

SUBJECTIVE EQUILIBRIA UNDER BELIEFS OF EXOGENOUS UNCERTAINTY: LINEAR QUADRATIC CASE

G. ARSLAN^{1,*}, S. YÜKSEL²

ABSTRACT. We consider a stochastic dynamic game where the players have their own linear state dynamics and quadratic cost functions. The players are coupled through some environment variables, generated by another linear system driven by the states and decisions of all players. Each player observes his own states realized up to the current time as well as the past realizations of his own decisions and the environment variables. Each player (incorrectly) believes that the environment variables are generated by an independent exogenous stochastic process. In this setup, we study the notion of “subjective equilibrium under beliefs of exogenous uncertainty (SEBEU)”. At an SEBEU, each player’s strategy is optimal with respect to his subjective belief; moreover, the objective probability distribution of the environment variables is consistent with players’ subjective beliefs. We construct an SEBEU in pure strategies, where each player strategy is an affine function of his own state and his estimate of the system state.

Keywords: Game Theory, Subjective Equilibrium, Distributional Consistency, Stochastic Dynamic Games.

AMS Subject Classification: 91A15, 93E20.

1. INTRODUCTION

In this paper, we consider a linear-quadratic non-cooperative dynamic game played in discrete-time. In our model, the players have their own state and decision variables; furthermore, our model includes the so-called “environment variables”, that are common to all players. The state and decision variables of all players as well as the environment variables belong to finite-dimensional Euclidian spaces. Each player’s state evolves according to a linear stochastic recursion driven by his own decisions as well as the environment variables. Environment variables, on the other hand, are generated by a separate linear stochastic recursion driven by the state and decision variables of all players. At each stage, the players select their decisions simultaneously based on their own information available at the time, which consist of their own state variables realized up to the current time as well as the past realizations of their own decisions and the common environment variables. Following the collective decisions of all players during each stage, the environment variables for that stage are realized, the players incur their stage-costs, and all of the system states evolve to the next stage. Each player’s stage-cost is determined according to a quadratic function of the player’s own state and decision variable as well as the common environment variables realized during the stage. Each player wishes to minimize the expected value of his own long-term cost, that is a discounted sum of his stage-costs.

In our setup, each player’s problem depends on the decisions of the other players only through the environment variables. However, a player has no knowledge about the other players, not even the presence of the other players. As mentioned above, each player observes the past realizations

¹Department of Electrical and Computer Engineering, University of Hawaii at Manoa, 2540 Dole Street, Honolulu, HI 96822

²Department of Mathematics and Statistics, Queens University, Kingston, ON K7L 3N6, Canada
e-mail: yuksel@mast.queensu.ca

*Corresponding author e-mail: gurdal@hawaii.edu

Manuscript received 16 October 2023.

of the environment variables during any stage but he has no knowledge of how the environment variables are generated. When the strategies of the other players are fixed, each player faces a partially observed sequential decision problem, however, the player is not aware of this fact. As such, each player incorrectly assumes that the environment variables, that are sequentially revealed to the player, is an independent random sequence that is exogenous to his problem. Each player has a belief represented by a subjective probability distribution for the environment variables. Under such beliefs, the players seek to minimize their expected long-term costs by choosing “optimal strategies”.

An equilibrium situation arises when players employ optimal strategies, with respect to their beliefs, and the resulting objective marginal probability distribution for the environment variables is consistent with their subjective beliefs. Such strategies and beliefs are said to form a subjective equilibrium under beliefs of exogenous uncertainty (SEBEU) [3]. At an SEBEU, if the objective marginal probability distribution of the environment variables is revealed to a player (who does not question his assumption of independence and exogeneity for the environment variables) would have no reason to change his belief; consequently, he/she would have no incentive to change his strategy.

The notion of the SEBEU is related to the notion of price-taking behavior in mathematical economics [2, 21] and can be regarded as a generalization of price-taking behavior to stochastic dynamic games. The notion of the SEBEU is distinct from the well-known notions of Nash equilibrium and mean-field equilibrium [16, 19] as well as from the other notions of subjective equilibrium in the literature [1, 10, 11, 18, 25, 26]. In particular, a Nash equilibrium in our setup would require each player to take into account the effects of his own decisions on the environment variables that, in turn, affects the player’s cost as well as his state evolution. In general, this leads to an infinite regression where each player needs to estimate the other players’ states, the other players’ estimates of their opponents’ states, and so on. The players at an SEBEU ignore such closed-loop effects, which essentially decouples player problems and leads to a significant simplification of the overall problem. The notion of stationary equilibrium in [1] is hinged on a condition of consistency between the player beliefs on the population states being constant and the long-run average of the actual population states (over an infinite-population); see also the analogous notion of oblivious equilibrium in [25]. We also selectively cite [13] where the players make optimal decisions by maintaining subjective nested beliefs over others types and their beliefs about others. We refer the reader to [3] for a review of the related literature vis-a-vis the notion of the SEBEU as well as for a motivational application on the Autonomous Demand Response (ADR) problem on price-sensitive control of energy usage in multiple (possibly very large number of) units, e.g., [9, 22].

The existence of the SEBEU and its various properties are studied in [3] for compact-metric space models, i.e., all state, decision, and environment variables belong to compact metric spaces. In particular, the existence of an SEBEU in mixed strategies is established; moreover, an SEBEU is shown to be an approximate Nash equilibrium in large-scale games. In this paper, our main objective is to obtain the counterparts of the results of [3] for linear-quadratic models described above, which fall outside the compact-metric models considered in [3]. Moreover, we aim to obtain more explicit and refined results by exploiting the specific linear-quadratic structure of the models.

A significant body of literature, pioneered in large part by Tamer Başar over several seminal publications studies multi-stage linear-quadratic games under various asymmetric information patterns and demonstrates the richness and fragility of equilibrium properties to informational changes. A linear-quadratic-Gaussian (LQG) game in which each player makes his own noisy linear observations of the system state in addition to having access to the past (but not the current) observations of the system state by his opponents (called one-step delay observation sharing pattern) is studied in [4, 5]. The existence of a unique Nash equilibrium, that is affine in the information variables, via contraction properties of best-response dynamics is established in

several studies including [4, 5], whereas exact expressions are presented in [5] for the two-player case; see also [20]. A remarkable insight is that if one also shares actions, leading to one-step-delay sharing, the Nash equilibria become non-unique even when the one-step delay observation sharing would lead to a unique equilibrium [4, 5]. For further results on informational fragility and existence of equilibria for static and dynamic problems, we refer the reader to [6, 7, 23] and [27, Chapter 10].

Among further relevant studies, [14] shows the existence of a unique Markov perfect equilibrium in two-player LQG games where the posterior beliefs of the players on the system state given the common information evolves independently of the player strategies. [24] considers two-player finite-horizon LQG games where each player observes his own state (but not his opponent's state) and the joint decisions. Under certain conditions, [24] constructs a perfect Bayesian equilibrium in such games. [8] shows that equilibrium strategies can be computed in zero-sum two-team linear-quadratic games with decentralized information structures under a mutual quadratic invariance condition. To the best of the authors' knowledge, the existing results in the literature are not applicable to the general linear-quadratic models considered in this paper.

Contributions. We present sufficient conditions, which involve only the primitive problem data, under which the SEBEU strategies can be explicitly constructed in pure strategies for the linear-quadratic models. Our construction, for both finite-horizon and infinite-horizon cases, lead to pure strategies that are affine functions of players' own states and their estimates of the system state. When the primitive random variables are Gaussian, the players' estimates of the system state can be generated by a linear recursion, i.e., a Kalman filter. We note that a SEBEU for the linear setup considered in our paper (when the primitive random variables are Gaussian) admits solutions that are always linear, whenever an equilibrium exists. On the other hand, a Nash equilibrium may or may not be linear and to our knowledge, beyond the setups studied in [4, 5] and noted in [27, Chapter 10], the existence and structure of Nash equilibria in partially observed settings has not been studied in the literature. Finally, when the number of decision makers is infinite, we show that an SEBEU is also a Nash equilibrium.

Notation. \mathbb{N}_0 and \mathbb{N} denote the nonnegative and positive integers, respectively; \mathbb{R} denotes the real numbers; $[\cdot]'$ denotes the transpose of a vector or matrix; $|\cdot|$ denotes the Euclidean norm for a vector; $|x|_Q^2 := x'Qx$ for a vector x and a matrix Q ; $A \succeq 0$ denotes a nonnegative definite matrix A , whereas $A \succ 0$ denotes a positive definite matrix A ; $\mathcal{P}(\mathbb{T})$ denotes the set of probability measures on the Borel sigma algebra $\mathcal{B}(\mathbb{T})$ of a topological space \mathbb{T} ; $P[\cdot]$, $E[\cdot]$, and $\text{cov}[\cdot]$ denote the probability, the expectation, and the covariance, respectively, $(E^\mu[\cdot])$ is also used to emphasize the underlying strategy μ ; $f(x) = \mathcal{O}(x)$ means $\limsup_{x \rightarrow 0} |f(x)/x|$ is finite.

2. SUBJECTIVE EQUILIBRIUM UNDER BELIEFS OF EXOGENOUS UNCERTAINTY

2.1. Model and definitions. Consider a decentralized stochastic system with N decision makers, $N \in \mathbb{N}$, where the i -th decision maker is referred to as DM^i , $i \in [1, N] := \{1, \dots, N\}$. Time variable $t \in [0, T) := \{0, \dots, T-1\}$ is an integer where $T \in \mathbb{N}$ is the time horizon; we will also allow for $T = \infty$ in which case $[0, T) = \mathbb{N}_0$. Each DM^i has its own state $x_t^i \in \mathbb{R}^{n_x^i}$, control input $u_t^i \in \mathbb{R}^{n_u^i}$, and random disturbance $w_t^i \in \mathbb{R}^{n_w^i}$ at time $t \in [0, T)$, where $n_x^i, n_u^i, n_w^i \in \mathbb{N}$. In addition, there are so-called environment variables denoted by $y_t \in \mathbb{R}^{n_y}$ at time $t \in [0, T)$, where $n_y \in \mathbb{N}$. Each DM^i has linear state dynamics as:

$$x_{t+1}^i = A_t^i x_t^i + B_t^i u_t^i + C_t^i y_t + w_t^i, \quad t \in [0, T) \quad (1)$$

starting from some (possibly random) initial state x_0^i , and a quadratic cost function given by:

$$c_t^i(x^i, u^i, y) = |x^i|_{Q_t^i}^2 + |u^i|_{R_t^i}^2 + 2y'(K_t^i u^i + L_t^i x^i), \quad (2)$$

where $A_t^i, B_t^i, C_t^i, Q_t^i, R_t^i, K_t^i, L_t^i$ are appropriate dimensional matrices ($Q_t^i \succ 0, R_t^i \succ 0$), for all $i \in [1, N], t \in [0, T)$. Each DM^i has also a quadratic terminal cost function $c_T^i(x^i) = |x^i|_{Q_T^i}^2$, where $Q_T^i \succ 0$, when $T \in \mathbb{N}$ ($c_T^i \equiv 0$ when $T = \infty$). The environment variables are generated by the linear dynamics

$$x_{t+1}^0 = A_t^0 x_t^0 + \sum_{i \in [1, N]} (B_t^{1,i} u_t^i + B_t^{2,i} x_t^i) + w_t^0, \tag{3}$$

$$y_t = D_t x_t^0 + \sum_{i \in [1, N]} (E_t^{1,i} u_t^i + E_t^{2,i} x_t^i) + \xi_t, \tag{4}$$

where $x_t^0 \in \mathbb{R}^{n_x^0}$ is the state of the environment dynamics (x_0^0 is some possibly random initial state), $\xi_t \in \mathbb{R}^{n_\xi}$, $w_t^0 \in \mathbb{R}^{n_w^0}$ are random disturbances, for some $n_x^0, n_\xi, n_w^0 \in \mathbb{N}$, and $A_t^0, B_t^{1,i}, B_t^{2,i}, D_t, E_t^{1,i}, E_t^{2,i}$ are appropriate dimensional matrices, for all $i \in [1, N], t \in [0, T)$.

Each DM^i has access to the information (I_t^i, Y_{t-1}) at time $t \in [0, T)$, where

$$I_t^i := ((x_k^i)_{k \in [0, t]}, (u_k^i)_{k \in [0, t]}), \quad Y_{t-1} := (y_k)_{k \in [0, t]}, \quad t \in [0, T). \tag{5}$$

A pure strategy for DM^i is a collection $s^i = (s_t^i)_{t \in [0, T)}$ of measurable mappings $s_t^i : \mathbb{I}_t^i \times \mathbb{Y}^{[0, t]} \rightarrow \mathbb{U}^i$, where $\mathbb{I}_t^i := (\mathbb{R}^{n_x^i})^{[0, t]} \times (\mathbb{R}^{n_u^i})^{[0, t]}$, $\mathbb{Y} := \mathbb{R}^{n_y}$, $t \in [0, T)$. For both $T \in \mathbb{N}$ and $T = \infty$, $(I_T^i, Y_{T-1}) \in \mathbb{I}_T^i \times \mathbb{Y}^{[0, T]}$ denotes DM^i 's entire history of observations. We let \mathbb{S}^i denote the set of pure strategies for each DM^i . If DM^i employs a pure strategy $s^i \in \mathbb{S}^i$, then

$$u_t^i = s_t^i(I_t^i, Y_{t-1}), \quad t \in [0, T). \tag{6}$$

For any joint pure strategy $s = (s^1, \dots, s^N) \in \mathbb{S} := \mathbb{S}^1 \times \dots \times \mathbb{S}^N$ employed, each DM^i incurs a long-term cost

$$\bar{J}^i(s^i, s^{-i}) = E \left[\sum_{t \in [0, T)} (\beta^i)^t c_t^i(x_t^i, u_t^i, y_t) + (\beta^i)^T c_T^i(x_T^i) \right], \tag{7}$$

where $s^{-i} \in \mathbb{S}^{-i} := \times_{j \neq i} \mathbb{S}^j$ denotes the pure strategies of all DMs other than DM^i , $\beta^i \in [0, 1]$ is DM^i 's discount factor ($\beta^i \in [0, 1)$ when $T = \infty$). The expectation in (7) is taken with respect to the probability distribution over the set $\mathbb{I}_T^i \times \mathbb{Y}^{[0, T)}$ of histories induced by $s \in \mathbb{S}$. Note that each DM^i 's cost is influenced by the other DMs only through the environment variables Y_{T-1} .

Suppose now that each DM^i views the sequence of environment variables as an **independent exogenous** (that is, primitive, generated by nature independently of the all other primitive random variables) random sequence $Z_{T-1} := (z_t)_{t \in [0, T)}$ with a given probability distribution $\zeta \in \mathcal{P}(\mathbb{Y}^{[0, T)})$ instead of the endogenous sequence Y_{T-1} generated by (1)-(6). We note that Z_{T-1} need not be a collection of independent random variables. We define DM^i 's cost corresponding to a pure strategy $s^i \in \mathbb{S}^i$ and a deterministic sequence of environment variables $\bar{Z}_{T-1} = (\bar{z}_t)_{t \in [0, T)} \in \mathbb{Y}^{[0, T)}$ as:

$$J^i(s^i, \bar{Z}_{T-1}) := E \left[\sum_{t \in [0, T)} (\beta^i)^t c_t^i(x_t^i, u_t^i, \bar{z}_t) + (\beta^i)^T c_T^i(x_T^i) \right], \tag{8}$$

where

$$\begin{aligned} x_{t+1}^i &= A_t^i x_t^i + B_t^i u_t^i + C_t^i \bar{z}_t + w_t^i, & t \in [0, T), \\ u_t^i &= s_t^i(I_t^i, \bar{Z}_{t-1}), & t \in [0, T), \end{aligned}$$

with $\bar{Z}_{t-1} := (\bar{z}_k)_{k \in [0, t)}$, for $t \in [0, T)$. Note that the expectation in (8) is taken with respect to the probability distribution over the set \mathbb{I}_T^i of DM^i 's state-control histories induced by $(s^i, \bar{Z}_{T-1}) \in \mathbb{S}^i \times \mathbb{Y}^{[0, T)}$. Accordingly, each DM^i who models the environment variables as an independent exogenous random sequence with the probability distribution $\zeta \in \mathcal{P}(\mathbb{Y}^{[0, T)})$ aims to minimize $E^\zeta[J^i(s^i, \cdot)]$ by choosing a pure strategy $s^i \in \mathbb{S}^i$. Therefore, from DM^i 's viewpoint,

no DM has any influence on the environment variables and the cost $E^\zeta[J^i(s^i, \cdot)]$ is independent of the strategies of all DMs other than DMⁱ.

To ensure that the long-term cost (7) and (8) are well-defined for any joint strategy $s \in \mathbb{S}$ and independent exogenous distribution $\zeta \in \mathcal{P}(\mathbb{Y}^{[0,T]})$ for the environment variables, we make the following assumption throughout this paper.

Assumption 1. $X_0, W_0, W_1, \dots, \xi_0, \xi_1, \dots$, are mutually independent and have finite second-order moments where $X_t := (x_t^i)_{i \in [0, N]}$, $W_t := (w_t^i)_{i \in [0, N]}$.

Definition 1. A joint strategy $s = (s^1, \dots, s^N) \in \mathbb{S}$ is called a *subjective equilibrium under beliefs of exogenous uncertainty (SEBEU) in pure strategies* if there exists subjective beliefs $\zeta^1, \dots, \zeta^N \in \mathcal{P}(\mathbb{Y}^{[0,T]})$, where ζ^i is DMⁱ's belief about the environment variables, such that the following two conditions hold.

- (i) Each DMⁱ's strategy s^i is optimal with respect to its subjective belief ζ^i , i.e.,

$$E^{\zeta^i}[J^i(s^i, \cdot)] \leq E^{\zeta^i}[J^i(\tilde{s}^i, \cdot)], \quad \forall i \in [1, N], \tilde{s}^i \in \mathbb{S}^i$$

- (ii) Each DMⁱ's subjective belief ζ^i is consistent with the objective probability distribution ζ_s of the environment variables, i.e.,

$$\zeta^i(B) = \zeta_s(B), \quad \forall i \in [1, N], B \in \mathcal{B}(\mathbb{Y}^{[0,T]}),$$

where $\zeta_s \in \mathcal{P}(\mathbb{Y}^{[0,T]})$ is generated by s endogenously through (1)-(6).

Note that, at an SEBEU, each DMⁱ ignores the influence of its own strategy s^i on the environment variables in minimizing its long-term cost. In contrast, the well-known concept of *Nash equilibrium in pure strategies* requires a joint pure strategy $s = (s^1, \dots, s^N) \in \mathbb{S}$ to satisfy

$$\bar{J}^i(s^i, s^{-i}) \leq \bar{J}^i(\tilde{s}^i, s^{-i}), \quad \forall i \in [1, N], \tilde{s}^i \in \mathbb{S}^i, \quad (9)$$

where each DMⁱ takes into account the entire influence of its own strategy s^i on its long-term cost including through the environment variables.

3. LINEAR QUADRATIC SYSTEMS

In this section, we show how to construct an SEBEU in pure strategies under certain conditions.

3.1. Finite-horizon case ($T \in \mathbb{N}$). Suppose that DMⁱ instead aims to minimize $E^\zeta[J^i(s^i, \cdot)]$ over the set of pure strategies \mathbb{S}^i for an independent exogenous sequence $(z_t)_{t \in [0, T]}$ of environment variables with a given probability distribution $\zeta \in \mathcal{P}(\mathbb{Y}^{[0, T]})$. If Assumption 1 holds and $\max_{t \in [0, T]} E[|z_t|^2] < \infty$, an optimal strategy minimizing $E^\zeta[J^i(s^i, \cdot)]$ over $s^i \in \mathbb{S}^i$ is obtained as:

$$s_t^i(I_t^i, Z_{t-1}) = F_t^i x_t^i + \sum_{k \in [t, T]} G_{t,k}^i E[z_k | Z_{t-1}] + H_t^i, \quad t \in [0, T],$$

where $(F_t^i)_{t \in [0, T]}$, $(G_{t,k}^i)_{0 \leq t \leq k < T}$, $(H_t^i)_{t \in [0, T]}$ are matrices of appropriate dimensions, which are pre-computable independently of the sequence $(z_t)_{t \in [0, T]}$ of environment variables; see Appendix A.1.

Therefore, a joint strategy $s = (s^i)_{i \in [1, N]} \in \mathbb{S}$ is an SEBEU if

$$s_t^i(I_t^i, Y_{t-1}) = F_t^i x_t^i + \sum_{k \in [t, T]} G_{t,k}^i E[y_k | Y_{t-1}] + H_t^i, \quad i \in [1, N], t \in [0, T] \quad (10)$$

and

$$y_t = D_t x_t^0 + \sum_{j \in [1, N]} E_t^{1,j} \left(F_t^j x_t^j + \sum_{k \in [t, T]} G_{t,k}^j E[y_k | Y_{t-1}] + H_t^j \right) + \sum_{j \in [1, N]} E_t^{2,j} x_t^j + \xi_t, \quad (11)$$

$$x_{t+1}^0 = A_t^0 x_t^0 + \sum_{j \in [1, N]} B_t^{1,j} \left(F_t^j x_t^j + \sum_{k \in [t, T]} G_{t,k}^j E[y_k | Y_{t-1}] + H_t^j \right) + \sum_{j \in [1, N]} B_t^{2,j} x_t^j + w_t^0, \quad (12)$$

$$x_{t+1}^j = A_t^j x_t^j + B_t^j \left(F_t^j x_t^j + \sum_{k \in [t, T]} G_{t,k}^j E[y_k | Y_{t-1}] + H_t^j \right) + C_t^j y_t + w_t^j, \quad j \in [1, N] \quad (13)$$

provided $\max_{t \in [0, T)} E[|y_t|^2] < \infty$. This reveals that an SEBEU of the form above can be obtained if a random sequence of environment variables having finite second-order moments and satisfying (11)-(13), called equilibrium environment variables, can be constructed. We next attempt to construct equilibrium environment variables via the conditional expectations $E[y_k | Y_{t-1}]$, $0 \leq t \leq k < T$.

We write (11)-(13) in the vector form as:

$$y_t = \mathcal{D}_t X_t + \sum_{k \in [t, T)} \mathcal{G}_{t,k}^p E[y_k | Y_{t-1}] + \mathcal{H}_t^p + \xi_t, \quad (14)$$

$$X_{t+1} = \mathcal{A}_t X_t + \sum_{k \in [t, T)} \mathcal{G}_{t,k}^X E[y_k | Y_{t-1}] + \mathcal{H}_t^X + C_t y_t + W_t, \quad (15)$$

where $\mathcal{D}_t, \mathcal{G}_{t,k}^p, \mathcal{H}_t^p, \mathcal{A}_t, \mathcal{G}_{t,k}^X, \mathcal{H}_t^X, C_t, 0 \leq t \leq k < T$, are matrices of appropriate dimension. For each $k \in [0, T)$ and arbitrary $\hat{X}_{k|k-1}$, we introduce the following linear equations

$$\left\{ \begin{array}{l} \hat{y}_{t|k-1} = \mathcal{D}_t \hat{X}_{t|k-1} + \sum_{n \in [t, T)} \mathcal{G}_{t,n}^p \hat{y}_{n|k-1} + \mathcal{H}_t^p + \hat{\xi}_t, \\ \hat{X}_{t+1|k-1} = \mathcal{A}_t \hat{X}_{t|k-1} + \sum_{n \in [t, T)} \mathcal{G}_{t,n}^X \hat{y}_{n|k-1} + \mathcal{H}_t^X + C_t \hat{y}_{t|k-1} + \hat{W}_t \end{array} \right\}, \quad t \in [k, T), \quad (16)$$

where $\hat{\xi}_t := E[\xi_t]$, $\hat{W}_t := E[W_t]$, and $\hat{y}_{t|k-1}$ are the unknowns representing the conditional expectations $E[y_t | Y_{k-1}]$, $t \in [k, T)$. We will refer to (16) as $Eq(k, \hat{X}_{k|k-1})$. Evidently, the existence of equilibrium environment variables $(y_t)_{t \in [0, T)}$ requires solutions $\hat{y}_{t|k-1}$ to $Eq(k, E[X_k | Y_{k-1}])$ for all realizations of Y_{k-1} that can be generated by (14)-(15).

Assumption 2. For each $k \in [0, T)$ and arbitrary $\hat{X}_{k|k-1}$, the linear equations $Eq(k, \hat{X}_{k|k-1})$ in (16) has a unique solution. We note that such a solution is necessarily in the form

$$\hat{y}_{t|k-1} = a_{t,k-1} \hat{X}_{k|k-1} + b_{t,k-1}, \quad t \in [k, T),$$

where $(a_{t,k-1})_{t \in [k, T)}$ and $(b_{t,k-1})_{t \in [k, T)}$ depend only on the constants $\mathcal{D}_t, \mathcal{G}_{t,n}^p, \mathcal{H}_t^p, \hat{\xi}_t, \mathcal{A}_t, \mathcal{G}_{t,n}^X, \mathcal{H}_t^X, C_t, \hat{W}_t, 0 \leq k \leq t \leq n < T$.

Assumption 2 is satisfied, for example, when the matrices $E_t^{1,j}, E_t^{2,j}, B_t^{1,j}, B_t^{2,j}$, for all $j \in [1, N]$, $t \in [0, T)$, are sufficiently small in size, i.e., when the sensitivity of the environment variables to the states and decisions of the DMs is sufficiently small. Note that the linear equations $Eq(k, \hat{X}_{k|k-1})$, for any $k \in [0, T)$, can be written as:

$$\begin{bmatrix} \hat{y}_{k|k-1} \\ \vdots \\ \hat{y}_{T-1|k-1} \end{bmatrix} = \Lambda_k \begin{bmatrix} \hat{y}_{k|k-1} \\ \vdots \\ \hat{y}_{T-1|k-1} \end{bmatrix} + \Upsilon_k \hat{X}_{k|k-1} + \Theta_k,$$

where $\Lambda_k, \Upsilon_k, \Theta_k$ are constants of appropriate dimension, which depend on $\mathcal{D}_t, \mathcal{G}_{t,n}^p, \mathcal{H}_t^p, \hat{\xi}_t, \mathcal{A}_t, \mathcal{G}_{t,n}^X, \mathcal{H}_t^X, C_t, \hat{W}_t, 0 \leq k \leq t \leq n < T$. Furthermore, the matrices $F_t^j, G_{t,k}^j, H_t^j$ defining

each DM^j's optimal controller are independent of $E_t^{1,j}$, $E_t^{2,j}$, $B_t^{1,j}$, $B_t^{2,j}$. Accordingly, when the matrices $E_t^{1,j}$, $E_t^{2,j}$, $B_t^{1,j}$, $B_t^{2,j}$, for all $j \in [1, N]$, $t \in [0, T)$, are sufficiently small in size, $I - \Lambda_k$ is invertible.

Under Assumption 1 and 2, it is natural to introduce the following candidate equilibrium environment variables, which is an independent exogenous sequence $(\bar{y}_t)_{t \in [0, T)}$ generated by the recursion, for $t \in [0, T)$,

$$\bar{y}_t = \mathcal{D}_t \bar{X}_t + \sum_{n \in [t, T)} \mathcal{G}_{t,n}^p(a_{n,t-1} E[\bar{X}_t | \bar{Y}_{t-1}] + b_{n,t-1}) + \mathcal{H}_t^p + \xi_t, \quad (17)$$

$$\bar{X}_{t+1} = \mathcal{A}_t \bar{X}_t + \sum_{n \in [t, T)} \mathcal{G}_{t,n}^X(a_{n,t-1} E[\bar{X}_t | \bar{Y}_{t-1}] + b_{n,t-1}) + \mathcal{H}_t^X + \mathcal{C}_t \bar{y}_t + W_t, \quad (18)$$

where $(a_{n,t-1})_{n \in [t, T)}$ and $(b_{n,t-1})_{n \in [t, T)}$ are as in Assumption 2, $\bar{Y}_{t-1} = (\bar{y}_k)_{k \in [0, t)}$, for $t \in [0, T)$, and $\bar{X}_0 = X_0$. It can be shown that the random sequence $(\bar{y}_t)_{t \in [0, T)}$ generated by (17)-(18) is indeed a sequence of equilibrium environment variables. This leads to the existence of the SEBEU.

Theorem 1. *Consider the finite-horizon linear quadratic problem of this subsection. If Assumption 1 and 2 hold, there exists an SEBEU strategy $s \in \mathbb{S}$ defined by:*

$$s_t^i(I_t^i, Y_{t-1}) = F_t^i x_t^i + G_t^i \hat{X}_{t|t-1}^i(Y_{t-1}) + \check{H}_t^i, \quad i \in [1, N], \quad t \in [0, T), \quad (19)$$

where $G_t^i := \sum_{k \in [t, T)} G_{t,k}^i a_{k,t-1}$, $\check{H}_t^i := \sum_{k \in [t, T)} G_{t,k}^i b_{k,t-1} + H_t^i$, and $\hat{X}_{t|t-1}^i(\cdot)$ denotes the state estimator mapping from \bar{Y}_{t-1} to $E[\bar{X}_t | \bar{Y}_{t-1}]$ based on (17)-(18).

Proof. Consider any $0 \leq k \leq \ell < T$ and arbitrary $\hat{X}_{k|k-1}$. The equations $Eq(k, \hat{X}_{k|k-1})$ for $t \in [k, T)$ coincide with the equations $Eq(\ell, \hat{X}_{\ell|k-1})$, where $\hat{X}_{\ell|k-1}$ is obtained by propagating $\hat{X}_{k|k-1}$ through (16). By Assumption 2, this implies that, for $t \in [k, T)$,

$$a_{t,k-1} \hat{X}_{k|k-1} + b_{t,k-1} = a_{t,\ell-1} \hat{X}_{\ell|k-1} + b_{t,\ell-1}. \quad (20)$$

By Assumption 2 and substituting (20) into $Eq(k, \hat{X}_{k|k-1})$, we obtain, for $t \in [k, T)$,

$$\hat{y}_{t|k-1} = \mathcal{D}_t \hat{X}_{t|k-1} + \sum_{n \in [t, T)} \mathcal{G}_{t,n}^p(a_{n,t-1} \hat{X}_{t|k-1} + b_{n,t-1}) + \mathcal{H}_t^p + \hat{\xi}_t, \quad (21)$$

$$\hat{X}_{t+1|k-1} = \mathcal{A}_t \hat{X}_{t|k-1} + \sum_{n \in [t, T)} \mathcal{G}_{t,n}^X(a_{n,t-1} \hat{X}_{t|k-1} + b_{n,t-1}) + \mathcal{H}_t^X + \mathcal{C}_t \hat{y}_{t|k-1} + \hat{W}_t, \quad (22)$$

where $\hat{y}_{t,k-1} = a_{t,k-1} \hat{X}_{k|k-1} + b_{t,k-1}$. Taking conditional expectations of both sides of (17)-(18) given \bar{Y}_{k-1} leads to, for $t \in [k, T)$,

$$\begin{aligned} E[\bar{y}_t | \bar{Y}_{k-1}] &= \mathcal{D}_t E[\bar{X}_t | \bar{Y}_{k-1}] + \sum_{n \in [t, T)} \mathcal{G}_{t,n}^p(a_{n,t-1} E[\bar{X}_t | \bar{Y}_{k-1}] \\ &\quad + b_{n,t-1}) + \mathcal{H}_t^p + \hat{\xi}_t, \end{aligned} \quad (23)$$

$$\begin{aligned} E[\bar{X}_{t+1} | \bar{Y}_{k-1}] &= \mathcal{A}_t E[\bar{X}_t | \bar{Y}_{k-1}] + \sum_{n \in [t, T)} \mathcal{G}_{t,n}^X(a_{n,t-1} E[\bar{X}_t | \bar{Y}_{k-1}] + b_{n,t-1}) \\ &\quad + \mathcal{H}_t^X + \mathcal{C}_t E[\bar{y}_t | \bar{Y}_{k-1}] + \hat{W}_t. \end{aligned} \quad (24)$$

The equations (21)-(22) hold, in particular, for $\hat{X}_{k|k-1} = E[\bar{X}_k | \bar{Y}_{k-1}]$, in which case $\hat{y}_{t,k-1} = a_{t,k-1} E[\bar{X}_k | \bar{Y}_{k-1}] + b_{t,k-1}$. Therefore, setting $\hat{X}_{k|k-1} = E[\bar{X}_k | \bar{Y}_{k-1}]$ in (21)-(22) and comparing

the right-hand-sides of (21)-(22) and (23)-(24) for $t = k$ yields

$$E[\bar{y}_k|\bar{Y}_{k-1}] = a_{k,k-1}E[\bar{X}_k|\bar{Y}_{k-1}] + b_{k,k-1} \quad \text{and} \quad E[\bar{X}_{k+1}|\bar{Y}_{k-1}] = \hat{X}_{k+1|k-1}.$$

Using $E[\bar{X}_{k+1}|\bar{Y}_{k-1}] = \hat{X}_{k+1|k-1}$ and comparing the right-hand-sides of (21)-(22) and (23)-(24) for $t = k + 1$ yields

$$E[\bar{y}_{k+1}|\bar{Y}_{k-1}] = a_{k+1,k-1}E[\bar{X}_k|\bar{Y}_{k-1}] + b_{k+1,k-1} \quad \text{and} \quad E[\bar{X}_{k+2}|\bar{Y}_{k-1}] = \hat{X}_{k+2|k-1}.$$

Using $E[\bar{X}_{k+2}|\bar{Y}_{k-1}] = \hat{X}_{k+2|k-1}$ and continuing in a similar fashion yields

$$E[\bar{y}_t|\bar{Y}_{k-1}] = a_{t,k-1}E[\bar{X}_k|\bar{Y}_{k-1}] + b_{t,k-1}, \quad t \in [k, T].$$

Using this in (17)-(18) shows that the environment variables $(\bar{y}_t)_{t \in [0, T]}$ generated by (17)-(18) satisfy (14)-(15). Furthermore, due to Assumption 2, $(\bar{y}_t)_{t \in [0, T]}$ have finite second-order moments. This implies that $(\bar{y}_t)_{t \in [0, T]}$ are equilibrium environment variables, and the strategy defined by (19) is an SEBEU. \square

We remark that Assumption 2 requires the existence of a unique solution to $Eq(k, \hat{X}_{k|k-1})$ for arbitrary $\hat{X}_{k|k-1}$ which is in general not necessary for the existence of equilibrium environment variables. Assumption 2 can be relaxed so that $Eq(k, \hat{X}_{k|k-1})$ has a solution only for those $\hat{X}_{k|k-1} = E[\bar{X}_k|\bar{Y}_{k-1}]$ generated by the dynamics (17)-(18). However, by also noting that an SEBEU must have the form (10) except on a set of observation histories of probability zero, we conclude that Assumption 2 cannot be relaxed in a significant way.

At an SEBEU, DMs model the environment variables as an independent exogenous stochastic process as in (17)-(18) where \bar{X}_t is a replica of the system state X_t . This model for the environment variables and the resulting distribution of the environment variables is consistent with the actual environment variables generated endogenously by the equilibrium strategies. Therefore, one can start with such a model for the environment variables (instead of an arbitrary distribution for the environment variables) and establish the existence of an SEBEU in an alternative fashion.

In the next subsection, we extend this result to infinite-horizon linear quadratic problems.

3.2. Infinite-horizon case ($T = \infty$). In this subsection, we further assume that the system is time-invariant. Hence, the matrices $A_t^i, B_t^i, C_t^i, A_t^0, B_t^{1,i}, B_t^{2,i}, D_t, E_t^{1,i}, E_t^{2,i}, Q_t^i, R_t^i, K_t^i, L_t^i$ are all assumed to be independent of the time variable t (hence the subscript t is dropped from these matrices). Recall that $\beta^i \in [0, 1)$, for all $i \in [1, N]$, in the infinite-horizon case.

Assumption 3.

- (i) A^0 is stable, (A^i, B^i) is stabilizable, for all $i \in [1, N]$
- (ii) $X_0, W_0, W_1, \dots, \xi_0, \xi_1, \dots$ are mutually independent and have finite second-order moments
- (iii) $(W_t)_{t \in \mathbb{N}_0}, (\xi_t)_{t \in \mathbb{N}_0}$ are each independent and identically distributed (iid).

Consider an independent exogenous sequence $(z_t)_{t \in [0, \infty)}$ of environment variables with a given probability distribution $\zeta \in \mathcal{P}(\mathbb{Y}^{[0, \infty)})$ satisfying $\sup_{t \in [0, \infty)} E[|z_t|^2] < \infty$. An optimal strategy minimizing $E^\zeta[J^i(s^i, \cdot)]$ over $s^i \in \mathbb{S}^i$ is given by:

$$u_t^i = s_t^i(I_t^i, Z_{t-1}) = F^i x_t^i + \sum_{n \in \mathbb{N}_0} G_n^i E[z_{t+n}|Z_{t-1}] + H^i, \quad t \in \mathbb{N}_0,$$

where $F^i, (G_n^i)_{n \in \mathbb{N}_0}, H^i$ are constants of appropriate dimensions, which are pre-computable independently of the sequence $(z_t)_{t \in [0, \infty)}$ of environment variables; see Appendix A.2.

Consider now a joint strategy $s = (s^i)_{i \in [1, N]} \in \mathbb{S}$ where the sequence $(y_t)_{t \in [0, \infty)}$ of environment variables generated under s has the probability distribution ζ_s and $\sup_{t \in [0, \infty)} E[|y_t|^2] < \infty$. The joint strategy s is an SEBEU if, for each $i \in [1, N]$, s^i minimizes $E^{\zeta_s}[J^i(s^i, \cdot)]$, e.g.,

$$s_t^i(I_t^i, Y_{t-1}) = F^i x_t^i + \sum_{n \in \mathbb{N}_0} G_n^i E[y_{t+n} | Y_{t-1}] + H^i, \quad t \in \mathbb{N}_0, \tag{25}$$

where

$$y_t = D x_t^0 + \sum_{j \in [1, N]} E^{1,j} \left(F^j x_t^j + \sum_{n \in \mathbb{N}_0} G_n^j E[y_{t+n} | Y_{t-1}] + H^j \right) + \sum_{j \in [1, N]} E^{2,j} x_t^j + \xi_t, \tag{26}$$

$$x_{t+1}^0 = A^0 x_t^0 + \sum_{j \in [1, N]} B^{1,j} \left(F^j x_t^j + \sum_{n \in \mathbb{N}_0} G_n^j E[y_{t+n} | Y_{t-1}] + H^j \right) + \sum_{j \in [1, N]} B^{2,j} x_t^j + w_t^0, \tag{27}$$

$$x_{t+1}^j = A^j x_t^j + B^j \left(F^j x_t^j + \sum_{n \in \mathbb{N}_0} G_n^j E[y_{t+n} | Y_{t-1}] + H^j \right) + C^j y_t + w_t^j, \quad j \in [1, N]. \tag{28}$$

Therefore, an SEBEU can be obtained via (25) provided a sequence $(y_t)_{t \in [0, \infty)}$ of environment variables satisfying (26)-(28) and $\sup_{t \in [0, \infty)} E[|y_t|^2] < \infty$, called equilibrium environment variables, exists. Next, we introduce a set of assumptions to ensure the existence of equilibrium environment variables.

We rewrite (26)-(28) in the vector form as:

$$y_t = \mathcal{D} X_t + \sum_{n \in \mathbb{N}_0} \mathcal{G}_n^p E[y_{t+n} | Y_{t-1}] + \mathcal{H}^p + \xi_t, \tag{29}$$

$$X_{t+1} = \mathcal{A} X_t + \sum_{n \in \mathbb{N}_0} \mathcal{G}_n^x E[y_{t+n} | Y_{t-1}] + \mathcal{H}^x + \mathcal{C} y_t + W_t, \tag{30}$$

where \mathcal{D} , $(\mathcal{G}_n^p)_{n \in \mathbb{N}_0}$, \mathcal{H}^p , \mathcal{A} , $(\mathcal{G}_n^x)_{n \in \mathbb{N}_0}$, \mathcal{H}^x , \mathcal{C} are some constants of appropriate dimensions. Taking the expectation of both sides of (29)-(30) shows that the existence of equilibrium environment variables requires a solution $(\hat{y}_t)_{t \in \mathbb{N}_0}$ to the infinite system of linear equations, for $t \in \mathbb{N}_0$,

$$\hat{y}_t = \mathcal{D} \hat{X}_t + \sum_{n \in \mathbb{N}_0} \mathcal{G}_n^p \hat{y}_{t+n} + \mathcal{H}^p + \hat{\xi}_0, \tag{31}$$

$$\hat{X}_{t+1} = \mathcal{A} \hat{X}_t + \sum_{n \in \mathbb{N}_0} \mathcal{G}_n^x \hat{y}_{t+n} + \mathcal{H}^x + \mathcal{C} \hat{y}_t + \hat{W}_0, \tag{32}$$

where $\hat{X}_0 = E[X_0]$. Accordingly, we make the following assumption.

Assumption 4. *The infinite system of linear equations (31)-(32) has a unique solution $(\hat{y}_t)_{t \in \mathbb{N}_0}$ for any \hat{X}_0 , which will necessarily be in the form*

$$\hat{y}_t = a_t \hat{X}_0 + b_t, \quad t \in \mathbb{N}_0, \tag{33}$$

where $(a_t)_{t \in \mathbb{N}_0}$, $(b_t)_{t \in \mathbb{N}_0}$ are assumed to be uniformly bounded and depend only on the constants \mathcal{D} , $(\mathcal{G}_n^p)_{n \in \mathbb{N}_0}$, \mathcal{H}^p , $\hat{\xi}_0$, \mathcal{A} , $(\mathcal{G}_n^x)_{n \in \mathbb{N}_0}$, \mathcal{H}^x , \mathcal{C} , \hat{W}_0 .

For any $k \in \mathbb{N}_0$, taking the conditional expectation of both sides of (29)-(30) given Y_{k-1} , where $t \in [k, \infty)$, also leads to (31)-(32) with \hat{X}_0 replaced with $E[X_k | Y_{k-1}]$. Therefore, under

Assumption 4, the conditional expectations of equilibrium environment variables satisfying (29)-(30), if exist, are given by:

$$E[y_{n+k}|Y_{k-1}] = a_n E[X_k|Y_{k-1}] + b_n, \quad n \in \mathbb{N}_0, k \in \mathbb{N}_0. \tag{34}$$

This prompts us to introduce the independent exogenous environment variables $(\bar{y}_t)_{t \in [0, \infty)}$ generated by the recursion, for $t \in \mathbb{N}_0$,

$$\bar{y}_t = \mathcal{D}\bar{X}_t + \mathcal{G}^p E[\bar{X}_t|\bar{Y}_{t-1}] + \check{\mathcal{H}}^p + \bar{\xi}_t, \tag{35}$$

$$\bar{X}_{t+1} = \mathcal{A}\bar{X}_t + \mathcal{G}^X E[\bar{X}_t|\bar{Y}_{t-1}] + \check{\mathcal{H}}^X + \mathcal{C}\bar{y}_t + \bar{W}_t, \tag{36}$$

where $\mathcal{G}^p := \sum_{n \in \mathbb{N}_0} \mathcal{G}_n^p a_n$, $\check{\mathcal{H}}^p := \sum_{n \in \mathbb{N}_0} \mathcal{G}_n^p b_n + \mathcal{H}^p$, $\mathcal{G}^X := \sum_{n \in \mathbb{N}_0} \mathcal{G}_n^X a_n$, $\check{\mathcal{H}}^X := \sum_{n \in \mathbb{N}_0} \mathcal{G}_n^X b_n + \mathcal{H}^X$,

the distribution of $(\bar{X}_0, (\bar{\xi}_t)_{t \in \mathbb{N}_0}, (\bar{W}_t)_{t \in \mathbb{N}_0})$ is identical to that of $(X_0, (\xi_t)_{t \in \mathbb{N}_0}, (W_t)_{t \in \mathbb{N}_0})$, and $\bar{Y}_{t-1} = (\bar{y}_k)_{k \in [0, t)}$. By construction, the environment variables $(\bar{y}_t)_{t \in [0, \infty)}$ generated by (35)-(36) satisfy (29)-(30); hence the environment variables $(\bar{y}_t)_{t \in [0, \infty)}$ are equilibrium environment variables, provided $\sup_{t \in [0, \infty)} E[|\bar{y}_t|^2] < \infty$.

Theorem 2. Consider the infinite-horizon linear quadratic problem of this subsection under Assumption 3-4. If $(\bar{y}_t)_{t \in [0, \infty)}$ generated by (35)-(36) satisfies $\sup_{t \in [0, \infty)} E[|\bar{y}_t|^2] < \infty$, there exists

an SEBEU strategy $s \in \mathbb{S}$ defined by

$$s_t^i(I_t^i, Y_{t-1}) = F^i x_t^i + G^i \hat{X}_{t|t-1}^i(Y_{t-1}) + \check{H}^i, \quad i \in [1, N], t \in \mathbb{N}_0, \tag{37}$$

where $G^i := \sum_{n \in \mathbb{N}_0} G_n^i a_n$, $\check{H}^i := \sum_{n \in \mathbb{N}_0} G_n^i b_n + H^i$, and $\hat{X}_{t|t-1}^i(\cdot)$ denotes the state estimator mapping $\bar{Y}_{t-1} \mapsto E[\bar{X}_t|\bar{Y}_{t-1}]$ based on (35)-(36).

Proof. The assumptions and the discussion preceding the theorem ensures that the environment variables generated by the joint strategy s defined by (37) are equilibrium environment variables (which are distributed identically to the independent exogenous environment variables generated by (35)-(36)). \square

Theorem 2 constructs an SEBEU under the unique solvability of (31)-(32) for any \hat{X}_0 and the uniform boundedness of the second moments of the environment variables. These conditions cannot be relaxed as long as we insist on the existence of an equilibrium (with environment variables having uniformly bounded second-order moments) for every initial state distribution. The uniform boundedness of the second moments of the environment variables generated under (37) can be ensured if the closed-loop dynamics (identical to (35)-(36)) are stable. The closed-loop stability will be discussed in more detail in the Gaussian case shortly.

3.3. Gaussian case. We now turn our attention to the Gaussian case where all primitive random variables are jointly Gaussian. With this Gaussian assumption, in both finite-horizon and infinite-horizon cases, the state estimator $\hat{X}_{t|t-1}^i(\cdot)$ mapping $\bar{Y}_{t-1} \mapsto E[\bar{X}_t|\bar{Y}_{t-1}]$ can be implemented as the following Kalman filter, for $t \in [0, T)$,

$$\begin{aligned} \hat{X}_{t+1|t} &= \mathcal{A}_t \hat{X}_{t|t-1} + \mathcal{G}_t^X \hat{X}_{t|t-1} + \check{\mathcal{H}}_t^X + \mathcal{C}_t \bar{y}_t + \hat{W}_t \\ &\quad + \mathcal{A}_t \Sigma_{t|t-1} \mathcal{D}'_t (\mathcal{D}_t \Sigma_{t|t-1} \mathcal{D}'_t + \text{cov}[\xi_t])^{-1} (\bar{y}_t - (\mathcal{D}_t + \mathcal{G}_t^p) \hat{X}_{t|t-1} - \check{\mathcal{H}}_t^p - \hat{\xi}_t), \end{aligned} \tag{38}$$

$$\Sigma_{t+1|t} = \mathcal{A}_t (\Sigma_{t|t-1} - \Sigma_{t|t-1} \mathcal{D}'_t (\mathcal{D}_t \Sigma_{t|t-1} \mathcal{D}'_t + \text{cov}[\xi_t])^{-1} \mathcal{D}_t \Sigma_{t|t-1}) \mathcal{A}'_t + \text{cov}[W_t], \tag{39}$$

where $\bar{X}_{0|-1} = E[X_0]$, $\Sigma_{0|-1} := \text{cov}[\bar{X}_0]$ (assuming $\mathcal{D}_t \Sigma_{t|t-1} \mathcal{D}'_t + \text{cov}[\xi_t]$ is invertible for $t \in [0, T)$). Therefore, under the Gaussian assumption, the players' decisions at an SEBEU can be generated by a linear time-varying system in both finite-horizon and infinite-horizon cases.

In the case of infinite-horizon, it is possible to obtain an SEBEU at which players' decisions can be generated by a linear time-invariant system. We recall that, in this case, $\mathcal{A}_t, \mathcal{C}_t, \mathcal{D}_t, \mathcal{G}_t^X, \tilde{\mathcal{H}}_t^X, \mathcal{G}_t^p, \tilde{\mathcal{H}}_t^p, \hat{W}_t, \text{cov}[W_t], \hat{\xi}_t, \text{cov}[\xi_t]$ are independent of t . Hence, a time-invariant SEBEU can be obtained if $\text{cov}[X_0] = \Sigma_{0|0} = \Sigma$ where Σ satisfies

$$\Sigma = \mathcal{A}(\Sigma - \Sigma \mathcal{D}'(\mathcal{D}\Sigma\mathcal{D}' + \text{cov}[\xi_0])^{-1}\mathcal{D}\Sigma)\mathcal{A}' + \text{cov}[W_0]. \tag{40}$$

This ensures $\Sigma_{t|t-1} = \Sigma$, for all $t \in \mathbb{N}_0$ and leads to the following corollary to Theorem 2.

Corollary 1. *Consider the infinite-horizon linear quadratic problem of subsection 3.2 under the assumptions of Theorem 2. Assume further that (i) $(X_0, (\xi_t)_{t \in \mathbb{N}_0}, (W_t)_{t \in \mathbb{N}_0})$ is Gaussian, (ii) $\mathcal{D}\Sigma\mathcal{D}' + \text{cov}[\xi_0]$ is invertible, and (iii) $\text{cov}[X_0] = \Sigma$ where Σ solves (40). The SEBEU strategy s in Theorem 2 can be implemented by the linear-time invariant system: $i \in [1, N], t \in \mathbb{N}_0$,*

$$\begin{aligned} s_t^i(I_t^i, Y_{t-1}) &= F^i x_t^i + G^i \hat{X}_{t|t-1} + \tilde{H}^i, \\ \hat{X}_{t+1|t} &= \mathcal{A}\hat{X}_{t|t-1} + \mathcal{G}^X \hat{X}_{t|t-1} + \tilde{\mathcal{H}}^X + \mathcal{C}y_t + \hat{W}_0, \\ &\quad + \mathcal{A}\Sigma\mathcal{D}'(\mathcal{D}\Sigma\mathcal{D}' + \text{cov}[\xi_0])^{-1}(y_t - (\mathcal{D} + \mathcal{G}^p)\hat{X}_{t|t-1} - \tilde{\mathcal{H}}^p - \hat{\xi}_0). \end{aligned}$$

(provided $(y_t)_{t \in [0, \infty)}$ generated under s satisfies $\sup_{t \in [0, \infty)} E[|y_t|^2] < \infty$).

Implicit in the steady-state initialization condition $\text{cov}[X_0] = \Sigma$ is that there exist $\Sigma \succeq 0$ solving (40), which can be established under reasonable assumptions ensuring closed-loop stability. Note that, since A^0 is stable, \mathcal{A} would be stable if $A^i + B^i F^i, i \in [1, N]$ are stable, for all $i \in [1, N]$. In other words, \mathcal{A} would be stable if each DM's own state dynamics (as the environment variables are seen as external inputs) is stabilized by its own control strategy. However, the internal dynamics of the environment variables and each DM's own state dynamics are coupled in a feedback loop. The strategies of DMs designed to optimize their individual objectives by viewing the process of environment variables as independent and exogenous cannot be expected to stabilize this interconnected feedback system. The overall stability of such an interconnected system (of stable subsystems) can be guaranteed under a condition of weak coupling, i.e., the coefficients $C^1, \dots, C^N, K^1, \dots, K^N, L^1, \dots, L^N$ are sufficiently small, which makes $\mathcal{C}, \mathcal{G}^X$ sufficiently small. The smallness of these coefficients de-emphasizes the role of the environment variables on the state dynamics as well as the cost functions of all DMs. This weak coupling condition along with the conditions for the stability of \mathcal{A} guarantees the existence of $\Sigma \succeq 0$ solving (40).

The condition $\text{cov}[X_0] = \Sigma$ is clearly a strong one. A possible alternative setting where steady-state initialization is not required would be the average cost setting.

3.4. Case of identical DMs. We formally introduce the notion of identical DMs as follows.

Definition 2. DM^{*i*} and DM^{*j*} are said to be identical if the following two conditions hold.

- (i) $(A_t^i, B_t^i, C_t^i, \beta^i, Q_t^i, R_t^i, K_t^i, L_t^i) = (A_t^j, B_t^j, C_t^j, \beta^j, Q_t^j, R_t^j, K_t^j, L_t^j)$, for all $t \in [0, T)$
- (ii) $(x_0^i, (w_t^i)_{t \in [0, T)})$ and $(x_0^j, (w_t^j)_{t \in [0, T)})$ have identical distributions.

When DM^{*i*} and DM^{*j*} are identical, $E^{\zeta^i}[J^i(s^i, \cdot)] = E^{\zeta^j}[J^j(s^j, \cdot)]$ for any $(s^i, \zeta^i) = (s^j, \zeta^j)$. We note that, due to the assumption $R_t^i \succ 0$ for all $t \in [0, T)$, a DM's optimal strategy with respect to any belief $\zeta \in \mathcal{P}(\mathbb{Y}^{[0, T)})$ is essentially unique, i.e., it can differ from that derived in Appendix only on a set of observation histories of probability zero. Therefore, the optimal strategies of identical DMs with respect to any common belief $\zeta \in \mathcal{P}(\mathbb{Y}^{[0, T)})$ can differ only on a set of observation histories of probability zero. As a result, the strategies of identical DMs at an SEBEU (if exists) can be taken as identical.

We should also point out that, when DM^i and DM^j are identical, the estimates $E[x_k^i|Y_{k-1}]$ and $E[x_k^j|Y_{k-1}]$ generated by an SEBEU would be identical as well. Therefore, when DM^i and DM^j are identical, DMs can implement their SEBEU strategies by generating only one of $E[x_k^i|Y_{k-1}]$ or $E[x_k^j|Y_{k-1}]$. For example, when all DMs are identical, the SEBEU strategies can be implemented by generating only $E[x_k^i|Y_{k-1}]$, for a generic DM^i , instead of generating the estimates $E[X_k|Y_{k-1}]$ of the entire system state.

3.5. Case of countably infinite set of identical DMs. We adopt the infinite-horizon setup of subsection 3.2 with a countable set $(DM^i)_{i \in \mathbb{N}}$ of identical DMs. The matrices $A^i, B^i, Q^i, R^i, K^i, L^i$ are all assumed to be independent of the index i (hence the superscript i is dropped from these matrices), and the matrices C^i are assumed to satisfy $C^i = 0$, for all $i \in \mathbb{N}$, i.e., the individual state dynamics are not directly controlled through the environment variables. Furthermore, the environment variable at each time $t \in \mathbb{N}_0$ is determined by the empirical average of the decisions as well as the states:

$$y_t = \limsup_{\bar{N} \rightarrow \infty} \frac{1}{\bar{N}} \sum_{j \in [1, \bar{N}]} (E^1 u_t^j + E^2 x_t^j) + \xi_t,$$

where E^1, E^2 are constants of appropriate dimension. In addition, we strengthen Assumption 3 as follows.

Assumption 5.

- (i) (A, B) is stabilizable
- (ii) $x_0^1, x_0^2, \dots, w_0^1, w_0^2, \dots, w_1^1, w_1^2, \dots, \xi_0, \xi_1, \dots$ are mutually independent and have finite second-order moments
- (iii) $(x_0^i)_{i \in \mathbb{N}}, (w_t^i)_{i \in \mathbb{N}, t \in \mathbb{N}_0}, (\xi_t)_{t \in \mathbb{N}_0}$ each is iid.

Consider an independent exogenous sequence $(z_t)_{t \in [0, \infty)}$ of environment variables which is iid with finite second-order moments and distribution ζ . As in the previous cases, suppose that each DM^i instead aims to minimize $E^\zeta[J^i(s^i, \cdot)]$ over $s^i \in \mathbb{S}^i$. An optimal strategy for each DM^i is obtained as:

$$u_t^i = s_t^i(I_t^i, Z_{t-1}) = Fx_t^i + GE[z_0] + H, \tag{41}$$

where $F, G := \sum_{n \in \mathbb{N}_0} G_n, H$ are constants of appropriate dimensions, which are pre-computable independently of the sequence $(z_t)_{t \in [0, \infty)}$ of environment variables; see Appendix A.2.

Consider now a joint strategy $s = (s^i)_{i \in \mathbb{N}} \in \mathbb{S}$ such that the sequence $(y_t)_{t \in [0, \infty)}$ of environment variables generated under s is iid with distribution ζ_s ($E[|y_t|^2] = E[|y_0|^2] < \infty$, for all $t \in \mathbb{N}_0$). The joint strategy s is an SEBEU if each s^i minimizes $E^{\zeta_s}[J^i(s^i, \cdot)]$, e.g.,

$$s_t^i(I_t^i, Y_{t-1}) = Fx_t^i + GE[y_0] + H, \tag{42}$$

where

$$y_t = (E^1 F + E^2) \limsup_{\bar{N} \rightarrow \infty} \frac{1}{\bar{N}} \sum_{j \in [1, \bar{N}]} x_t^j + E^1 GE[y_0] + E^1 H + \xi_t \tag{43}$$

$$x_{t+1}^j = Ax_t^j + B(Fx_t^j + GE[y_0] + H) + w_t^j, \quad j \in \mathbb{N}. \tag{44}$$

Therefore, an SEBEU can be constructed via (42) provided an iid sequence $(y_t)_{t \in [0, \infty)}$ of environment variables satisfying (43)-(44) (as a random sequence) can be obtained. Due to Assumption 5, $(x_t^i)_{i \in \mathbb{N}}$ generated by (44) would be iid with finite second-order moments, for any $t \in \mathbb{N}_0$. As a result, we would have $\limsup_{\bar{N} \rightarrow \infty} \frac{1}{\bar{N}} \sum_{j \in [1, \bar{N}]} x_t^j = E[x_t^1]$, for any $t \in \mathbb{N}_0$, and

$$y_t = (E^1 F + E^2)E[x_t^1] + E^1 GE[y_0] + E^1 H + \xi_t, \\ E[x_{t+1}^1] = AE[x_t^1] + B(FE[x_t^1] + GE[y_0] + H) + E[w_t^1].$$

Thus, if the linear equations

$$\hat{y}_0 = (E^1 F + E^2)\hat{x}_0 + E^1 G \hat{y}_0 + E^1 H + \hat{\xi}_0, \tag{45}$$

$$\hat{x}_0 = A\hat{x}_0 + B(F\hat{x}_0 + G\hat{y}_0 + H) + \hat{w}_0^1. \tag{46}$$

have a solution (\hat{y}_0, \hat{x}_0) and $(E[y_0], E[x_0^1]) = (\hat{y}_0, \hat{x}_0)$, then the environment variables generated by the strategies

$$u_t^i = s_t^i(I_t^i, Y_{t-1}) = Fx_t^i + G\hat{y}_0 + H, \quad i \in \mathbb{N} \tag{47}$$

would be iid with finite second-order moments and satisfy (43)-(44). This leads to the following result whose proof follows from the preceding discussion.

Theorem 3. *Consider the infinite-horizon linear quadratic problem of subsection 3.2 with a countably infinite set of identical DMs under Assumption 5. If $(E[y_0], E[x_0^1]) = (\hat{y}_0, \hat{x}_0)$ where (\hat{y}_0, \hat{x}_0) solves (45)-(46), there exists an SEBEU strategy $s \in \mathbb{S}$ defined by (47).*

In the case where the environment variables are determined by the average decisions of a countable set of identical DMs, no individual DM can change the sequence of environment variables by unilaterally deviating to an alternative strategy; therefore, an SEBEU is also a Nash equilibrium, which is in fact a mean-field equilibrium [15, 17, 19].

4. AN EXAMPLE

We present an example where $T = 2$, all DMs are identical, and the environment variables are generated by:

$$y_t = \frac{1}{N} \underbrace{\sum_{i \in [1, N]} x_t^i}_{:= \bar{x}_t} + \xi_t, \quad t \in [0, 2).$$

We also assume, in addition to Assumption 1, that all primitive random variables

$$x_0^1, \dots, x_0^N, w_0^1, w_0^N, \dots, w_1^1, \dots, w_1^N, \xi_0, \xi_1$$

are mutually independent Gaussian random variables with zero-mean and $\text{cov}[\xi_0] > 0$.

In this case, the SEBEU established by Theorem 1 takes the form, for each $i \in [1, N]$,

$$\begin{aligned} s_0^i(x_0^i) &= F_0 x_0^i + G_{0,0} \hat{x}_0 + G_{0,1} \hat{x}_1, \\ s_1^i(x_1^i, y_0) &= F_1 x_1^i + G_{1,1} \hat{x}_{1|0}(y_0), \end{aligned}$$

where $F_0, F_1, G_{0,0}, G_{0,1}, G_{1,1}$, are given in Appendix A.1 and

$$\begin{aligned} \hat{x}_0 &= E[\bar{x}_0], \\ \hat{x}_1 &= (A_0 + B_0 F_0) \hat{x}_0 + B_0 (G_{0,0} \hat{x}_0 + G_{0,1} \hat{x}_1) + C_0 \hat{x}_0, \\ \hat{x}_{1|0}(y_0) &= (A_0 + B_0 F_0) \hat{x}_{0|0}(y_0) + B_0 (G_{0,0} \hat{x}_0 + G_{0,1} \hat{x}_1) + C_0 y_0, \\ \hat{x}_{0|0}(y_0) &= E[\bar{x}_0 | y_0] = \text{cov}[\bar{x}_0] (\text{cov}[\bar{x}_0] + \text{cov}[\xi_0])^{-1} y_0. \end{aligned}$$

Since $E[\bar{x}_0] = 0$ (by assumption) and

$$(I - B_0 G_{0,1}) \hat{x}_1 = (A_0 + B_0 F_0 + B_0 G_{0,0} + C_0) E[\bar{x}_0] \tag{48}$$

an SEBEU exists corresponding to any \hat{x}_1 in the null space of $I - B_0 G_{0,1}$. On the other hand, if \bar{x}_0 was not assumed to be zero-mean and (48) does not have a solution for \hat{x}_1 , an SEBEU does not exist.

Let us now consider the case where $I - B_0G_{0,1}$ is non-singular (and $E[\bar{x}_0] = 0$ as assumed). The SEBEU, which is essentially unique, is

$$s_0^i(x_0^i) = F_0x_0^i, \tag{49}$$

$$s_1^i(x_1^i, y_0) = F_1x_1^i + G_1^N y_0, \tag{50}$$

where

$$G_1^N = G_{1,1}((A_0 + B_0F_0)\text{cov}[\bar{x}_0](\text{cov}[\bar{x}_0] + \text{cov}[\xi_0])^{-1} + C_0).$$

Due to $\text{cov}[\bar{x}_0] = \frac{1}{N}\text{cov}[x_0^1]$, we have $\lim_{N \rightarrow \infty} G_1^N = G_{1,1}C_0$; the state feedback gains F_0, F_1 are independent of the number of DMs. Obtaining the set of Nash equilibria is not as straightforward; however, it is possible to show that, for any $\epsilon > 0$, the SEBEU (49)-(50) constitutes an ϵ -Nash equilibrium if the number of DMs is sufficiently large. To show this, we obtain DMⁱ's optimal response $\tilde{s}^{i,N} \in \mathbb{S}^i$ to the strategies (49)-(50) as follows.

$$\begin{aligned} \tilde{V}_1^i(I_1^i, y_0) := \min_{u_1^i} \{ & |x_1^i|_{Q_1}^2 + |u_1^i|_{R_1}^2 + 2E[y_1^i|I_1^i, y_0](K_1u_1^i + L_1x_1^i) \\ & + \beta E[|A_1x_1^i + B_1u_1^i + C_1y_1 + w_1^i|_{Q_2}^2|I_1^i, y_0] \}. \end{aligned}$$

DMⁱ's optimal response at stage $t = 1$ is

$$\tilde{s}_1^{i,N}(I_1^i, y_0) = -(R_1 + \beta B_1'Q_2B_1)^{-1}(\beta B_1'Q_2A_1x_1^i + (K_1' + \beta B_1'Q_2C_1)E[y_1^i|I_1^i, y_0]).$$

We write

$$E[y_1^i|I_1^i, y_0] = \frac{1}{N}x_1^i + (A_0 + B_0F_0)E[\bar{x}_0^{-i}|I_1^i, y_0] + \frac{N-1}{N}C_0y_0,$$

where $\bar{x}_0^{-i} := \frac{1}{N} \sum_{j \neq i} x_0^j$. From $y_0 = \frac{1}{N}x_0^i + \bar{x}_0^{-i} + \xi_0$, we have

$$E[\bar{x}_0^{-i}|I_1^i, y_0] = \text{cov}[\bar{x}_0^{-i}](\text{cov}[\bar{x}_0^{-i}] + \text{cov}[\xi_0])^{-1} \left(y_0 - \frac{1}{N}x_0^i \right).$$

Therefore,

$$\tilde{s}_1^{i,N}(I_1^i, y_0) = \tilde{F}_1^N x_1^i + \tilde{G}_1^N y_0 + \tilde{F}_{1,0}^N x_0^i,$$

where

$$\begin{aligned} \tilde{F}_1^N &:= F_1 + \mathcal{O}(1/N), \\ \tilde{G}_1^N &:= G_{1,1}C_0 + \mathcal{O}(1/N), \\ \tilde{F}_{1,0}^N &:= \mathcal{O}(1/N). \end{aligned}$$

Let

$$\tilde{V}_0^i(x_0^i) := \min_{u_0^i} \{ |x_0^i|_{Q_0}^2 + |u_0^i|_{R_0}^2 + 2E[y_0^i|x_0^i](K_0u_0^i + L_0x_0^i) + \beta E[\tilde{V}_1^i(I_1^i, y_0)|x_0^i] \}.$$

DM's optimal response at stage $t = 0$ can be derived as:

$$\tilde{s}_0^{i,N}(x_0^i) = \tilde{F}_0^N x_0^i$$

where

$$\begin{aligned} \tilde{F}_0^N &:= -(R_0 + \beta B_0'M_1B_0 + \mathcal{O}(1/N))^{-1}(\beta B_0'M_1A_0 + \mathcal{O}(1/N)), \\ M_1 &= Q_1 + \beta A_1'(Q_2 - \beta Q_2B_1(R_1 + \beta B_1'Q_2B_1)^{-1}B_1'Q_2)A_1. \end{aligned}$$

It is straightforward to see that $\lim_{N \rightarrow \infty} (\tilde{F}_0^N, \tilde{F}_1^N, \tilde{G}_1^N, \tilde{F}_{1,0}^N) = (F_0, F_1, G_{1,1}C_0, 0)$. This implies that, for any $\epsilon > 0$, the SEBEU (49)-(50) is an ϵ -Nash equilibrium if the number of DMs is sufficiently large.

We should also remark that the SEBEU strategies (49)-(50) are always linear mappings of each DM^i 's information variables whereas the same does not always hold for Nash equilibria for such linear quadratic Gaussian models [7, Chapter 7].

5. CONCLUSION

We extended some of the results in our earlier work on the notion of subjective equilibrium under beliefs of exogenous uncertainty (SEBEU) to the linear-quadratic case where the players interact through some environment variables. We established the existence of an SEBEU in pure strategies under the condition that a linear system of algebraic equations are solvable. Our results are constructive and the players strategies at an SEBEU are affine functions of their local state observations and their estimates of the system state. In the Gaussian case, the estimates of the system state can be generated by a Kalman filter. We pointed out that an SEBEU is also a Nash equilibrium in the case of countable number of identical players. Showing that an SEBEU is an approximate Nash equilibrium when there is a large but finite number of players remains as a future research topic.

REFERENCES

- [1] Adlakha, S., Johari, R., Weintraub, G.Y. Equilibria of dynamic games with many players: Existence, approximation, and market structure, *J. Econ. Theory*, V.156, 2015, pp.269-316.
- [2] Arrow, K.J., Debreu, G. Existence of an equilibrium for a competitive economy, *Econometrica*, 1954, pp.265-290.
- [3] Arslan, G., Yüksel, S. Subjective equilibria under beliefs of exogenous uncertainty for dynamic games, *SIAM J. Cont. Optim.*, V.61, N.3, 2023, pp.1038-1062.
- [4] Başar, T. Decentralized multicriteria optimization of linear stochastic systems, *IEEE Trans. Autom. Control*, V.23, N.2, 1978, pp.233-243.
- [5] Başar, T. Two-criteria LQG decision problems with one-step delay observation sharing pattern, *Infor. Control*, V.38, N.1, 1978, pp.21-50.
- [6] Başar, T. Stochastic differential games and intricacy of information structures, In Josef Haunschmied, Vladimir M. Veliov, and Stefan Wrzaczek, editors, *Dynamic Games in Economics*, V.16 of *Dynamic Modeling and Econometrics in Economics and Finance*, Springer Berlin Heidelberg, 2014, pp.23-49.
- [7] Başar, T., Olsder, G.J. *Dynamic Noncooperative Game Theory: Second Edition*, Classics in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, 1999, 519 p.
- [8] Colombino, M., Smith, R.S., Summers, T.H. Mutually quadratically invariant information structures in two-team stochastic dynamic games, *IEEE Trans. Autom. Control*, V.63, N.7, 2017, pp.2256-2263.
- [9] Constantopoulos, P., Schweppe, F.C., Larson, R.C. Estia: A real-time consumer control scheme for space conditioning usage under spot electricity pricing, *Comput. Oper. Res.*, V.18, N.8, 1991, pp.751-765.
- [10] Dudebout, N., Shamma, J.S. Empirical evidence equilibria in stochastic games, In *51st IEEE Conf. Decision and Control*, 2012, pp.5780-5785.
- [11] Dudebout, N., Shamma, J.S. Exogenous empirical-evidence equilibria in perfect-monitoring repeated games yield correlated equilibria, In *53rd IEEE Conf. Decis. Control*, 2014, pp.1167-1172.
- [12] Duncan, T.E., Pasik-Duncan, B. Discrete time linear quadratic control with arbitrary correlated noise, *IEEE Trans. Autom. Control*, V.58, N.5, 2012, pp.1290-1293.
- [13] Gmytrasiewicz, P.J., Doshi, P. A framework for sequential planning in multi-agent settings, *J. Artif. Intel. Res.*, V.24, 2005, pp.49-79.
- [14] Gupta, A., Nayyar, A., Langbort, C., Başar, T. Common information based Markov perfect equilibria for linear-Gaussian games with asymmetric information, *SIAM J. Control Optim.*, V.52, N.5, 2014, pp.3228-3260.
- [15] Huang, M., Caines, P.E., Malhamé, R.P. Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ϵ -Nash equilibria, *IEEE Trans. Autom. Control*, V.52, 2007, pp.1560-1571.
- [16] Huang, M., Malhamé, R.P., Caines, P.E. Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle, *Commun. Inf. Syst.*, V.6, N.3, 2006, pp.221-252.
- [17] Jovanovic, B., Rosenthal, R.W. Anonymous sequential games, *J. Math. Econ.*, V.17, N.1, 1988, pp.77-87.
- [18] Kalai, E., Lehrer, E. Subjective games and equilibria, *Games Econ. Behav.*, V.8, N.1, 1995, pp.123-163.
- [19] Lasry, J.-M., Lions, P.-L. Mean field games, *Japanese Journal of Mathematics*, V.2, V.1, 2007, pp.229-260.

- [20] Papavassilopoulos, G. On the linear-quadratic-Gaussian Nash game with one-step delay observation sharing pattern, *IEEE Trans. Autom. Control*, V.27, N.5, 1982, pp.1065-1071.
- [21] Radner, R. Competitive equilibrium under uncertainty, *Econometrica*, 1968, pp.31-58.
- [22] Roozbehani, M., Dahleh, M.A., Mitter, S.K. Volatility of power grids under real-time pricing, *IEEE Trans. Power Syst.*, V.27, N.4, 2012, pp.1926-1940.
- [23] Sanjari, S., Başar, T., Yüksel, S. Isomorphism properties of optimality and equilibrium solutions under equivalent information structure transformations: Stochastic dynamic games and teams, *SIAM J. Control Optim.*, V.61, N.5, 2023, pp.3102-3130.
- [24] Vasal, D., Anastasopoulos, A. Signaling equilibria for dynamic LQG games with asymmetric information, *IEEE Trans. Control Netw. Syst.*, V.8, N.3, 2021, pp.1177-1188.
- [25] Weintraub, G.Y., Benkard, C.L., Van Roy, B. Industry dynamics: Foundations for models with an infinite number of firms, *J. Econ. Theory*, V.146, N.5, 2011, pp.1965-1994.
- [26] Wiszniewska-Matyskiel, A. Redefinition of belief distorted Nash equilibria for the environment of dynamic games with probabilistic beliefs, *J. Optim. Theory Appl.*, V.172, N.3, 2017, pp.984-1007.
- [27] Yüksel, S., Başar, T. *Stochastic Teams, Games and Control under Information Constraints*, Springer, Cham, 2024, 935 p.

APPENDIX A

We drop the superscript i and derive an optimal response for a generic DM with respect to independent exogenous environment variables $(z_t)_{t \in [0, T]}$ satisfying $\sup_{t \in [0, T]} E[|z_t|^2] < \infty$. DM has

the knowledge of the distribution of $(z_t)_{t \in [0, T]}$ and has access to the realizations $Z_{t-1} = (z_k)_{k \in [0, t]}$ at time $t \in [0, T)$ prior to choosing u_t . For any random vector ζ_t , we use the notation $\hat{\zeta}_t = E[\zeta_t]$ and $\hat{\zeta}_{t|k} = E[\zeta_t | Z_k]$. Assumption 1 holds in the finite-horizon case, whereas Assumption 3 holds in the infinite-horizon case. By completing the squares, we write DM's undiscounted cost at time $t \in [0, T)$ as:

$$c_t(x_t, u_t, z_t) = \tilde{c}_t(x_t, u_t, z_t) - |z_t|_{L_t Q_t^{-1} L_t' + K_t R_t^{-1} K_t'}^2,$$

where

$$\tilde{c}_t(x_t, u_t, z_t) := |x_t + Q_t^{-1} L_t' z_t|_{Q_t}^2 + |u_t + R_t^{-1} K_t' z_t|_{R_t}^2.$$

Since $E\left[\sum_{t \in [0, T)} \beta^t |z_t|_{K_t Q_t^{-1} K_t' + L_t R_t^{-1} L_t'}^2\right]$ is finite and independent of the DM's strategy, we assume

that DM's undiscounted cost function at time $t \in [0, T)$ is \tilde{c}_t without loss of generality.

DM's optimal response in the finite-horizon case can be obtained using the main result in [12] which studies the finite-horizon linear quadratic control problem with arbitrary correlated noise; however, for completeness, we provide a brief derivation of DM's optimal response below.

A.1. Finite-horizon case. We define DM's cost-to-go functions as, for all $k \in [0, T]$ and $(I_k, Z_{k-1}) \in \mathbb{I}_k \times \mathbb{Y}^{[0, k]}$,

$$V_k(I_k, Z_{k-1}) := \min_{s_k, \dots, s_{T-1}} E \left[\sum_{t \in [k, T)} \beta^{t-k} \tilde{c}_t(x_t, u_t, z_t) + \beta^{T-k} c_T(x_T) \middle| I_k, Z_{k-1} \right], \quad k \in [0, T)$$

with $V_T(I_T, Z_{T-1}) := c_T(x_T)$, where $u_n = s_n(I_n, Z_{n-1})$, $s_n : (\mathbb{I}_n, \mathbb{Y}^{[0, n]}) \rightarrow \mathbb{U}$, for $n \in [k, T)$. These cost-to-go functions satisfy, for $k \in [0, T)$,

$$V_k(I_k, Z_{k-1}) = \min_{s_k} E \left[\tilde{c}_k(x_k, u_k, z_k) + \beta V_{k+1}(I_{k+1}, Z_k) \middle| I_k, Z_{k-1} \right]. \quad (51)$$

As an induction hypothesis, assume that V_{k+1} has the following form, for $k \in [0, T)$,

$$V_{k+1}(I_{k+1}, Z_k) = |x_{k+1}|_{M_{k+1}}^2 + 2x_{k+1}' N_{k+1}(Z_k) + O_{k+1}(Z_k), \quad (52)$$

where $M_{k+1} \succeq 0$ is an appropriate dimensional matrix, N_{k+1}, O_{k+1} are appropriate dimensional functions of Z_k . V_T satisfies this hypothesis with $M_T = Q_T$, $N_T = 0$, $O_T = 0$. The minimizing control in (51) is obtained as:

$$u_k = s_k(I_k, Z_{k-1}) := F_k x_k + G_{k,k} \hat{z}_{k|k-1} + \tilde{H}_k(Z_{k-1}), \quad (53)$$

where

$$F_k = -\beta S_k^{-1} B'_k M_{k+1} A_k, \quad (54)$$

$$G_{k,k} = -S_k^{-1} (K'_k + \beta B'_k M_{k+1} C_k), \quad (55)$$

$$\tilde{H}_k(Z_{k-1}) = -\beta S_k^{-1} B'_k (\hat{N}_{k+1|k-1} + M_{k+1} \hat{w}_k), \quad (56)$$

$$S_k = R_k + \beta B'_k M_{k+1} B_k. \quad (57)$$

Substituting the minimizing control (53)-(57) in (51)-(52) results as the following:

$$\begin{aligned} V_k(I_k, Z_{k-1}) = & |x_k|_{Q_k}^2 + |F_k x_k + G_{k,k} \hat{z}_{k|k-1} + \tilde{H}_k(Z_{k-1})|_{R_k}^2 \\ & + 2\hat{z}'_{k|k-1} (K_k (F_k x_k + G_{k,k} \hat{z}_{k|k-1} + \tilde{H}_k(Z_{k-1})) + L_k x_k) \\ & + 2E[|z_k|_{K_k Q_k^{-1} K'_k + L_k R_k^{-1} L'_k}^2 | Z_{k-1}] \\ & + \beta E[(A_k + B_k F_k) x_k + B_k (G_{k,k} \hat{z}_{k|k-1} \\ & + \tilde{H}_k(Z_{k-1})) + C_k z_k + w_k]_{M_{k+1}}^2 | Z_{k-1}] \\ & + \beta E[2((A_k + B_k F_k) x_k + B_k (G_{k,k} \hat{z}_{k|k-1} \\ & + \tilde{H}_k(Z_{k-1})) + C_k z_k + w_k)' N_{k+1}(Z_k) | Z_{k-1}] \\ & + \beta E[O_{k+1}(Z_k) | Z_{k-1}]. \end{aligned} \quad (58)$$

Grouping the constant, linear, and quadratic terms in x_k for fixed Z_{k-1} , we rewrite $V_k(I_k, Z_{k-1})$ as:

$$V_k(I_k, Z_{k-1}) = |x_k|_{M_k}^2 + 2x'_k N_k(Z_{k-1}) + O_k(Z_{k-1}),$$

where

$$M_k = Q_k + \beta A'_k (M_{k+1} - \beta M_{k+1} B_k S_k^{-1} B'_k M_{k+1}) A_k, \quad (59)$$

$$N_k(Z_{k-1}) = \beta (A_k + B_k F_k)' \hat{N}_{k+1|k-1} + \hat{v}_{k|k-1}, \quad (60)$$

$$v_k = (F'_k K'_k + L'_k) z_k + \beta (A_k + B_k F_k)' M_{k+1} (C_k z_k + w_k), \quad (61)$$

$$O_k(Z_{k-1}) = \beta \hat{O}_{k+1|k-1} + \hat{\lambda}_{k|k-1} \quad (62)$$

$$\begin{aligned} \lambda_k = & |G_{k,k} \hat{z}_{k|k-1} + \tilde{H}_k(Z_{k-1})|_{R_k}^2 + 2\hat{z}'_k K_k (G_{k,k} \hat{z}_{k|k-1} + \tilde{H}_k(Z_{k-1})) \\ & + |z_k|_{K_k Q_k^{-1} K'_k + L_k R_k^{-1} L'_k}^2 + \beta |B_k (G_{k,k} \hat{z}_{k|k-1} + \tilde{H}_k(Z_{k-1})) + C_k z_k + w_k|_{M_{k+1}}^2 \\ & + 2\beta (B_k (G_{k,k} \hat{z}_{k|k-1} + \tilde{H}_k(Z_{k-1})) + C_k z_k + w_k)' N_{k+1}(Z_k). \end{aligned} \quad (63)$$

The recursion (59) is the well-known Riccati difference equation for discrete time linear quadratic problems (without the environment variables), $N_k(Z_{k-1})$ is obtained from the right hand side of (58) by collecting the linear terms in x_k for fixed Z_{k-1} and using $R_k F_k + \beta B'_k M_{k+1} (A_k + B_k F_k) = 0$, and $O_k(Z_{k-1})$ equals the right hand side of (58) for $x_k = 0$. Since $V_k(I_k, Z_{k-1})$ satisfies the induction hypothesis, an optimal strategy for the DM is given by (53). This optimal strategy is rewritten as, for $k \in [0, T)$,

$$s_k(I_k, Z_{k-1}) = F_k x_k + \sum_{t \in [k, T)} G_{k,t} \hat{z}_{t|k-1} + H_k,$$

where

$$\begin{aligned} G_{k,k} &= -S_k^{-1}(K'_k + \beta B'_k M_{k+1} C_k), \\ G_{k,t} &= -\beta S_k^{-1} B'_k (\Phi_{k,t} M_{t+1} C_t + \Phi_{k,t-1} (F'_t K'_t + L'_t)), \quad t \in [k+1, T), \\ H_k &= -\beta S_k^{-1} B'_k \sum_{t \in [k, T)} \Phi_{k,t} M_{t+1} \hat{w}_t, \\ \Phi_{k,k} &= I, \quad \Phi_{k,t} = \prod_{n \in [k+1, t]} \beta (A_n + B_n F_n)', \quad t \in [k+1, T). \end{aligned}$$

A.2. Infinite-horizon case. Recall that $A_t, B_t, C_t, K_t, L_t, Q_t, R_t$ are independent of t (hence the subscript t is dropped) and $\beta \in (0, 1)$.

Let $J_k^*(x_k, Z_{k-1})$ denote DM's optimal cost for the infinite-horizon problem starting from (x_k, Z_{k-1}) at time $k \in \mathbb{N}_0$. For any $\tilde{T} \in \mathbb{N}$, let $J_{k, \tilde{T}}^*(x_k, Z_{k-1})$ denote DM's optimal cost for the finite-horizon version of the problem starting from (x_k, Z_{k-1}) at time $k \in \mathbb{N}_0$ and running over the horizon $[k, \tilde{T})$ (with no terminal cost). We have, for $k \in \mathbb{N}_0$,

$$J_{k, k+1}^*(x_k, Z_{k-1}) \leq J_{k, k+2}^*(x_k, Z_{k-1}) \leq \dots \leq J_k^*(x_k, Z_{k-1}) < \infty,$$

where $J_k^*(x_k, Z_{k-1}) < \infty$ is due to the stabilizability of (A, B) and $\sup_{t \in \mathbb{N}_0} E[|z_t|^2] < \infty$. Hence, for any $k \in \mathbb{N}_0$ and (x_k, Z_{k-1}) , $\lim_{\tilde{T} \rightarrow \infty} J_{k, \tilde{T}}^*(x_k, Z_{k-1}) = J_{k, \infty}^*(x_k, Z_{k-1})$ for some $J_{k, \infty}^*(x_k, Z_{k-1}) \leq J_k^*(x_k, Z_{k-1})$. Note that

$$J_{k, \tilde{T}}^*(x_k, Z_{k-1}) = |x_k|_{M_k}^2 + 2x_k' N_k(Z_{k-1}) + O_k(Z_{k-1}),$$

where $M_k, N_k(Z_{k-1}), O_k(Z_{k-1})$, $k \in \mathbb{N}_0$, are generated by the backward recursions (59)-(63) with the boundary conditions $M_{\tilde{T}} = N_{\tilde{T}} = O_{\tilde{T}} = 0$. This implies that, for any fixed $k \in \mathbb{N}_0$ and Z_{k-1} , $M_k, N_k(Z_{k-1}), O_k(Z_{k-1})$ converge to some finite values as $\tilde{T} \rightarrow \infty$.

It is well known that, as $\tilde{T} \rightarrow \infty$, M_k for any fixed $k \in \mathbb{N}_0$ tends to the unique solution $M \succeq 0$ of the algebraic Riccati equation

$$M = Q + \beta A'(M - \beta M B S^{-1} B' M) A \quad (64)$$

within the set of positive semi-definite matrices since (A, B) is stabilizable and $(A, Q^{1/2})$ is detectable. Therefore, as $\tilde{T} \rightarrow \infty$, $F_k, G_{k,k}, S_k$ for any fixed $k \in \mathbb{N}_0$ introduced in (54)-(57) converge to

$$\begin{aligned} F &:= -\beta S^{-1} B' M A, \\ G &:= -S^{-1} (K' + \beta B' M C), \\ S &:= R + \beta B' M B, \end{aligned}$$

respectively, where F renders the matrix $\sqrt{\beta}(A + BF)$ asymptotically stable. It follows from (59)-(63) that, as $\tilde{T} \rightarrow \infty$, $N_k(Z_{k-1}), \tilde{H}_k(Z_{k-1}), O_k(Z_{k-1})$ for any fixed $k \in \mathbb{N}_0$ and Z_{k-1} tend to

$$\begin{aligned} N_{k, \infty}(Z_{k-1}) &:= \sum_{n \in \mathbb{N}_0} (\beta(A + BF)')^n \hat{v}_{k+n|k-1}, \\ \tilde{H}_{k, \infty}(Z_{k-1}) &:= -\beta S^{-1} B' \left(\sum_{n \in \mathbb{N}_0} (\beta(A + BF)')^n \hat{v}_{k+n+1|k-1} + M \hat{w}_0 \right), \\ O_{k, \infty}(Z_{k-1}) &= \sum_{n \in \mathbb{N}_0} \beta^n \hat{\lambda}_{k+n|k-1}, \end{aligned}$$

respectively, where

$$\begin{aligned} \bar{v}_k &:= (F'K' + L)z_k + \beta(A + BF)'M(Cz_k + w_k), \\ \bar{\lambda}_k &:= |G\hat{z}_{k|k-1} + \tilde{H}_{k,\infty}(Z_{k-1})|_R^2 + 2z_k'K(G\hat{z}_{k|k-1} + \tilde{H}_{k,\infty}(Z_{k-1})) \\ &\quad + |z_k|_{KQ^{-1}K'+LR^{-1}L'}^2 + \beta|B(G\hat{z}_{k|k-1} + \tilde{H}_{k,\infty}(Z_{k-1})) + Cz_k + w_k|_M^2 \\ &\quad + 2\beta(B(G\hat{z}_{k|k-1} + \tilde{H}_{k,\infty}(Z_{k-1})) + Cz_k + w_k)'N_{k+1,\infty}(Z_k). \end{aligned}$$

Straightforward computation shows that $(J_{k,\infty})_{k \in \mathbb{N}_0}$ satisfy the Bellman equations, i.e.,

$$\begin{aligned} J_{k,\infty}^*(x_k, Z_{k-1}) &= |x_k|_M^2 + 2x_k'N_{k,\infty}(Z_{k-1}) + O_{k,\infty}(Z_{k-1}) \\ &= \min_{u_k \in \mathbb{U}} E[|x_k + Q^{-1}L'z_k|_Q^2 + |u_k + R^{-1}K'z_k|_R^2 \\ &\quad + \beta J_{k+1,\infty}^*(x_{k+1}, Z_k)|x_k, Z_{k-1}] \end{aligned} \tag{65}$$

with the minimizing control

$$u_k^* = s_k^*(x_k, Z_{k-1}) := Fx_k + G\hat{z}_{k|k-1} + \tilde{H}(Z_{k-1}). \tag{66}$$

Now, let $J_{k,s^*}(x_k, Z_{k-1})$ denote the cost achieved by the strategy $s^* = (s_0^*, s_1^*, \dots)$ starting from (x_k, Z_{k-1}) at time $k \in \mathbb{N}_0$. Let $(x_{n+1}^*, u_n^*)_{n \in [k, \tilde{T}]}$ be generated by $(s_n^*, z_n)_{n \in [k, \tilde{T}]}$ from $(x_k^* = x_k, Z_{k-1})$. Since $(J_{k,\infty}^*)_{k \in \mathbb{N}_0}$ is non-negative valued, we write

$$\begin{aligned} J_{k,s^*}(x_k, Z_{k-1}) &\leq \lim_{\tilde{T} \rightarrow \infty} E \left[\sum_{n \in [k, \tilde{T}]} \beta^{n-k} \tilde{c}_n(x_n^*, u_n^*, z_n) + \beta^{\tilde{T}-k} J_{\tilde{T},\infty}^*(x_{\tilde{T}}^*, Z_{\tilde{T}-1}) \middle| x_k, Z_{k-1} \right] \\ &= \lim_{\tilde{T} \rightarrow \infty} E \left[\sum_{n \in [k, \tilde{T}-1]} \beta^{n-k} \tilde{c}_n(x_n^*, u_n^*, z_n) + \beta^{\tilde{T}-k-1} J_{\tilde{T}-1,\infty}^*(x_{\tilde{T}-1}^*, Z_{\tilde{T}-2}) \middle| x_k, Z_{k-1} \right] \\ &\quad \vdots \\ &= J_{k,\infty}^*(x_k, Z_{k-1}), \end{aligned}$$

where the equalities are obtained by repeated application of (65)-(66). Therefore, the strategy s^* is optimal.

Finally, consider any optimal strategy $\check{s} = (\check{s}_0, \check{s}_1, \dots)$. We must have, for $k \in \mathbb{N}_0$,

$$\begin{aligned} J_k^*(x_k, Z_{k-1}) &= E[\tilde{c}_k(x_k, \check{u}_k, z_k) + \beta J_{k+1}^*(Ax_k + B\check{u}_k + Cz_k + w_k, Z_k)|x_k, Z_{k-1}] \\ &= \min_{u_k} E[\tilde{c}_k(x_k, u_k, z_k) + \beta J_{k+1}^*(Ax_k + Bu_k + Cz_k + w_k, Z_k)|x_k, Z_{k-1}], \end{aligned}$$

where $\check{u}_k = \check{s}_k(x_k, Z_{k-1})$. Since the minimum above is uniquely achieved by $s_k^*(x_k, Z_{k-1})$, we have $\check{s} = s^*$. The optimal strategy is rewritten as:

$$u_k^* = s_k^*(x_k, Z_{k-1}) = Fx_k + \sum_{n \in \mathbb{N}_0} G_n \hat{z}_{k+n|k-1} + H,$$

where

$$\begin{aligned} G_0 &= -S^{-1}(K' + \beta B'MC), \\ G_n &= -\beta S^{-1}B'(\beta(A + BF)')^n((F'K' + L') + \beta(A + BF)'MC), \quad n \in \mathbb{N}, \\ H &= -\beta S^{-1}B'(I - \beta(A + BF)')^{-1}M\hat{w}_0. \end{aligned}$$



Gürdal Arslan - received his Ph.D. degree in Electrical Engineering from the University of Illinois at Urbana-Champaign in 2001. From 2001 to 2004, he was an Assistant Researcher in the Department of Mechanical and Aerospace Engineering, University of California, Los Angeles. In August 2004, he joined the Electrical Engineering Department at the University of Hawaii, Manoa, where he is currently an Associate Professor. His current research interests lie in the design of cooperative multi-agent systems using game theoretic methods. He received the National Science Foundation Career Award in May 2006. He currently serves as an Associate Editor for *Automatica*.



Serdar Yüksel - received his B.Sc. degree in Electrical and Electronics Engineering from Bilkent University in 2001, M.Sc. and Ph.D. degrees in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign in 2003 and 2006, respectively. He was a post-doctoral researcher at Yale University before joining Queen's University as an Assistant Professor in the Department of Mathematics and Statistics, where he is now a Professor. He has been awarded the 2013 CAIMS/PIMS Early Career Award in Applied Mathematics.

His research interests are on stochastic control, decentralized control, information theory and probability. He has been an Associate Editor for the *IEEE Transactions on Automatic Control*, *Automatica*, *Systems and Control Letters*, and *Mathematics of Control, Signals, and Systems*.