Deep Riccati equation). We refer to the stacked version of Riccati equations in (11) and (12) as one deep Riccati equation that identifies a scalable solution for the big optimization problem in (10), as described in Theorem 1.

It is to be noted that the dimension of the matrices in deep Riccati equation is independent of the number of agents $n_k$, $\forall k \in \mathbb{N}_K$. An interesting feature of the deep Riccati equation is that it decomposes into $K + 1$ decoupled Riccati equations for control (team) problems whereas such decomposition does not generally hold for game problems.

Prior to the operation of the system, every agent $i$ of sub-population $k \in \mathbb{N}_K$ can independently solve two Riccati equations: a local Riccati equation in (11) and a global Riccati equation in (12). During the control process, every agent $i$ of sub-population $k \in \mathbb{N}_K$ coordinates itself within its sub-population at any time $t \in [0, T]$ based on several factors: (a) the solution of local Riccati equation (11), (b) private information $\{x_t^{i,k}, r_t^{i,k}, \alpha^{i,k}\}$, and (c) public information $\{x_t, r_t, \theta, \sigma_t^{w,k}, \mu_k, n_k\}$. Simultaneously, the agent $i$ of sub-population $k \in \mathbb{N}_K$ coordinates itself within the population based on global factors: (d) the solution of global Riccati equation (12) and (e) public information $\{x_t, r_t, \theta, \{\mu_k, \sigma_t^{w,k}, n_k\}_{k=1}^{K}\}$. Note that the only piece of information that needs to be shared at any time $t$ is $x_t$.

Remark 1 (Risk-neutral model). When the risk factor in Theorem 1 is set to zero (i.e., $\theta = 0$), the solution of Theorem 1 reduces to the solution of risk-neutral problem [4, 8]. In contrast to the risk-neutral model wherein the Riccati equations are independent of the number of agents and probability distribution of driving noises, Riccati equations in the risk-sensitive model depend on the above parameters.

To further emphasize the complexity of the risk-sensitive case compared to the risk-neutral one, consider a case wherein the local noises are correlated. In this situation, properties P1 and P2 in the proof of Theorem 1 do not necessarily hold; hence, the decomposition proposed in Theorem 1 will not hold either. In contrary, this difficulty does not arise in the risk-neutral problem because of the celebrated certainty equivalence principle.
Remark 2 (Common noise). Suppose that local noises $(w_i^k)_{i=1}^{n_k}$, $k \in \mathbb{N}_K$, are correlated through an additive common noise $\tilde{w}_k^k$ such that $w_i^k := \tilde{w}_k^k + \tilde{w}_i^k$, where $(\tilde{w}_i^k)_{i=1}^{n_k}$ and $\tilde{w}_k^k$ are mutually independent. In this case, relations P.1 and P.2 still hold, i.e., the decomposition in Theorem 1 is valid.

Definition 6 (Price of Robustness). The price of robustness (PoR) is defined as a factor to quantify the loss of performance by taking the robustness into account, i.e.,

$$
PoR(\theta) := \gamma_T(\theta) - \gamma_T(0).
$$

We impose the following assumption to ensure that the state dynamics and cost function remain bounded as the number of agents goes to infinity.

Assumption 3. All matrices defined in the dynamics and cost functions (2) and (3) are uniformly bounded with respect to $n^k$, $\forall k \in \mathbb{N}_K$.

For uniformly bounded $\mu^k$, $\forall k \in \mathbb{N}_K$, local and global Riccati equations in (11) and (12) reduce to their counterparts in the risk-neutral model, as the number of agents goes to infinity. This implies that there is no loss of optimality in restricting attention to risk-neutral models when the number of agents is very large. It is to be noted that this relationship does not hold for game problems.

Corollary 1. Let Assumptions 2 and 3 hold. From Theorem 2 and equations (11) and (12), it follows that the price of robustness converges to zero as $n^k \to \infty$, $\forall k \in \mathbb{N}_K$.

In the next two subsections, we propose two sub-optimal solutions for the PDSS structure. To quantify the performance of each sub-optimal solution, we define a measure, called price of information, inspired by a similar notion in [19].

Definition 7 (Price of Information). The price of information (PoI) of a PDSS strategy $\tilde{g}$ is defined as the optimality gap between strategy $\tilde{g}$ and optimal DSS strategy $g^*$, i.e.,

$$
PoI_T(\theta) := \gamma_{\tilde{g}}^T(\theta) - \gamma_{g^*}^T(\theta).
$$

A. Finite-model strategy for Problem 2

To distinguish from the optimal solution, let $\tilde{x}_i^k$ and $\tilde{u}_i^k$ denote the state and action of agent $i$ under the sub-optimal strategy $\tilde{g}$ at time $t$. For any sub-population $k \in \mathbb{N}_K$ at any time $t \in [0, T]$, define:

$$
\tilde{x}_i^k := \frac{1}{n_k^k} \sum_{i=1}^{n_k^k} a^{i,k} \tilde{x}_i^k, \quad \tilde{u}_i^k := \frac{1}{n_k^k} \sum_{i=1}^{n_k^k} a^{i,k} \tilde{u}_i^k.
$$

Let $z_i := \text{vec}(z_i^1, \ldots, z_i^{K})$ be an estimate for $x_i$, where for any sub-population $k \in S$, $z_i^k := \tilde{x}_i^k$ and for any $k \notin S$, $z_i^k := m_i^k \cdot \tilde{x}_i^k$. Denote

$$
A_i^{\theta,k} := B_i^k F_i^{\theta,k} + A_i^k + B_i^k F_i^\theta,
$$

$$
\eta_i^{\theta,k} := B_i^k q_i^{\theta,k} + B_i^k R_i^{-1} B_i^T_i \phi_i.
$$

Define a Kalman-like filter as follows:

$$
\dot{\tilde{z}}_i^k = A_i^{\theta,k} \tilde{z}_i^k + A_i^{\theta,k} z_i^k + \eta_i^{\theta,k} + H_k(\tilde{x}_i^k - A_i^{\theta,k} \tilde{z}_i^k - \eta_i^{\theta,k}),
$$

where the observer gain $H_k^k = 1$ if $k \in S$, and $H_k^k = 0$ if $k \notin S$. Notice that $z_i$ is not necessarily the best possible estimate, but it is always measurable with respect to PDSS information structure. The interested reader is referred to [20] for more details on the above Kalman-like filter that emerges in the trade-off between data collection and data estimation.

We now propose a PDSS strategy wherein the action of any agent $i$ in any sub-population $k \in \mathbb{N}_K$ at any time $t \in [0, T]$ is given by:

$$
\tilde{u}_i^k = F_i^{\theta,k}(\tilde{x}_i^k - \alpha^{i,k} z_i^k) + \alpha^{i,k} F_i^{\theta,k} z_i^k + q_i^{\theta,i,k} + \alpha^{i,k} q_i^{\theta,k}.
$$

(16)

Theorem 2. Let Assumptions 2 and 3 hold. Then, the PDSS strategy (16) is a solution for Problem 2.

Proof. Define a relative distance between the deep state under optimal strategy (13), i.e., $\tilde{x}_i^k$, and its estimate under sub-optimal strategy (16), i.e., $z_i^k$, for any sub-population $k \in \mathbb{N}_K$ at time $t \in [0, T]$, i.e.,

$$
e_i^k := \tilde{x}_i^k - z_i^k.
$$

Similarly, define a relative distance between the deep state under the sub-optimal strategy, i.e., $\tilde{x}_i^k$, and the estimate $\tilde{z}_i^k$, for any sub-population $k$ at any time $t$ as follows:

$$
\xi_i^k := \tilde{x}_i^k - \tilde{z}_i^k.
$$

(17)

(18)

Denote $e_i := \text{vec}(e_i^1, \ldots, e_i^K)$ and $\xi_i := \text{vec}(\xi_i^1, \ldots, \xi_i^K)$. From equations (1), (2), (6), (13), (15), (16), (17) and (18), it follows that for any $k \notin S$,

$$
\begin{cases}
\dot{d}(e_i) = (A_i^k e_i^k + A_i^{\theta,k} e_i) dt + C_i^k d(\tilde{w}_i^k), \\
\dot{d}(\xi_i) = (A_i^k + B_i^k F_i^{\theta,k}) \xi_i^k \\
\quad + (A_i^k + B_i^k \text{diag}(F_i^{\theta,1}, \ldots, F_i^{\theta,K})) \xi_i dt + C_i^k d(\tilde{w}_i^k).
\end{cases}
$$

(19)

and for any $k \in S$,

$$
\begin{cases}
\dot{e}_i^k = A_i^k e_i^k + A_i^{\theta,k} e_i - (A_i^k + B_i^k F_i^{\theta,k}) \xi_i^k \\
\quad - (A_i^k + B_i^k \text{diag}(F_i^{\theta,1}, \ldots, F_i^{\theta,K})) \xi_i dt + (1 - H_k^k) C_i e_i d(\tilde{w}_i^k).
\end{cases}
$$

Let $H^S := \text{vec}(H^1, \ldots, H^K)$, where $H^S$ depends on the set of sub-populations whose deep states are observed. The above dynamics can be expressed in a more compact form as follows:

$$
\begin{align}
\dot{d}(e_i) &= (A_i + B_i^k F_i^\theta) e_i dt + (1 - H^S) C_i d(\tilde{w}_i) \\
&- H^S (A_i + B_i^k \text{diag}(F_i^{\theta,1}, \ldots, F_i^{\theta,K})) \xi_i dt, \\
\dot{d}(\xi_i) &= (1 - H^S) (A_i + B_i^k \text{diag}(F_i^{\theta,1}, \ldots, F_i^{\theta,K})) \xi_i dt \\
&+ (1 - H_k^k) C_i e_i d(\tilde{w}_i^k).
\end{align}
$$

(19)

From the strong law of large numbers, linear dynamics [19] and Assumption 3 one can conclude that the relative distances in (17) and (18) converge to zero with probability one as $n^* \to \infty$ because the following holds with probability one:

$$
\lim_{n^* \to \infty} \tilde{x}_i = \bar{x}_i = z_i, \quad \lim_{n^* \to \infty} \tilde{w}_i = 0, \quad \forall k \notin S.
$$

Therefore, it results that the estimate $z_i$ with the update rule (15) is asymptotically optimal, i.e., PDSS strategy (16) is a solution for Problem 2.
B. Infinite-model strategy for Problem 2

The strategy 16 depends on the number of agents and covariance matrices of all sub-populations including those whose deep states are not observed, because $A_{k_i}^{\theta_i}$ and $\eta_{k_i}^{\theta_i}$ in the update rule 15 are computed based on the finite-population model with $P_t$ and $\phi_t$ in 12. However, it is possible to ignore the number of agents as well as the covariance matrices of sub-populations $k \notin S$ and propose a Kalman-like filter similar to 15 based on an infinite-population model, where $P_t$ and $\phi_t$ are given by 12, with $n^k$ is set to infinity, $\forall k \notin S$. In particular, for any sub-population $k \notin S$, one has $(\mu^k/n^k)\sigma_{k,v}^{\infty} = 0$ for every $k \notin S$ in the covariance matrix $\sigma_t$ of the resultant Kalman filter. For the special case of no-sharing information structure, i.e. $S = \emptyset$, matrix $\sigma_t$ is zero, which implies that the infinite-population Kalman filter of the risk-sensitive model is identical to that of the risk-neutral model (where $\theta = 0$).

Consider a PDSS strategy similar to 16 wherein the action of any agent $i$ of sub-population $k \in \mathbb{N}_K$ at time $t \in [0,T]$ is given by:

$$\tilde{u}_i^k = F^{\infty,\theta,k}_t(\tilde{x}_i^k - \alpha_{i,k}^{\infty,k}z_i^k) + \alpha_{i,k}^{\infty,k}P^{\infty,\theta,k}_tz_i^k + q_t^{\infty,\theta,i,k} + \alpha_{i,k}^t q_t^{\infty,\theta,k},$$

(20)

where the above matrices, drifts and estimates are computed based on an infinite-population model, where $n^k$ is set to infinity for every $k \notin S$.

Corollary 2. Let Assumptions 1–5 hold. Then, infinite-model PDSS strategy 20 is a solution for Problem 2.

Proof. The proof directly follows from the fact that Riccati equations 11 and 12 are bounded and continuous with respect to $n^k$, $k \in \mathbb{N}_K$.

Remark 3. Although both finite- and infinite-model PDSS strategies 16 and 20 converge to the same unique solution as $n^k \rightarrow \infty$, they have different prices of information when applied to the finite-population model. In particular, the strategy 16 takes the number of agents and covariance matrices into account while the strategy 20 ignores the above information. On the other hand, strategy 20 is simpler for analysis because the effect of an individual agent can be neglected in the infinite-population model.

IV. GENERALIZATIONS

A. Infinite-horizon cost function

In this subsection, we extend the results of Theorems 1 and 2 to the infinite-horizon cost function. Suppose that the model is time-homogeneous, and the cost function is given by:

$$\gamma_{\infty}(\theta) = \lim_{T \rightarrow \infty} \sup_{\theta_T} \frac{1}{T} \gamma_T(\theta).$$

(21)

The following stability assumption is imposed on the model.

Assumption 4. Let $(A^k, B^k)$ and $(A, B)$ be stabilizable, and $(A^{\infty}, Q^{\infty})$ and $(A, Q)$ be detectable, $\forall k \in \mathbb{N}_K$. In addition, suppose that algebraic forms of Riccati equations 11 and 12 admit positive definite solutions $P^k$, $k \in \mathbb{N}_K$, and $P$.

Theorem 3. Let Assumptions 1–5 hold. The following holds:

- Theorem 1 extends to the infinite-horizon cost function 21 such that the strategy 13 becomes stationary.
- Theorem 2 and Corollary 2 extend to the infinite-horizon cost function 21 under an additional condition that matrix $A + B \text{diag}(F^{\theta,1}_1, \ldots, F^{\theta,K}_1)$ must be Hurwitz. The above matrix is always Hurwitz when agents are dynamically decoupled.

Proof. The proof of the first part follows directly from the standard stability and detectability conditions 9, where algebraic forms of Riccati equations 11 and 12 admit bounded solutions. The proof of the second part, however, requires an additional condition making sure that the relative errors, defined in 17 and 18, will remain bounded as $T \rightarrow \infty$. In the case that matrix $A + B \text{diag}(F^{\theta,1}_1, \ldots, F^{\theta,K}_1)$ is Hurwitz, the dynamics of the errors, given by 19, becomes stable; hence, the limit to infinity exists. Consequently, when the dynamics of agents are decoupled, i.e., $A = \text{diag}(A^1, \ldots, A^K)$ and $B = \text{diag}(B^1, \ldots, B^K)$, matrix $(A + B \text{diag}(F^{\theta,1}_1, \ldots, F^{\theta,K}_1)$ becomes Hurwitz due to the fact that $A^k + B^k F^{\theta,k}_1$ is Hurwitz for any $k \in \mathbb{N}_K$ (as a stabilizing solution of the Riccati equation 11).

B. Multiple impact factors in a sub-population

So far, we have assumed that every sub-population has only one set of impact factors. In this subsection, we show that our results naturally extend to multiple sets of impact factors, where the interaction between agents in each sub-population is modelled by a number of orthonormal linear regressions. For any sub-population $k \in \mathbb{N}_K$, consider $m^k \in \mathbb{N}$ different sets of impact factors such that for every $m \in \mathbb{N}_{m^k}$:

$$x_t^m = \frac{1}{n_k} \sum_{i=1}^{n_k} \alpha_i^{m,k} \tilde{x}_i^k, \quad u_t^m = \frac{1}{n_k} \sum_{i=1}^{n_k} \alpha_i^{m,k} \tilde{u}_i^k, \quad \gamma_t^m = \frac{1}{n_k} \sum_{i=1}^{n_k} \alpha_i^{m,k} \tilde{\gamma}_i^k.$$

Denote by $\tilde{x}_t = ((\tilde{x}_t^m)^{m=1}_{m=m^k})^{K=1}_{K}$ the deep states of all sub-populations at time $t \in [0, \infty)$. The coupling in the dynamics 2 can be extended to the following form:

$$\sum_{m=1}^{m^k} \alpha_i^{m,k} \tilde{A}_t^{m,k} \tilde{x}_t^m \quad \text{and} \quad \sum_{m=1}^{m^k} \alpha_i^{m,k} \tilde{B}_t^{m,k} \tilde{u}_t^m.$$

The coupling in the cost function 3 remains as same as before. The gauge transformation 7 takes the following form: for every agent $i \in \mathbb{N}_{n^k}$ of sub-population $k \in \mathbb{N}_K$:

$$\Delta x_i^m := x_i^m - \sum_{m'=1}^{m^k} \alpha_i^{m',k} x_i^{m'}, \quad \Delta u_i^m := u_i^m - \sum_{m'=1}^{m^k} \alpha_i^{m',k} u_i^{m'}, \quad \Delta r_i^m := r_i^m - \sum_{m'=1}^{m^k} \alpha_i^{m',k} r_i^{m'},$$

(22)

Assumption 5. Let the impact factors be orthonormal vectors across every sub-population $k \in \mathbb{N}_K$ such that for any $m, m' \in \mathbb{N}_{m^k}$:

$$\frac{1}{n_k} \sum_{i=1}^{n_k} \alpha_i^{m,k} \alpha_i^{m',k} = 1 \quad (m \neq m').$$

Proposition 3. Let Assumption 5 hold. Then, the linear dependence and orthogonal properties introduced by the gauge
transformation hold. In particular, for every \( m \in \mathbb{N}_{m_k} \) and \( k \in \mathbb{N}_K \), the following relations hold:

\[
\sum_{i=1}^{n_k} \alpha^{i,m,k} \Delta x^i_t = \bar{0}_x \times 1, \quad \sum_{i=1}^{n_k} \alpha^{i,m,k} \Delta u^i_t = \bar{0}_u \times 1,
\]

and

\[
\begin{align*}
\left\{ \sum_{i=1}^{n_k} \alpha^{i,m,k} (\Delta x^i_t - \Delta x^i_0) \right\} \overline{Q}_t (z_{i,m,k}^T - z_{i,m,k}^0) &= 0, \\
\sum_{i=1}^{n_k} \alpha^{i,m,k} (\Delta u^i_t) \overline{R}_t (z_{i,m,k}^T - z_{i,m,k}^0) &= 0.
\end{align*}
\]

Proof. The proof follows directly from Assumption 5 and gauge transformation 22.

Remark 4 (Full decomposition). Consider a special case wherein all matrices in (6) are block diagonal, i.e., the dynamics in each sub-population \( k \in \mathbb{N}_K \) are coupled through the following components:

\[
\sum_{m=1}^{m_k} \alpha^{i,m,k} A^{m,k}_t \bar{x}^{m,k}_t + \sum_{m=1}^{m_k} \alpha^{i,m,k} B^{m,k}_t \bar{u}^{m,k}_t,
\]

and weighting matrices \( \overline{Q}_t^{k} \) and \( \overline{R}_t^{k} \), \( k \in \mathbb{N}_K \), are block diagonal. In this case, the global Riccati equation (12) decomposes to \( \sum_{k=1}^{K} m^k \) Riccati equations.

V. A Supply Chain Management Example

Consider a supplier that provides a particular product to its consumers (e.g., the bandwidth provided by telecommunication company). The product must be distributed to the consumers through a number of distributors (hubs), each of which has its own operating condition and capacity. The objective is to find a risk-sensitive solution for the supplier and distributors such that the delivered product is as close as possible to the supplier’s production level while the conditions of distributors are respected. To this end, let the supplier be the only agent in the first sub-population. Denote by \( x^i_t \) and \( u^i_t \), \( n^i = 1 \), respectively, the production level and control input of the supplier, normalized with respect to the number of distributors, at time \( t \in [0,T] \). The evolution of the state of the supplier is described by the following linearized dynamics:

\[
dx^i_t = (A^i x^i_t + B^i u^i_t)dt + dw^i_t,
\]

where \( w^i_t \) represents the uncertainty of the market, affecting the production level. Let the cost of the production be quadratic in the state and control input as follows:

\[
\frac{1}{2} \| x^i_t \|^2 Q + \int_0^T (\| x^i_t \|^2 Q + \| u^i_t \|^2 R_t) dt.
\]

The second sub-population is comprised of \( n^2 \) distributors with the following dynamics:

\[
dx^i_t = (A^2 x^i_t + B^2 u^i_t)dt + dw^i_t , \quad i \in \mathbb{N}_{n^2},
\]

where \( u^i_t \) and \( w^i_t \) are the control input and uncertainty of the distributor \( i \) at time \( t \in [0,T] \), respectively. Let \( r^i \) denote the desired operating point of the distributor \( i \) with the cost function:

\[
J_T^{1,2} = \| x^i_T - r^i \|^2 Q^2 + \int_0^T (\| x^i_t - r^i \|^2 Q^2 + \| u^i_t \|^2 R_t) dt.
\]

In addition, denote by \( \alpha^i \) the impact factor of the distributor \( i \in \mathbb{N}_{n^2} \), indicating its contribution in delivering the product, and by \( n^2 x^i_t \) the total distributed (delivered) product to consumers. We add a penalty function to the supplier’s cost function for the mismatch between the production level \( n^2 (2) d^i (1) \) and distributed product \( n^2 x^i_T \). Therefore, we define the following normalized cost function for the supplier:

\[
J_T := \| \bar{x}_t \|^2 Q^1 + \int_0^T (\| \bar{x}_t \|^2 Q^1 + \| \bar{u}_t \|^2 R_t) dt,
\]

where \( Q^1 := \left[ \begin{array}{c} Q^p' \quad -Q^p \\ -Q^p \quad Q^p \end{array} \right] \) and \( Q^p \) is the weighting matrix associated with the penalty function. Given the risk factor \( \theta \in \mathbb{R}_{>0} \), we define the following risk-sensitive cost function: \( \gamma_T (\theta) = \frac{1}{\theta} \log \mathbb{E} [e^{\theta (J_T + \sum_{i=1}^{n^2} J^i_{T}^{1,2})}] \).

Numerical examples

Since the complexity of the proposed strategies is independent of the number of distributors, we choose small \( n^2 \) in our simulations for ease of display. In particular, we consider the following parameters under DSS information structure:

\[
n^2 = 20, \quad T = 10, \quad \theta = 1, \quad A^1 = 0.4, \quad B^1 = 0.8, \quad Q^1 = 1, \quad R^1 = 1, \quad Q^p = 10, \quad A^2 = 2, \quad B^2 = 1, \quad Q^2 = 1, \quad R^2 = 0.1, \quad \sigma_{w,1} = 0.36, \quad \sigma_{w,2} = 9,
\]

where the sample time is 0.01. In the figures, the thick-red curve represents the trajectory of the normalized production level of the supplier \( x^1_t = \bar{x}^1_t \), the dashed-black curve shows the average of the distributed product \( x^2_T \), and the remaining curves are the trajectories of the distributors.

Example 1. Consider the case where all distributors have equal weights (impact factors), i.e. \( \alpha^i = 1, \forall i \in \mathbb{N}_{n^2} \). For any \( i \in \mathbb{N}_{n^2} \), let the desired operating point of distributor \( i \) be \( r^i = 0.5 + 2 o^i \), where \( o^i \) is a uniformly distributed random variable in the interval \([0, 1]\). A sample path of the states of the supplier and distributors under the optimal strategy is depicted in Figure 3. It is shown that the production level meets the distribution level while each distributor is rather close to its local desired operating point.

Example 2. Consider now the case where the weights (impact factors) of the distributors are not equal. To visualize the results better, let the variance of the distributors’ noise be \( \sigma_{w,2} = 1 \), and the desired operating point \( r^i \) be a uniformly distributed random variable in the interval \([0, 1]\), \( \forall i \in \mathbb{N}_{n^2} \). In Figure 3(a), the distribution is performed mainly by one distributor and it is observed that this distributor has to deviate from its local desired operating point to compensate for others. Similarly, in Figure 3(b) mainly two distributors perform the distribution function. In Figure 3(c), other distributors become more involved, and Figure 3(d) demonstrates the case where the distribution is carried out by all distributors equally, resulting in smoother trajectories.
Fig. 1. The trajectories of the supplier and distributors in Example 1.

Fig. 2. The trajectories of the states of the supplier and distributors in Example 2. (a) The impact factor of one distributor is 4.45 and all others 0.1; (b) impact factors of two distributors are 3.14 and all others are 0.1; (c) half of the distributors have the impact factor 0.1 and the other half have 1.41, and (d) all distributors have the same impact factor.

VI. CONNECTION TO DEEP NEURAL NETWORKS

The mathematics of neural networks can be traced back to the seminal work of Gauss [21] in regression theory and their application to the study of cat’s visual cortex in physiology [22], [23]. Due to computational limitations and early negative results, however, they did not receive much attention until recently [1], [2]. In general, a neural network consists of three layers: input, hidden and output, and when the number of hidden layers is more than one, the network is called deep. The main objective in deep learning is to find an end-to-end mapping from input to output by using some training data. The standard approach is to define a cost function, e.g., least square or cross entropy to penalize the distance between the network output and desired output (called supervised learning). To iteratively update the weights of the nodes in hidden layers, a stochastic gradient descent algorithm can be devised in a back-propagation manner based on the chain rule. When control action is added to deep learning, the resultant problem is called deep reinforcement learning, which is a more challenging problem.

With all the recent progresses, there are still some important challenges that need to be addressed in deep neural networks, some of which are described below.

- **Performance guarantee and interpretability:** Since deep neural networks are mostly designed based on intuition, empirical data, and trial and error experiments, it is often difficult to guarantee how well the trained network can perform under other sets of data. In addition, it is not clear how to interpret and connect the parameters of a deep neural network to those of a real-world system.

- **Numerous parameters and tunability:** There are so many parameters to tune such as depth, width, activation functions, initial weights and step sizes, each of which can lead to undesirable consequences such as under-fitting/over-fitting and vanishing/exploding gradient. Furthermore, since the optimization problem with respect to weights is a non-convex optimization problem, one would face with a large number of spurious local minima and saddle points wherein the gradient algorithm get stuck.

- **Troubleshooting:** A deep neural network is a giant black-box that maps an input to an output without providing much insight about how it does it. This ambiguity can be challenging, for example, when it comes to debugging an odd behaviour and troubleshooting the network.

In what follows, we highlight some aspects of deep control that can be useful for deep learning and more importantly for deep reinforcement learning. In general, deep control may be viewed as a reverse problem of deep learning wherein the model is known but the structure of the solution (which is a deep neural network) is unknown.

For simplicity of presentation, consider a special case of one sub-population with dynamically decoupled agents. Let $M \in \mathbb{N}$ denote the number of orthonormal features (that can be greater than the number of agents). The state dynamics of agent $i \in \mathbb{N}_n$ under the optimal strategy can be expressed by:

$$
\dot{x}_i^j = \sum_{j=1}^{n} W_t^{i,j}(\theta, \alpha)x_i^j + b_i^j(\theta, \alpha),
$$

where the optimal weight matrix is described as follows:

$$
W_t^{i,i}(\theta, \alpha) := A_t + B_t((1 - \frac{1}{n} \sum_{m=1}^{M} (\alpha^{i,m})^2)F_\theta^t + \frac{1}{n} \sum_{m=1}^{M} \sum_{m' = 1}^{M} \alpha^{i,m} \alpha^{i,m'} F_{\theta,m,m'}^t),
$$

and for $j \neq i$,

$$
W_t^{j,i}(\theta, \alpha) := \frac{1}{n} \sum_{m=1}^{M} \sum_{m' = 1}^{M} \alpha^{i,m} (-\alpha^{j,m} F_\theta^t + \sum_{m' = 1}^{M} \alpha^{j,m'} F_{\theta,m,m'}^t),
$$

where the bias term is given by:

$$
b_i^j(\theta, \alpha) := B_t(q_i^j + \sum_{m=1}^{M} \alpha^{i,m} q_{i,m}^\theta).$$
An interesting observation is that the structure of the optimal network \((23)\) remains independent of the number of orthonormal features \(M\), although the complexity of computing the optimal weight matrix increases with the number of features. In other words, a simple deep neural network can incorporate arbitrary number of orthonormal features into a fixed number of neurons \(n\). This finding is consistent with the experimental successes of deep neural networks to model highly complex structures.

One salient characteristic of the structure \((23)\) is the fact that it is interpretable, i.e., it provides a one-to-one relationship between the parameters of a deep neural network and those of a control system presented in Section \[\] In addition, since weight initialization is a key step in training a deep neural network, it is possible to use the above optimal weight matrix for computing initial conditions (that is computed by a deep Riccati equation for a network consisting of \(n\) neurons with depth \(T\), where each neuron has its own impact factor and nominal operating point). Moreover, to efficiently incorporate a certain design, it is sometimes necessary to adjust the weights at all layers in a consistent manner, simultaneously. For example, the robust parameter \(\theta\) can automatically tune the weights in such a way that the resultant design is resilient to uncertainties or by selecting large penalty functions on control actions, it would be possible to promote sparsity. This design property can be useful to ensure that the network is not over-fitted (i.e., there is room for uncertainties) or does not contain excessive number of coefficients. Furthermore, the above results can be extended to constraint deep control systems where quadratic programming and model predictive control can be utilized. For instance, ReLU (rectified linear unit) activation function can be viewed as a sufficient condition for an inequality constraint (i.e., \(x^i_t \geq 0, \; i \in \mathbb{N}_n\)) such that:

\[
\dot{x}_t^i = f\left(\sum_{j=1}^{n} W_{ij}^\theta(\theta, \alpha)x_j^i + b_i(\theta, \alpha)\right), \tag{24}
\]

where \(f(\cdot) := \max(0, \cdot)\). This formulation is consistent with the fact that ReLU activation function has been successfully used in classification problems, e.g., image processing applications, where states of interest are in the form of probability (frequency) of the occurrence of a feature, which is a non-negative number. Hence, ReLU activation function may be interpreted as a sufficient condition ensuring the feasible states are non-negative.

VII. CONCLUSIONS

Inspired by deep learning that deals with big data and provides an efficient representation for modelling complex systems, we introduced deep linear quadratic control problem in this paper in order to control complex systems with a large number of decision makers, by developing a low-dimensional deep Riccati equation. In particular, two non-classical information structures were studied, i.e., deep-state sharing and partial deep-state sharing, where the optimal solution for the former information structure and two sub-optimal solutions for the latter structure were obtained. The main results were also extended to infinite-horizon cost function. To illustrate the efficacy of the proposed results, two examples of supply-chain management system were provided.

In addition, the potential impact of the obtained results in enhancing our understanding of deep neural networks was demonstrated. An interesting extension is to use reinforcement learning methods in deep control systems to learn how to control complex networked systems. The authors have an upcoming paper on the above extension.

REFERENCES

[1] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” nature, vol. 521, no. 7553, p. 436, 2015.
[2] J. Schmidhuber, “Deep learning in neural networks: An overview,” Neural Networks, vol. 61, pp. 85–117, 2015.
[3] J. Arabneydi and A. G. Aghdam, “Deep teams: Decentralized decision making with finite and infinite number of agents,” in IEEE Transactions on Automatic Control, 2020.
[4] J. Arabneydi, “New concepts in team theory: Mean field teams and reinforcement learning,” Ph.D. dissertation, McGill University, 2016.
[5] J. Arabneydi and A. Mahajan, “Team-optimal solution of finite number of mean-field coupled LQG subsystems,” in Proceedings of the 54th IEEE Conference on Decision and Control, 2015, pp. 5308–5313.
[6] ——, “Linear quadratic mean field teams: Optimal and approximately optimal decentralized solutions,” Available at https://arxiv.org/abs/1609.00056, 2016.
[7] J. Arabneydi, M. Baharloo, and A. G. Aghdam, “Optimal distributed control for leader-follower networks: A scalable design,” in Proceedings of the 31st Canadian Conference on Electrical and Computer Engineering, 2018, pp. 1–4.
[8] M. Baharloo, J. Arabneydi, and A. G. Aghdam, “Near-optimal control strategy in leader-follower networks: A case study for linear quadratic mean-field teams,” in Proceedings of the 57th IEEE Conference on Decision and Control, 2018, pp. 3288–3293.
[9] T. Başar and P. Bernhard, H-infinity optimal control and related minimax design problems: A dynamic game approach. Birkhäuser Basel, 2008.
[10] J. A. Dieudonné and J. B. Carrell, “Invariant theory, old and new,” Advances in mathematics, vol. 4, no. 1, pp. 1–80, 1970.
[11] E. Noether, “Invariant variation problems,” Transport Theory and Statistical Physics, vol. 1, no. 3, pp. 186–207, 1971.
[12] D. Jacobson, “Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games,” IEEE Transactions on Automatic control, vol. 18, no. 2, pp. 124–131, 1973.
[13] P. Whittle, “Risk-sensitive linear/quadratic/Gaussian control,” Advances in Applied Probability, vol. 13, no. 4, pp. 764–777, 1981.
[14] A. Jadbaie, J. Lin, and A. S. Morse, “Coordination of groups of mobile autonomous agents using nearest neighbor rules,” IEEE Transactions on Automatic Control, vol. 48, no. 6, pp. 988–1001, 2003.
[15] L. Xiao and S. Boyd, “Fast linear iterations for distributed averaging,” Systems and Control Letters, vol. 53, no. 1, pp. 65–78, 2004.
[16] K. Moriyasu, An elementary primer for gauge theory. World Scientific, 1983.
where for any set of coefficients $P$ of matrix set of solutions that is invariant to right and left multiplication $PA = PB$.

In particular, one has:

$$
\|Px_t\|_{Q_t} = x_t^T P_t Q_t x_t = \text{Tr}(x_t^T P_t \sum_{h=-H}^{H} q_t(h)P^h P x_t)
$$

$$
= \sum_{h=-H}^{H} q_t(h) \text{Tr}(x_t^T S A^{h+2} S^T x_t)
$$

$$
= \sum_{h=-H}^{H} q_t(h) \text{Tr}(x_t^T \sum_{m=1}^{n} \lambda_m^{h+2} v_m v_m^T x_t)
$$

$$
= \sum_{h=-H}^{H} \lambda_m^{2h} \text{Tr}(x_t^T v_m (\sum_{h=-H}^{H} \lambda_m^{h} q_t(h)) v_m^T x_t)
$$

$$
= \sum_{m=1}^{n} \lambda_m^{2h} \|v_m^T x_t\|_{Q_t}^2 = \sum_{m=1}^{n} \lambda_m^{2h} q_t(h).
$$

(25)

and for every $m \in \mathbb{N}_n$,

$$
\|proj(x_t,v_m)\|_{Q_t} = (\|x_t^T v_m\|^2 \text{Tr}(v_m^T S A^{h+2} S^T v_m) v_m)
$$

$$
= (x_t^T v_m)^2 \text{Tr}(\sum_{h=-H}^{H} q_t(h)P^h v_m)
$$

$$
= (x_t^T v_m)^2 (\sum_{h=-H}^{H} q_t(h) \text{Tr}(v_m^T S A^{h+2} S^T v_m))
$$

$$
= \|v_m^T x_t\|_{Q_t}^2 \sum_{h=-H}^{H} \lambda_m^{2h} q_t(h).
$$

(26)

From (25) and (26), it results that the condition of equivariant cost holds. A similar argument is valid for $R_t$. The proof is completed, on noting that $v_m^T x_t = \sqrt{n} \bar{x}_m$ and $v_m^T u_t = \sqrt{n} \bar{u}_m$.

### Appendix B

**Proof of Proposition 2**

The condition of equivariant dynamics reduces to that of exchangeable dynamics in [4, Definition 2.1] for every permutation matrix. In addition, it is well known that for any permutation matrix $\{\lambda_m, -\lambda_m\} = 1$, $m \in \mathbb{N}_n$. Hence, the equivariant cost reduces to exchangeable cost, where for any permutation matrices $P$ and $P'$:

$$
\sum_{m=1}^{n} \|proj(x_t,v_m)\|_{Q_t} = \sum_{m=1}^{n} \|proj(x_t,v'_m)\|_{Q_t}.
$$

By writing the above vectors, i.e. $\{x_t, (v_m,v'_m)_{m=1}^{n}\}$, in the standard basis, one can show that the structure of matrix $Q_t$ has to be in the form of $q_t I_{n \times n} + n \bar{q}_t I_{n \times n}$. For more details, see [4, Proposition 2.1].

**Appendix A

**Proof of Proposition 1**

Since $P$ is a symmetric real-valued matrix, there exists a spectral decomposition $P = S \Lambda S^T$, where orthogonal matrix $S$ and diagonal matrix $\Lambda$ contain eigenvector and eigenvalues of $P$, respectively. The condition of the equivariant dynamics in Definition 1 holds if $PA_t = A_t' P$ and $PB_t = B_t' P$. One set of solutions that is invariant to right and left multiplication of matrix $P$ is Laurent polynomial functions of $P$ and $P^{-1}$, where for any set of coefficients $a_t(h) \in \mathbb{R}$, $h \in \mathbb{Z}_H$, $H \in \mathbb{N}$:

$$
A_t = \sum_{h=-H}^{H} a_t(h) P^h = \sum_{h=-H}^{H} a_t(h) S \Lambda^h S^T
$$

$$
= \sum_{h=-H}^{H} a_t(h) (\sum_{m=1}^{n} \lambda_m^h v_m v_m^T) = a_t I_{n \times n} + \sum_{m=1}^{n} \bar{a}_m^h v_m v_m^T,
$$

where $a_t = a_t(0)$ and $\bar{a}_m^h = \sum_{h=-H, h \neq 0}^{H} \lambda_m^h a_t(h)$, $m \in \mathbb{N}_n$.

A similar argument holds for $PB_t = B_t' P$. In addition, any Laurent polynomial function of $P$ and $P^{-1}$ is a solution for $Q_t$ wherein for any coefficients $q_t(h)$, $h \in \mathbb{Z}_H$, $H \in \mathbb{N}$:

$$
Q_t := \sum_{h=-H}^{H} q_t(h) P^h = \sum_{h=-H}^{H} q_t(h) S \Lambda^h S^T
$$

$$
= \sum_{h=-H}^{H} q_t(h) (\sum_{m=1}^{n} \lambda_m^h v_m v_m^T) = q_t I_{n \times n} + \sum_{m=1}^{n} \bar{q}_m^h v_m v_m^T,
$$

where $q_t = q_t(0)$ and $\bar{q}_m^h = \sum_{h=-H, h \neq 0}^{H} \lambda_m^h q_t(h)$, $m \in \mathbb{N}_n$. For more details, see [4, Proposition 2.1].