Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Non-Cooperative Outcomes for Stochastic Nash Games: Decision Strategies towards Multi-Attribute Performance Robustness ⋆

Khanh D. Pham *

* *Air Force Research Laboratory, Kirtland AFB, NM 87117 USA (E-mail: AFRL.RVSV@kirtland.af.mil).*

Abstract: The advantages of compactness from logic of state-space model description and quantitativity from probabilistic knowledge of stochastic disturbances have been exploited to construct a situational awareness which then provides essential common knowledge to non-cooperative decision makers about the adverse and dynamic environment within a linear-quadratic class of nonzero-sum stochastic games. It incorporates the perception of relevant attributes of the decision problem, comprehension of the meaning of the shared interaction model in combination with and in relation to various goals of non-cooperative decision makers so that future projection of higher-order characteristics of the Chi-squared random measures of performance is obtained with high confidence. New solution concepts, called the multi-cumulant Nash strategies are proposed to directly influence respective performance distributions and to effectively guarantee performance robustness for non-cooperative decision makers.

Keywords: Integral-quadratic performance-measure, information statistics, stochastic Nash games, non-cooperative decision makers, performance robustness, dynamic programming, cumulant-based control, certainty equivalence principle

1. INTRODUCTION

In non-cooperative multi-player differential games, there may be more than two decision makers who wish to optimize different utility functions. Each decision maker manipulates his strategies to a shared interaction process, described by a stochastic differential equation and chooses a criterion of performance that reflects his a-prior interest and finally a strategy is attained via optimization of his measure of performance. To the best knowledge of the author, most studies on nonzero-sum differential games for instance, Basar (1982), Cruz (2001) and references therein have concentrated on the strategy selection part of a Nash equilibrium. This equilibrium ensures that no decision maker has incentives to unilaterally deviate in order to improve his expected performance with respect to all realizations of the environmental disturbances. Very little work, if any, has been published on the subject of higher-order characteristic performance variability introduced by the process noise stochasticity that involves: 1) whether it is possible to construct beliefs that reflect genuine uncertainty about the environmental disturbances, namely Nature's mixed random realizations, yet are narrow enough to permit learning on performance robustness with respect to variability in the stochastic environment and 2) the support of beliefs should, on the one hand, encompass mixed random strategies, reflecting uncertainty about Nature and, on the other hand, needs to include strategies that actually are optimal for non-cooperative decision makers given the support of their beliefs. In economics context, the importance of incorporating aversion to specification uncertainty has been considered by Hansen (1999). Rather recently, a new stochastic control theory of robust design is also already in place wherein the accounts of Pham (2005, 2007a,b, 2008) have addressed risk aversion for performance uncertainty in stochastic regulators and cooperative decision-making. The main contribution of the research investigation considered here is the extension of the aforementioned results Pham (2007b, 2008) to another stochastic class of non-cooperative multi-person games where all respective measures of performance of non-cooperative decision makers are viewed as random variables with Nature's mixed random realizations. The action space of Nature regarding all realizations of the underlying stochastic process is assumed to be common knowledge, but a realized measure of performance is private and known only to that particular decision maker. Included with the present work are some innovative answers to completely unexplored research areas as such: i) Non-cooperative decision makers are not reasonably content with such average measures of performance; ii) An efficient and tractable procedure that calculates exactly higher-order characteristics of performance distributions; and iii) New optimal strategies that guarantee performance robustness with something much stronger than a stochastically averaging measure of performance.

2. PERFORMANCE-MEASURE STATISTICS

A class of stochastic decision problems with N non-cooperative decision makers, identified as $u^1, \ldots, u^N$ is considered on a finite horizon $[t_0, t_f]$. The initial decision

⋆ This work was supported in part by the Frank M. Freimann Chair in Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556 U.S.A.

978-1-1234-7890-2/08/$20.00 © 2008 IFAC **10.3182/20080706-5-KR-1001.0692**

state at time $t = t_0$, $x(t_0) \triangleq x_0 \in \mathbb{R}^n$ is fixed. It is assumed that a Nash decision system of uncertainty and robustness is engaged against Nature in a stationary (stochastic) environment whose representation $w(t) \triangleq w(t,\omega) : [t_0, t_f] \times \Omega \mapsto \mathbb{R}^p$ is an p-dimensional stationary Wiener process defined with $\{\mathcal{F}_t\}_{t\geq t_0>0}$ being its filtration on a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t\geq t_0>0}, \mathcal{P})$ over $[t_0, t_f]$ with the correlation of independent increments

$$E\left\{[w(\tau) - w(\xi)][w(\tau) - w(\xi)]^T\right\} = W|\tau - \xi|, \quad W > 0.$$

Next, non-cooperative decision makers choose actions in strategy spaces $\mathcal{U}^i \in L^2_{\mathcal{F}_t}(\Omega; \mathcal{C}([t_0, t_f]; \mathbb{R}^{m_i}))$ and $i = 1, \ldots, N$ being the subsets of Hilbert space of $\mathbb{R}^{m_i}$-valued square integrable processes on $[t_0, t_f]$ that are adapted to the sigma-field $\mathcal{F}_t$ generated by $w(t)$. Associated with each admissible (N+1)-tuple $(x(\cdot); u^1(\cdot), \ldots, u^N(\cdot))$ is a random measure of performance $J^i : \mathbb{R}^n \times \mathcal{U}^1 \times \cdots \times \mathcal{U}^N \mapsto \mathbb{R}^+$

$$J^i(x_0; u^1, \ldots, u^N) = x^T(t_f) Q^i_f x(t_f) \tag{1}$$
$$+ \int_{t_0}^{t_f} \left[x^T(\tau) Q^i(\tau) x(\tau) + \sum_{j=1}^{N} (u^j)^T(\tau) R^{ij}(\tau) u^j(\tau) \right] d\tau$$

for i-th decision maker where states of the interaction process, $x(t) \triangleq x(t,\omega) : [t_0, t_f] \times \Omega \mapsto \mathbb{R}^n$ belong to the Hilbert space $L^2_{\mathcal{F}_t}(\Omega; \mathcal{C}([t_0, t_f]; \mathbb{R}^n))$ with $E\left\{\int_{t_0}^{t_f} x^T(\tau) x(\tau) d\tau\right\}$ finite and evolve via the stochastic differential equation

$$dx(t) = \left[A(t)x(t) + \sum_{i=1}^{N} B^i(t) u^i(t) \right] dt + G(t) dw(t)\,,$$
$$x(t_0) = x_0 \tag{2}$$

in which $A \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{n\times n})$, $B^i \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{n\times m_i})$, and $G \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{n\times p})$ are deterministic matrix-valued functions. Design parameters $Q^i_f \in \mathbb{R}^{n\times n}$, $Q^i \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{n\times n})$, and $R^{ij} \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{m_i\times m_i})$ representing relative weightings for terminal state, transient state trajectories, and actions are deterministic and positive semidefinite with $R^{ii}(t)$ invertible.

Since the interaction model (2) is linear, it is often argued that the actions of non-cooperative decision makers should be a function of the states. The restriction of strategy spaces to the set of so-called Markov functions can be justified by the assumption that non-cooperative decision makers participate in the Nash game where they only have access to the current time and state of the interaction. Therefore, it amounts to considering only those feedback Nash equilibria which permit a linear feedback synthesis $\gamma^i : [t_0, t_f] \times L^2_{\mathcal{F}_t}(\Omega; \mathcal{C}([t_0, t_f]; \mathbb{R}^n)) \mapsto L^2_{\mathcal{F}_t}(\Omega; \mathcal{C}([t_0, t_f]; \mathbb{R}^{m_i}))$

$$u^i(t) = \gamma^i(t, x(t)) \triangleq K^i(t) x(t) \tag{3}$$

where the admissible feedback gains $K^i \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{m_i\times n})$ will be appropriately defined. In particular, this subclass of feedback Nash equilibria should satisfy the requirement.

Definition 1. Feedback Nash Equilibrium.
A set of equilibrium actions $u^i_*(t)$ is called strongly time consistent if, for all $t_1 \in [t_0, t_f)$, these actions constitute a Nash equilibrium for the truncated decision problem defined on $[t_1, t_f]$ where $x(t_1)$ is an arbitrarily chosen state which is reachable from some initial state at $t = t_0$.

For the given (t_0, x_0) and subject to the strategies (3), the dynamics of the interaction process (2) follows

$$dx(t) = \left[A(t) + \sum_{i=1}^{N} B^i(t) K^i(t) \right] x(t) dt + G(t) dw(t)\,,$$
$$x(t_0) = x_0\,, \tag{4}$$

and the performance-measure associated with i-th decision maker becomes

$$J^i(x_0; K^1, \ldots, K^N) = x^T(t_f) Q^i_f x(t_f) \tag{5}$$
$$+ \int_{t_0}^{t_f} x^T(\tau) \left[Q^i(\tau) + \sum_{j=1}^{N} (K^j)^T(\tau) R^{ij}(\tau) K^j(\tau) \right] x(\tau) d\tau\,.$$

In addition, the stochastic interaction system (2) in the absence of process noises is assumed to be uniformly exponentially stable. That is, there exist positive constants η_1 and η_2 such that the pointwise matrix norm of the closed-loop state transition matrix satisfies the inequality

$$||\Phi(t,\tau)|| \leq \eta_1 e^{-\eta_2 (t-\tau)} \qquad \forall\ t \geq \tau \geq t_0\ .$$

The pair $(A(t), [B^1(t) \ldots B^N(t)])$ is stabilizable if there exist bounded matrix-valued functions $K^1(t), \ldots, K^N(t)$ such that $dx(t) = \left(A(t) + \sum_{i=1}^{N} B^i(t) K^i(t)\right) x(t) dt$ is uniformly exponentially stable.

Within the view of the linear-quadratic structure of the decision problem, the performance-measure (5) for i-th decision maker is clearly a random variable with Chi-squared type. Hence, the uncertainty of performance distribution must be assessed via a complete set of higher-order statistics beyond the statistical averaging. It is therefore necessary to develop a mathematical construct and support of the beliefs on performance uncertainty to extract the knowledge in definite terms of performance-measure statistics for each decision maker. This is done by adopting the results in Pham (2007b).

Theorem 1. Performance-Measure Statistics.
Let decision makers choose actions $u^i(t) = K^i(t) x(t)$, the pair $(A, [B^1 \ldots B^N])$ be uniformly stabilizable on $[t_0, t_f]$, and the multi-person Nash decision system be governed by (4) and (5). For $k_i \in \mathbb{Z}^+$ fixed, the k_i-th cumulant associated with i-th decision maker is given by

$$\kappa^i_k(t_0, x_0) = x_0^T H^i(t_0, k_i) x_0 + D^i(t_0, k_i) \tag{6}$$

where the cumulant-generating components $\{H^i(\alpha, r)\}_{r=1}^{k_i}$ and $\{D^i(\alpha, r)\}_{r=1}^{k_i}$ evaluated at $\alpha = t_0$ satisfy the supporting matrix-valued differential equations (with the dependence of $H^i(\alpha, r)$ and $D^i(\alpha, r)$ upon the admissible gains $K^1, \ldots, K^N$ suppressed)

$$\frac{d}{d\alpha} H^i(\alpha, 1) = -\left[A(\alpha) + \sum_{j=1}^{N} B^j(\alpha) K^j(\alpha) \right]^T H^i(\alpha, 1)$$
$$- H^i(\alpha, 1) \left[A(\alpha) + \sum_{j=1}^{N} B^j(\alpha) K^j(\alpha) \right]$$
$$- Q^i(\alpha) - \sum_{j=1}^{N} (K^j)^T(\alpha) R^{ij}(\alpha) K^j(\alpha) \tag{7}$$

and, for $2 \leq r \leq k_i$

$$\frac{d}{d\alpha}H^i(\alpha,r) = -\left[A(\alpha)+\sum_{j=1}^{N}B^j(\alpha)K^j(\alpha)\right]^T H^i(\alpha,r)$$
$$- H^i(\alpha,r)\left[A(\alpha)+\sum_{j=1}^{N}B^j(\alpha)K^j(\alpha)\right]$$
$$-\sum_{s=1}^{r-1}\frac{2r!}{s!(r-s)!}H^i(\alpha,s)G(\alpha)WG^T(\alpha)H^i(\alpha,r-s) \quad (8)$$

together with $1 \le r \le k_i$

$$\frac{d}{d\alpha}D^i(\alpha,r) = -\mathrm{Tr}\left\{H^i(\alpha,r)G(\alpha)WG^T(\alpha)\right\} \quad (9)$$

where the terminal-value conditions $H^i(t_f,1) = Q^i_f$, $H^i(t_f,r) = 0$ for $2 \le r \le k_i$ and $D^i(t_f,r) = 0$ for $1 \le r \le k_i$.

Clearly then, the compactness offered by logic from the state-space model description (4) has been successfully combined with the quantitativity from a-priori probabilistic knowledge of Nature's noise characteristics. Thus, the uncertainty of individual performance (5) can now be represented in a compact and robust way. Subsequently, the time-backward differential equations (7)-(9) not only offer a tractable procedure for the calculation of (6) but also allow the incorporation of a subclass of linear feedback Nash equilibria. Such performance-measure statistics are therefore, referred as "information" statistics which are extremely valuable for shaping i-th decision maker's performance distribution.

3. PROBLEM STATEMENTS

Suffice it to say here that all the performance-measure statistics, or equivalently cumulants (6) depend in part of the known initial condition $x(t_0)$. Although different states $x(t)$ will result in different values for the traditional "performance-to-come", the cumulant values are however, functions of time-backward evolutions of the cumulant-generating variables $H^i(\alpha,r)$ and $D^i(\alpha,r)$ that totally ignore all the values $x(t)$. This fact therefore makes the new optimization problem as being considered in cumulant-based control particularly unique as compared with the more traditional dynamic programming class of investigations. In other words, the time-backward trajectories (7)-(9) should be considered as the "new" dynamical equations from which the resulting Mayer optimization and associated value function in the framework of dynamic programming Fleming (1975) thus depend on these "new" state variables $H^i(\alpha,r)$ and $D^i(\alpha,r)$.

For notational simplicity, k_i-tuple variables $\mathcal{H}^i$ and $\mathcal{D}^i$ are introduced as the states of performance uncertainty for i-th decision maker with $\mathcal{H}^i(\cdot) \triangleq \left(\mathcal{H}^i_1(\cdot),\ldots,\mathcal{H}^i_{k_i}(\cdot)\right)$ and $\mathcal{D}^i(\cdot) \triangleq \left(\mathcal{D}^i_1(\cdot),\ldots,\mathcal{D}^i_{k_i}(\cdot)\right)$ wherein each element $\mathcal{H}^i_r \in \mathcal{C}^1([t_0,t_f];\mathbb{R}^{n\times n})$ of $\mathcal{H}^i$ and each element $\mathcal{D}^i_r \in \mathcal{C}^1([t_0,t_f];\mathbb{R})$ of $\mathcal{D}^i$ have the representations of $\mathcal{H}^i_r(\cdot) = H^i(\cdot,r)$ and $\mathcal{D}^i_r(\cdot) = D^i(\cdot,r)$, with the right members satisfying the dynamic equations (7)-(9). The convenient mappings are defined by

$$\mathcal{F}^i_r : [t_0,t_f]\times(\mathbb{R}^{n\times n})^{k_i}\times\mathbb{R}^{m_1\times n}\times\cdots\times\mathbb{R}^{m_N\times n}\mapsto\mathbb{R}^{n\times n}$$
$$\mathcal{G}^i_r : [t_0,t_f]\times(\mathbb{R}^{n\times n})^{k_i}\mapsto\mathbb{R}$$

with the actions

$$\mathcal{F}^i_1(\alpha,\mathcal{H}^i,K^1,\ldots,K^N) \triangleq$$
$$-\left[A(\alpha)+\sum_{j=1}^{N}B^j(\alpha)K^j(\alpha)\right]^T\mathcal{H}^i_1(\alpha)$$
$$-\mathcal{H}^i_1(\alpha)\left[A(\alpha)+\sum_{j=1}^{N}B^j(\alpha)K^j(\alpha)\right]$$
$$-Q^i(\alpha)-\sum_{j=1}^{N}(K^j)^T(\alpha)R^{ij}(\alpha)K^j(\alpha),$$

$$\mathcal{F}^i_r(\alpha,\mathcal{H}^i,K^1,\ldots,K^N) \triangleq$$
$$-\left[A(\alpha)+\sum_{j=1}^{N}B^j(\alpha)K^j(\alpha)\right]^T\mathcal{H}^i_r(\alpha)$$
$$-\mathcal{H}^i_r(\alpha)\left[A(\alpha)+\sum_{j=1}^{N}B^j(\alpha)K^j(\alpha)\right]$$
$$-\sum_{s=1}^{r-1}\frac{2r!}{s!(r-s)!}\mathcal{H}^i_s(\alpha)G(\alpha)WG^T(\alpha)\mathcal{H}^i_{r-s}(\alpha),\quad 2\le r\le k_i$$

$$\mathcal{G}^i_r(\alpha,\mathcal{H}^i) \triangleq -\mathrm{Tr}\left\{\mathcal{H}^i_r(\alpha)G(\alpha)WG^T(\alpha)\right\},\quad 1\le r\le k_i.$$

Now it is straightforward to establish the product mappings $\mathcal{F}^i_1\times\cdots\times\mathcal{F}^i_{k_i} : [t_0,t_f]\times(\mathbb{R}^{n\times n})^{k_i}\times\mathbb{R}^{m_1\times n}\times\cdots\times\mathbb{R}^{m_N\times n}\mapsto(\mathbb{R}^{n\times n})^{k_i}$ and $\mathcal{G}^i_1\times\cdots\times\mathcal{G}^i_{k_i} : [t_0,t_f]\times(\mathbb{R}^{n\times n})^{k_i}\mapsto\mathbb{R}^{k_i}$ along with the corresponding notations $\mathcal{F}^i \triangleq \mathcal{F}^i_1\times\cdots\times\mathcal{F}^i_{k_i}$ and $\mathcal{G}^i \triangleq \mathcal{G}^i_1\times\cdots\times\mathcal{G}^i_{k_i}$. Thus, the dynamic equations of performance uncertainty (7)-(9) can be rewritten as

$$\frac{d}{d\alpha}\mathcal{H}^i(\alpha) = \mathcal{F}^i(\alpha,\mathcal{H}^i(\alpha),K^1(\alpha),\ldots,K^N(\alpha)), \quad (10)$$
$$\frac{d}{d\alpha}\mathcal{D}^i(\alpha) = \mathcal{G}^i(\alpha,\mathcal{H}^i(\alpha)) \quad (11)$$

where the terminal-value conditions $\mathcal{H}^i(t_f) \triangleq \mathcal{H}^i_f = (Q^i_f,0,\ldots,0)$ and $\mathcal{D}^i(t_f) \triangleq \mathcal{D}^i_f = (0,\ldots,0)$.

Note $\mathcal{H}^i = \mathcal{H}^i(\cdot,K^1,\ldots,K^N)$ and $\mathcal{D}^i = \mathcal{D}^i(\cdot,K^1,\ldots,K^N)$. The performance index for the multi-person stochastic decision problem can be formulated in $K^1,\ldots,K^N$.

Definition 2. Performance Index.
Fix $k_i \in \mathbb{Z}^+$, the sequence $\mu^i = \{\mu^i_r \ge 0\}_{r=1}^{k_i}$ with $\mu^i_1 > 0$. Then for the given initial condition (t_0,x_0), the i-th decision maker minimizes his own performance index

$$\phi^i_0 : \{t_0\}\times(\mathbb{R}^{n\times n})^{k_i}\times\mathbb{R}^{k_i}\mapsto\mathbb{R}^+$$

over a finite optimization horizon is defined as

$$\phi^i_0\left(t_0,\mathcal{H}^i(t_0,K^1,\ldots,K^N),\mathcal{D}^i(t_0,K^1,\ldots,K^N)\right)$$
$$\triangleq\sum_{r=1}^{k_i}\mu^i_r\kappa^i_r(K^1,\ldots,K^N), \quad (12)$$
$$=\sum_{r=1}^{k_i}\mu^i_r\left[x_0^T\mathcal{H}^i_r(t_0,K^1,\ldots,K^N)x_0+\mathcal{D}^i_r(t_0,K^1,\ldots,K^N)\right]$$

where parametric design freedom μ^i_r chosen by i-th decision maker represent different levels of influence as they deem important to his performance distribution. Solutions

$\{\mathcal{H}_r^i(\alpha, K^1, \ldots, K^N)\}_{r=1}^{k_i}$ and $\{\mathcal{D}_r^i(\alpha, K^1, \ldots, K^N)\}_{r=1}^{k_i}$ when evaluated at $\alpha = t_0$ satisfy the dynamic equations (10)-(11) with the terminal-value conditions $\mathcal{H}_f^i = (Q_f^i, 0, \ldots, 0)$, and $\mathcal{D}_f^i = (0, \ldots, 0)$.

Remark 1. The performance index (12) associated with i-th decision maker is a weighted summation of some information statistics with μ_r^i representing multiple degrees of shaping the probability density function of (5). If all the cumulants of (5) remain bounded as (5) arbitrarily closes to 0, the first cumulant dominates the summation and the cumulant-based optimization problem reduces to the classical Linear-Quadratic-Gaussian (LQG) control problem.

For the given $(t_f, \mathcal{H}_f^i, \mathcal{D}_f^i)$, the class $\mathcal{K}^i_{t_f, \mathcal{H}_f^i, \mathcal{D}_f^i; \mu^i}$ of admissible feedback gains is then defined.

Definition 3. Admissible Feedback Gain.
Let a compact subset $\overline{K}^i \subset \mathbb{R}^{m_i \times n}$ be the set of allowable gain values chosen by i-th decision maker. For the given $k_i \in \mathbb{Z}^+$, $\mu^i = \{\mu_r^i \geq 0\}_{r=1}^{k_i}$ with $\mu_1^i > 0$, the sets of admissible feedback gain strategies $\mathcal{K}^i_{t_f, \mathcal{H}_f^i, \mathcal{D}_f^i; \mu^i}$ is assumed to be the class of $\mathcal{C}([t_0, t_f]; \mathbb{R}^{m_i \times n})$ with values $K^i(\cdot) \in \overline{K}^i$ for which solutions to the dynamic equations (10)-(11) exist on the finite horizon $[t_0, t_f]$.

Since cooperation can not be enforced in the multi-person decision problem with multi-criteria objectives, a Nash equilibrium solution concept ensures that no decision makers have incentive to unilaterally deviate from the equilibrium decision laws in order to further reduce their performance indices. Thus, the Nash game-theoretic framework is suitable to capture the nature of conflicts as actions of a decision maker are tightly coupled with those of other remaining decision makers.

Definition 4. Nash Equilibrium Solution.
Let $\phi_0^i\big(t_0, \mathcal{H}^i\left(t_0, K^1, \ldots, K^{i-1}, K^i, K^{i+1}, \ldots, K^N\right),$ $\mathcal{D}^i\left(t_0, K^1, \ldots, K^{i-1}, K^i, K^{i+1}, K^N\right)\big)$ associate with i-th decision maker. Then, the N-tuple of admissible feedback gains $\left(K_*^1, \ldots, K_*^N\right)$ provides a Nash equilibrium solution, if $\phi_0^i\Big(t_0, \mathcal{H}^i\left(t_0, K_*^1, \ldots, K_*^{i-1}, K_*^i, K_*^{i+1}, \ldots, K_*^N\right),$ $\mathcal{D}^i\left(t_0, K_*^1, \ldots, K_*^{i-1}, K_*^i, K_*^{i+1}, \ldots, K_*^N\right)\Big)$ is at most equal to $\phi_0^i\Big(t_0, \mathcal{H}^i\left(t_0, K_*^1, \ldots, K_*^{i-1}, K^i, K_*^{i+1}, \ldots, K_*^N\right),$ $\mathcal{D}^i(t_0, K_*^1, \ldots, K_*^{i-1}, K^i, K_*^{i+1}, \ldots, K_*^N)\Big)$ where K^i is any admissible feedback gain of i-th decision maker.

When solving for a Nash equilibrium solution, it is very important to realize that N decision makers have different performance indices to minimize. A standard approach for a potential solution from the set of N inequalities as stated above is to solve jointly N optimal control problems defined by these inequalities, each of which depends structurally on the other decision maker's decision laws. However, a Nash equilibrium solution even under the feedback information structure to this class of problems can not be unique due to informational nonuniqueness as indicated in Basar (1982). Therein, problems with informational nonuniqueness under the feedback information pattern and the need for more satisfactory resolution have been addressed via the requirement of the Nash equilibrium solution to have the additional property that its restriction on the interval $[t_0, \alpha]$ is a Nash solution to the truncated version of the original problem, defined on $[t_0, \alpha]$. With such a restriction so defined, the solution is now termed as a feedback Nash equilibrium solution which is now free of any informational nonuniqueness, and thus whose derivation allows a dynamic programming type argument, as proposed in the sequel development.

Definition 5. Feedback Nash Equilibrium Solution.
Let N-tuple $\left(K_*^1, \ldots, K_*^N\right)$ be feedback Nash equilibrium in $\times_{r=1}^N \mathcal{K}^r_{t_f, \mathcal{H}_f^r, \mathcal{D}_f^r; \mu^r}$ and $(\mathcal{H}_*^i, \mathcal{D}_*^i)$ the corresponding trajectory solutions of the dynamic equations associated with i-th decision maker

$$\frac{d}{d\alpha}\mathcal{H}^i(\alpha) = \mathcal{F}^i\left(\alpha, \mathcal{H}^i(\alpha), K^1(\alpha), \ldots, K^N(\alpha)\right)$$
$$\frac{d}{d\alpha}\mathcal{D}^i(\alpha) = \mathcal{G}^i\left(\alpha, \mathcal{H}^i(\alpha)\right).$$

Then, the N-tuple of feedback gains $\left(K_*^1, \ldots, K_*^N\right)$ when restricted to the interval $[t_0, \alpha]$ is still a feedback Nash equilibrium solution for the optimal control problem with the appropriate terminal-value condition $\left(\alpha, \mathcal{H}_*^i(\alpha), \mathcal{D}_*^i(\alpha)\right)$ for all $\alpha \in [t_0, t_f]$.

Next, one may state the corresponding cumulant-based control optimization problems for which the set of static Nash problems needs to be solved for a feedback Nash equilibrium solution and satisfies the already mentioned Nash inequalities. The optimization problem for the multi-person stochastic Nash decision problem is then stated.

Definition 6. Optimization Problem.
Suppose that $k_i \in \mathbb{Z}^+$ is fixed. Under the memoryless perfect-state information pattern, the optimization for i-th decision maker in the multi-person stochastic Nash decision problem over a finite horizon is given by the minimization of $\phi_0^i\left(t_0, \mathcal{H}^i(t_0, K^1, \ldots, K^N), \mathcal{D}^i(t_0, K^1, \ldots, K^N)\right)$ over all $K^i(\cdot) \in \mathcal{K}^i_{t_f, \mathcal{H}_f^i, \mathcal{D}_f^i; \mu^i}$, $\mu^i = \{\mu_r^i \geq 0\}_{r=1}^{k_i}$ with $\mu_1^i > 0$, and subject to (10)-(11) for $\alpha \in [t_0, t_f]$.

Since a direct dynamic programming approach is taken, it is therefore necessary to introduce the value function $\mathcal{V}^i(\varepsilon, \mathcal{Y}^i, \mathcal{Z}^i)$ for i-th decision maker starting at $(\varepsilon, \mathcal{Y}^i, \mathcal{Z}^i)$.

Definition 7. Value Function.
The value function $\mathcal{V}^i : [t_0, t_f] \times (\mathbb{R}^{n \times n})^{k_i} \times \mathbb{R}^{k_i} \mapsto \mathbb{R}^+ \cup \{+\infty\}$ associated with i-th decision maker is defined as $\mathcal{V}^i\left(\varepsilon, \mathcal{Y}^i, \mathcal{Z}^i\right)$ equal to the minimization of $\phi_0^i\left(t_0, \mathcal{H}^i(t_0, K^1, \ldots, K^N), \mathcal{D}^i(t_0, K^1, \ldots, K^N)\right)$ for all $K^i(\cdot) \in \mathcal{K}^i_{\varepsilon, \mathcal{Y}^i, \mathcal{Z}^i; \mu^i}$ and $(\varepsilon, \mathcal{Y}^i, \mathcal{Z}^i) \in [t_0, t_f] \times (\mathbb{R}^{n \times n})^{k_i} \times \mathbb{R}^{k_i}$.

Note that the value function $\mathcal{V}^i\left(\varepsilon, \mathcal{Y}^i, \mathcal{Z}^i\right)$ for $i = 1, \ldots, N$ is supposed to be continuously differentiable in $\left(\varepsilon, \mathcal{Y}^i, \mathcal{Z}^i\right)$.

Definition 8. Actionable Sets.
Let a non-empty actionable set associated with i-th decision maker be defined by the expression $\mathcal{Q}^i \triangleq \Big\{\left(\varepsilon, \mathcal{Y}^i, \mathcal{Z}^i\right) \in [t_0, t_f] \times (\mathbb{R}^{n \times n})^{k_i} \times \mathbb{R}^{k_i} : \mathcal{K}^i_{\varepsilon, \mathcal{Y}^i, \mathcal{Z}^i; \mu^i} \neq 0\Big\}$.

When i-th decision maker is confident that other $N-1$ decision makers choose their Nash equilibrium strategies, i.e., $(K_*^1, \ldots, K_*^{i-1}, K_*^{i+1}, \ldots, K_*^N)$. He then uses his Nash strategy, K_*^i. Thus, the definition of Nash equilibrium is used in an obvious extension of the usual dynamic programming argument. As the result, the value function $\mathcal{V}^i\left(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i\right)$ satisfies the following necessary condition.

Theorem 2. HJB Equation-Mayer Problem.
Let $\left(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i\right)$ be any interior point of the actionable set $\mathcal{Q}^i$ for i-th decision maker at which the value function $\mathcal{V}^i\left(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i\right)$ is differentiable. If there exist a Nash equilibrium strategy set $(K_*^1,\ldots,K_*^N)\in\times_{r=1}^N\mathcal{K}^r_{\varepsilon,\mathcal{Y}^r,\mathcal{Z}^r;\mu^r}$, then the partial differential equation of dynamic programming

$$0=\min_{K^i\in\overline{K}^i}\left\{\frac{\partial}{\partial\varepsilon}\mathcal{V}^i\left(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i\right)\right.+\frac{\partial}{\partial\operatorname{vec}(\mathcal{Z}^i)}\mathcal{V}^i\left(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i\right)\cdot\operatorname{vec}\left(\mathcal{G}^i\left(\varepsilon,\mathcal{Y}^i\right)\right)+\frac{\partial}{\partial\operatorname{vec}(\mathcal{Y}^i)}\mathcal{V}^i\left(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i\right)\cdot\left.\operatorname{vec}\left(\mathcal{F}^i(\varepsilon,\mathcal{Y}^i,K_*^1,\ldots,K_*^{i-1},K^i,K_*^{i+1},\ldots,K_*^N)\right)\right\}\quad(13)$$

is satisfied where $\mathcal{V}^i\left(t_0,\mathcal{H}_0^i,\mathcal{D}_0^i\right)=\phi_0^i\left(t_0,\mathcal{H}_0^i,\mathcal{D}_0^i\right)$.

Finally, the sufficient condition used to verify the Nash strategy of i-th decision maker is given in the sequel.

Theorem 3. Verification Theorem.
Fix $k_i\in\mathbb{Z}^+$ and let $\mathcal{W}^i(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i)$ be continuously differentiable solution of the HJB equation (13) which satisfies the boundary condition

$$\mathcal{W}^i(t_0,\mathcal{H}_0^i,\mathcal{D}_0^i)=\phi_0^i\left(t_0,\mathcal{H}_0^i,\mathcal{D}_0^i\right).\quad(14)$$

Let 3-tuple $\left(t_f,\mathcal{H}_f^i,\mathcal{D}_f^i\right)$ be point of $\mathcal{Q}^i$; strategy set $(K_*^1,\ldots,K_*^{i-1},K^i,K_*^{i+1},\ldots,K_*^N)$ in $\times_{r=1}^N\mathcal{K}^r_{t_f,\mathcal{H}_f^r,\mathcal{D}_f^r;\mu^r}$; and the i-th decision maker corresponding solutions $(\mathcal{H}^i,\mathcal{D}^i)$ of the equations (10)-(11). $\mathcal{W}^i\left(\alpha,\mathcal{H}^i(\alpha),\mathcal{D}^i(\alpha)\right)$ is then a time-backward increasing function of α.

If $(K_*^1,\ldots,K_*^{i-1},K_*^i,K_*^{i+1},\ldots,K_*^N)$ is a set of strategies in $\times_{r=1}^N\mathcal{K}^r_{t_f,\mathcal{H}_f^r,\mathcal{D}_f^r;\mu^r}$ defined on $[t_0,t_f]$ with corresponding i-th decision maker solutions, $\left(\mathcal{H}_*^i,\mathcal{D}_*^i\right)$ of the preceding equations such that for $\alpha\in[t_0,t_f]$

$$0=\frac{\partial}{\partial\varepsilon}\mathcal{W}^i\left(\alpha,\mathcal{H}_*^i(\alpha),\mathcal{D}_*^i(\alpha)\right)+\frac{\partial}{\partial\operatorname{vec}(\mathcal{Y}^i)}\mathcal{W}^i\left(\alpha,\mathcal{H}_*^i(\alpha),\mathcal{D}_*^i(\alpha)\right)\cdot\operatorname{vec}\left(\mathcal{F}^i\left(\alpha,\mathcal{H}_*^i(\alpha),K_*^1(\alpha),\ldots,K_*^N(\alpha)\right)\right)+\frac{\partial}{\partial\operatorname{vec}(\mathcal{Z}^i)}\mathcal{W}^i\left(\alpha,\mathcal{H}_*^i(\alpha),\mathcal{D}_*^i(\alpha)\right)\operatorname{vec}\left(\mathcal{G}^i\left(\alpha,\mathcal{H}_*^i(\alpha)\right)\right)\quad(15)$$

then K_*^i is the Nash strategy in $\mathcal{K}^i_{t_f,\mathcal{H}_f^i,\mathcal{D}_f^i;\mu^i}$ and

$$\mathcal{W}^i\left(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i\right)=\mathcal{V}^i\left(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i\right)\quad(16)$$

where $\mathcal{V}^i\left(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i\right)$ is the value function associated with i-th decision maker.

4. NON-COOPERATIVE DECISION STRATEGIES

Recall that the optimization problem being considered herein is in "Mayer form" and can be solved by applying an adaptation of the Mayer form verification theorem of dynamic programming given in Fleming (1975). In the framework of dynamic programming, the i-th decision maker is often required to denote the terminal time and states of a family of optimization problems as $(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i)$ rather than $(t_f,\mathcal{H}_f^i,\mathcal{D}_f^i)$. That is, for $\varepsilon\in[t_0,t_f]$ and $1\le l\le k_i$, the states of the system (10)-(11) defined on the interval $[t_0,\varepsilon]$ have the terminal values denoted by $\mathcal{H}^i(\varepsilon)\equiv\mathcal{Y}^i$ and $\mathcal{D}^i(\varepsilon)\equiv\mathcal{Z}^i$. Since the performance index (12) is quadratic affine in terms of arbitrarily fixed x_0, this observation suggests a solution to the HJB equation (13) is of the form as follows.

Theorem 4. Candidate Value-Function.
Fix $k_i\in\mathbb{Z}^+$ and let $(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i)$ be any interior point of the actionable set $\mathcal{Q}^i$ at which the real-valued function

$$\mathcal{W}^i(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i)=x_0^T\sum_{l=1}^{k_i}\mu_l^i\left(\mathcal{Y}_l^i+\mathcal{E}_l^i(\varepsilon)\right)x_0+\sum_{l=1}^{k_i}\mu_l^i\left(\mathcal{Z}_l^i+\mathcal{T}_l^i(\varepsilon)\right)\quad(17)$$

is differentiable. The parametric functions of time $\mathcal{E}_l^i\in\mathcal{C}^1([t_0,t_f];\mathbb{R}^{n\times n})$ and $\mathcal{T}_l^i\in\mathcal{C}^1([t_0,t_f];\mathbb{R})$ are yet to be determined for i-th decision maker. Furthermore, the time derivative of $\mathcal{W}^i(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i)$ is given by

$$\frac{d}{d\varepsilon}\mathcal{W}^i(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i)=\sum_{l=1}^{k_i}\mu_l^i\left(\mathcal{G}_l^i(\varepsilon,\mathcal{Y}^i)+\frac{d}{d\varepsilon}\mathcal{T}_l^i(\varepsilon)\right)+x_0^T\sum_{l=1}^{k_i}\mu_l^i\left(\mathcal{F}_l^i(\varepsilon,\mathcal{Y}^i,K^1,\ldots,K^N)+\frac{d}{d\varepsilon}\mathcal{E}_l^i(\varepsilon)\right)x_0.\quad(18)$$

The substitution of this hypothesized solution (17) into the Hamilton-Jacobi-Bellman (HJB) equation (13) and making use of the result (18) yields the necessary condition

$$0\equiv\min_{K^i\in\overline{K}^i}\left\{x_0^T\left(\sum_{l=1}^{k_i}\mu_l^i\frac{d}{d\varepsilon}\mathcal{E}_l^i(\varepsilon)\right)x_0+\sum_{l=1}^{k_i}\mu_l^i\frac{d}{d\varepsilon}\mathcal{T}_l^i(\varepsilon)+\sum_{l=1}^{k_i}\mu_l^i\mathcal{G}_l^i(\varepsilon,\mathcal{Y}^i)+x_0^T\left(\sum_{l=1}^{k_i}\mu_l^i\mathcal{F}_l^i(\varepsilon,\mathcal{Y}^i,K^1,\ldots,K^N)\right)x_0\right\}.\quad(19)$$

Differentiating the expression within the bracket of (19) with respect to K^i yields the necessary condition for an extremum of (12) on $[t_0,\varepsilon]$,

$$-2(B^i)^T(\varepsilon)\sum_{l=1}^{k_i}\mu_l^i\mathcal{Y}_l^iM_0-2\mu_1^iR^{ii}(\varepsilon)K^iM_0=0.$$

Clearly, the matrix $M_0\triangleq x_0x_0^T$ is arbitrarily rank-one, it must be true that

$$K^i(\varepsilon,\mathcal{Y}^i,\mathcal{Z}^i)=-(R^{ii})^{-1}(\varepsilon)(B^i)^T(\varepsilon)\sum_{r=1}^{k_i}\widehat{\mu}_r^i\mathcal{Y}_r^i\quad(20)$$

where $\widehat{\mu}_r^i\triangleq\mu_l^i/\mu_1^i$ for $\mu_1^i>0$. Substituting the gain expressions (20) into the right member of the HJB equation (13) yields the value of the minimum

$$x_0^T\Bigg\{\sum_{l=1}^{k_i}\mu_l^i\frac{d}{d\varepsilon}\mathcal{E}_l^i(\varepsilon)-A^T(\varepsilon)\sum_{l=1}^{k_i}\mu_l^i\mathcal{Y}_l^i-\sum_{l=1}^{k_i}\mu_l^i\mathcal{Y}_l^iA(\varepsilon)$$
$$+\left[\sum_{j=1}^{N}\sum_{r=1}^{k_j}\widehat{\mu}_r^j\mathcal{Y}_r^jB^j(\varepsilon)(R^{jj})^{-1}(\varepsilon)(B^j)^T(\varepsilon)\right]\sum_{l=1}^{k_i}\mu_l^i\mathcal{Y}_l^i$$
$$+\sum_{l=1}^{k_i}\mu_l^i\mathcal{Y}_l^i\left[\sum_{j=1}^{N}B^j(\varepsilon)(R^{jj})^{-1}(\varepsilon)(B^j)^T(\varepsilon)\sum_{s=1}^{k_j}\widehat{\mu}_s^j\mathcal{Y}_s^j\right]$$
$$-\mu_1^iQ^i(\varepsilon)-\mu_1\sum_{j=1}^{N}\left[\sum_{r=1}^{k_j}\widehat{\mu}_r^j\mathcal{Y}_r^jB^j(\varepsilon)(R^{jj})^{-1}(\varepsilon)\cdot\right.$$
$$\left.R^{ij}(\varepsilon)(R^{jj})^{-1}(\varepsilon)(B^j)^T(\varepsilon)\sum_{s=1}^{k_j}\widehat{\mu}_s^j\mathcal{Y}_s^j\right]$$
$$-\sum_{l=2}^{k_i}\mu_l^i\sum_{q=1}^{l-1}\frac{2l!}{q!(l-q)!}\mathcal{Y}_q^iG(\varepsilon)WG^T(\varepsilon)\mathcal{Y}_{l-q}^i\Bigg\}x_0$$
$$+\sum_{l=1}^{k_i}\mu_l^i\frac{d}{d\varepsilon}\mathcal{T}_l^i(\varepsilon)-\sum_{l=1}^{k_i}\mu_l^i\mathrm{Tr}\left\{\mathcal{Y}_l^iG(\varepsilon)WG^T(\varepsilon)\right\}. \quad (21)$$

It is now necessary to exhibit $\left\{\mathcal{E}_p^i(\cdot)\right\}_{p=1}^{k_i}$ and $\left\{\mathcal{T}_p^i(\cdot)\right\}_{p=1}^{k_i}$ which render the left side of (21) equal to zero for $\varepsilon\in[t_0,t_f]$, when $\{\mathcal{Y}_p^i\}_{p=1}^{k_i}$ are evaluated along the solution trajectories.

Studying the expression (21) reveals that $\mathcal{E}_p^i(\cdot)$ and $\mathcal{T}_p^i(\cdot)$ for $1\le p\le k_i$ satisfying the differential equations

$$\frac{d}{d\varepsilon}\mathcal{E}_1^i(\varepsilon)=A^T(\varepsilon)\mathcal{H}_1^i(\varepsilon)+\mathcal{H}_1^i(\varepsilon)A(\varepsilon)+Q^i(\varepsilon)$$
$$-\mathcal{H}_1^i(\varepsilon)\left[\sum_{j=1}^{N}B^j(\varepsilon)(R^{jj})^{-1}(\varepsilon)(B^j)^T(\varepsilon)\sum_{s=1}^{k_j}\widehat{\mu}_s^j\mathcal{H}_s^j(\varepsilon)\right]$$
$$-\left[\sum_{j=1}^{N}\sum_{r=1}^{k_j}\widehat{\mu}_r^j\mathcal{H}_r^j(\varepsilon)B^j(\varepsilon)(R^{jj})^{-1}(\varepsilon)(B^j)^T(\varepsilon)\right]\mathcal{H}_1^i(\varepsilon)$$
$$+\sum_{j=1}^{N}\left[\sum_{r=1}^{k_j}\widehat{\mu}_r^j\mathcal{H}_r^j(\varepsilon)B^j(\varepsilon)(R^{jj})^{-1}(\varepsilon)\cdot\right.$$
$$\left.R^{ij}(\varepsilon)(R^{jj})^{-1}(\varepsilon)(B^j)^T(\varepsilon)\sum_{s=1}^{k_j}\widehat{\mu}_s^j\mathcal{H}_s^j(\varepsilon)\right] \quad (22)$$

and, for $2\le p\le k_i$

$$\frac{d}{d\varepsilon}\mathcal{E}_p(\varepsilon)=A^T(\varepsilon)\mathcal{H}_p^i(\varepsilon)+\mathcal{H}_p^i(\varepsilon)A(\varepsilon)$$
$$-\mathcal{H}_p^i(\varepsilon)\left[\sum_{j=1}^{N}B^j(\varepsilon)(R^{jj})^{-1}(\varepsilon)(B^j)^T(\varepsilon)\sum_{s=1}^{k_j}\widehat{\mu}_s^j\mathcal{H}_s^j(\varepsilon)\right]$$
$$-\left[\sum_{j=1}^{N}\sum_{r=1}^{k_j}\widehat{\mu}_r^j\mathcal{H}_r^j(\varepsilon)B^j(\varepsilon)(R^{jj})^{-1}(\varepsilon)(B^j)^T(\varepsilon)\right]\mathcal{H}_p^i(\varepsilon)$$
$$+\sum_{q=1}^{p-1}\frac{2p!}{q!(p-q)!}\mathcal{H}_q^i(\varepsilon)G(\varepsilon)WG^T(\varepsilon)\mathcal{H}_{p-q}^i(\varepsilon)\,, \quad (23)$$

together with, for $1\le p\le k_i$

$$\frac{d}{d\varepsilon}\mathcal{T}_p^i(\varepsilon)=\mathrm{Tr}\left\{\mathcal{H}_p^i(\varepsilon)G(\varepsilon)WG^T(\varepsilon)\right\} \quad (24)$$

will work. Furthermore, at the boundary condition, it is necessary to have $\mathcal{W}^i\left(t_0,\mathcal{H}_0^i,\mathcal{D}_0^i\right)=\phi_0^i\left(t_0,\mathcal{H}_0^i,\mathcal{D}_0^i\right)$. Or equivalently, it yields

$$x_0^T\sum_{l=1}^{k_i}\mu_l(\mathcal{H}_{l0}^i+\mathcal{E}_l^i(t_0))\,x_0+\sum_{l=1}^{k_i}\mu_l^i(\mathcal{D}_{l0}^i+\mathcal{T}_l^i(t_0))=$$
$$x_0^T\sum_{l=1}^{k_i}\mu_l^i\mathcal{H}_{l0}^ix_0+\sum_{l=1}^{k_i}\mu_l^i\mathcal{D}_{l0}^i\,.$$

Thus, matching the boundary condition yields the corresponding initial value conditions $\mathcal{E}_p(t_0)=0$ and $\mathcal{T}_p(t_0)=0$ for the equations (22)-(24). Applying the feedback gains specified in (20) along the solution trajectories of the equations (10)-(11), these equations become the time-backward Riccati-type differential equations

$$\frac{d}{d\varepsilon}\mathcal{H}_1^i(\varepsilon)=-A^T(\varepsilon)\mathcal{H}_1^i(\varepsilon)-\mathcal{H}_1^i(\varepsilon)A(\varepsilon)-Q^i(\varepsilon)$$
$$+\mathcal{H}_1^i(\varepsilon)\left[\sum_{j=1}^{N}B^j(\varepsilon)(R^{jj})^{-1}(\varepsilon)(B^j)^T(\varepsilon)\sum_{s=1}^{k_j}\widehat{\mu}_s^j\mathcal{H}_s^j(\varepsilon)\right]$$
$$+\left[\sum_{j=1}^{N}\sum_{r=1}^{k_j}\widehat{\mu}_r^j\mathcal{H}_r^j(\varepsilon)B^j(\varepsilon)(R^{jj})^{-1}(\varepsilon)(B^j)^T(\varepsilon)\right]\mathcal{H}_1^i(\varepsilon)$$
$$-\sum_{j=1}^{N}\left[\sum_{r=1}^{k_j}\widehat{\mu}_r^j\mathcal{H}_r^j(\varepsilon)B^j(\varepsilon)(R^{jj})^{-1}(\varepsilon)\cdot\right.$$
$$\left.R^{ij}(\varepsilon)(R^{jj})^{-1}(\varepsilon)(B^j)^T(\varepsilon)\sum_{s=1}^{k_j}\widehat{\mu}_s^j\mathcal{H}_s^j(\varepsilon)\right] \quad (25)$$

and, for $2\le p\le k_i$

$$\frac{d}{d\varepsilon}\mathcal{H}_p^i(\varepsilon)=-A^T(\varepsilon)\mathcal{H}_p^i(\varepsilon)-\mathcal{H}_p^i(\varepsilon)A(\varepsilon)$$
$$+\mathcal{H}_p^i(\varepsilon)\left[\sum_{j=1}^{N}B^j(\varepsilon)(R^{jj})^{-1}(\varepsilon)(B^j)^T(\varepsilon)\sum_{s=1}^{k_j}\widehat{\mu}_s^j\mathcal{H}_s^j(\varepsilon)\right]$$
$$+\left[\sum_{j=1}^{N}\sum_{r=1}^{k_j}\widehat{\mu}_r^j\mathcal{H}_r^j(\varepsilon)B^j(\varepsilon)(R^{jj})^{-1}(\varepsilon)(B^j)^T(\varepsilon)\right]\mathcal{H}_p^i(\varepsilon)$$
$$-\sum_{q=1}^{p-1}\frac{2p!}{q!(p-q)!}\mathcal{H}_q^i(\varepsilon)G(\varepsilon)WG^T(\varepsilon)\mathcal{H}_{p-q}^i(\varepsilon)\,, \quad (26)$$

together with, for $1\le p\le k_i$

$$\frac{d}{d\varepsilon}\mathcal{D}_p^i(\varepsilon)=-\mathrm{Tr}\left\{\mathcal{H}_p^i(\varepsilon)G(\varepsilon)WG^T(\varepsilon)\right\} \quad (27)$$

where the terminal-value conditions $\mathcal{H}_1^i(t_f)=Q_f^i$, $\mathcal{H}_p^i(t_f)=0$ for $2\le p\le k_i$ and $\mathcal{D}_p^i(t_f)=0$ for $1\le p\le k_i$. Thus, whenever these equations (25)-(27) admit solutions $\{\mathcal{H}_p^i(\cdot)\}_{p=1}^{k_i}$ and $\{\mathcal{D}_p^i(\cdot)\}_{p=1}^{k_i}$, then the existence of $\{\mathcal{E}_p(\cdot)\}_{p=1}^{k_i}$ and $\{\mathcal{T}_p^i(\cdot)\}_{p=1}^{k_i}$ satisfying the equations (22)-(24) are assured. By comparing equations (22)-(24) to those of (25)-(27), one may recognize that these sets of equations are related to one another by $\frac{d}{d\varepsilon}\mathcal{E}_p^i(\varepsilon)=-\frac{d}{d\varepsilon}\mathcal{H}_p^i(\varepsilon)$ and $\frac{d}{d\varepsilon}\mathcal{T}_p^i(\varepsilon)=-\frac{d}{d\varepsilon}\mathcal{D}_p^i(\varepsilon)$ for $1\le p\le k_i$. Enforcing the initial value conditions of $\mathcal{E}_p^i(t_0)=0$ and $\mathcal{T}_p^i(t_0)=0$ uniquely implies that $\mathcal{E}_p^i(\varepsilon)=\mathcal{H}_p^i(t_0)-\mathcal{H}_p^i(\varepsilon)$ and $\mathcal{T}_p^i(\varepsilon)=\mathcal{D}_p^i(t_0)-\mathcal{D}_p^i(\varepsilon)$ for all $\varepsilon\in[t_0,t_f]$ and yields a value function $\mathcal{W}^i(\varepsilon, \mathcal{Y}^i, \mathcal{Z}^i) = x_0^T \sum_{l=1}^{k_i} \mu_l^i \mathcal{H}_l^i(t_0) x_0 + \sum_{l=1}^{k_i} \mu_l^i \mathcal{D}_l^i(t_0)$ for which the sufficient condition (15) of the verification theorem is satisfied. Therefore, the decision laws (20) minimizing (12) become optimal

$$K_*^i(\varepsilon) = -(R^{ii})^{-1}(\varepsilon)(B^i)^T(\varepsilon) \sum_{r=1}^{k_i} \widehat{\mu}_r^i \mathcal{H}_{r*}^i(\varepsilon). \quad (28)$$

Theorem 5. Multi-Cumulant Nash Decision Strategies. Consider the stochastic Nash game (4)-(5) against Nature whose $(A, [B^1 \dots B^N])$ is uniformly stabilizable on $[t_0, t_f]$. Let $k_i \in \mathbb{Z}^+$ and the sequence $\mu^i = \{\mu_r^i \geq 0\}_{r=1}^{k_i}$ with $\mu_1^i > 0$. Then the optimal decision law associated with i-th decision maker is given by

$$K_*^i(\alpha) = -(R^{ii})^{-1}(\alpha)(B^i)^T(\alpha) \sum_{r=1}^{k_i} \widehat{\mu}_r^i \mathcal{H}_{r*}^i(\alpha) \quad (29)$$

where $\widehat{\mu}_r^i \triangleq \mu_l^i / \mu_1^i$ are mutually chosen by i-th decision maker for different levels of robustness on his own performance and $\{\mathcal{H}_{r*}^i(\alpha)\}_{r=1}^{k_i}$ are optimal solutions of the time-backward differential equations

$$\frac{d}{d\alpha}\mathcal{H}_{1*}^i(\alpha) = -\left[A(\alpha) + \sum_{j=1}^N B^j(\alpha) K_*^j(\alpha)\right]^T \mathcal{H}_{1*}^i(\alpha) - \mathcal{H}_{1*}^i(\alpha)\left[A(\alpha) + \sum_{j=1}^N B^j(\alpha) K_*^i(\alpha)\right] - Q^i(\alpha) - \sum_{j=1}^N (K_*^j)^T(\alpha) R^{ij}(\alpha) K_*^j(\alpha)\,, \quad (30)$$

and, for $2 \leq r \leq k_i$

$$\frac{d}{d\alpha}\mathcal{H}_{r*}^i(\alpha) = -\left[A(\alpha) + \sum_{j=1}^N B^j(\alpha) K_*^j(\alpha)\right]^T \mathcal{H}_{r*}^i(\alpha) - \mathcal{H}_{r*}^i(\alpha)\left[A(\alpha) + \sum_{j=1}^N B^j(\alpha) K_*^j(\alpha)\right] - \sum_{s=1}^{r-1} \frac{2r!}{s!(r-s)!} \mathcal{H}_{s*}^i(\alpha) G(\alpha) W G^T(\alpha) \mathcal{H}_{r-s,*}^i(\alpha) \quad (31)$$

with the terminal-value conditions $\mathcal{H}_{1*}^i(t_f) = Q_f^i$, and $\mathcal{H}_{r*}^i(t_f) = 0$ when $2 \leq r \leq k_i$.

Remark 2. It is observed that to have the Nash strategy (29) of i-th decision maker be defined and continuous for all $\alpha \in [t_0, t_f]$, the solutions $\mathcal{H}_{r*}^i(\alpha)$ to the equations (30)-(31) when evaluated at $\alpha = t_0$ must also exist. Therefore, it is necessary that $\mathcal{H}_{r*}^i(\alpha)$ are finite for all $\alpha \in [t_0, t_f)$. Moreover, the solutions of (30)-(31) exist and are continuously differentiable in a neighborhood of t_f. Applying a result from Dieudonne (1960), these solutions can further be extended to the left of t_f as long as $\mathcal{H}_{r*}^i(\alpha)$ remain finite. Hence, the existences of unique and continuously differentiable solutions to the equations (30)-(31) are certain if $\mathcal{H}_{r*}^i(\alpha)$ are bounded for all $\alpha \in [t_0, t_f)$. As the result, the candidate value functions $\mathcal{W}^i(\alpha, \mathcal{H}^i, \mathcal{D}^i)$ are continuously differentiable as well.

Theorem 6. Necessary and Sufficient Conditions. $\left(K_*^1\left(\alpha, \mathcal{H}_*^1\right), \dots, K_*^N\left(\alpha, \mathcal{H}_*^N\right)\right)$ is a Nash equilibrium strategy set if and only if $\left(\mathcal{H}_*^1(\alpha), \dots, \mathcal{H}_*^N(\alpha)\right)$ is bounded for all $\alpha \in [t_0, t_f)$.

5. CONCLUSIONS

This research article presents a new innovative and generalized solution concept to address multiple resolutions of performance robustness often seen as unresolved issues in stochastic control and Linear-Quadratic (LQ) decision problems. Within the cumulant-based control framework for decision-making, rational decision makers have ability to maintain preferable shapes of realized performance indices with respect to Nature's mixed random realizations via multi-cumulant based actions that now have access to the current time and interaction states. These decentralized feedback Nash gains (29) operate dynamically on the time-backward histories of the cumulant-supporting equations (30)-(31) from the final to the current time. Interestingly enough, it is noted that these cumulant-supporting equations also depend on Nature's a-priori probabilistic characteristics. Thus, non-cooperative decision makers employing optimal decision gains (29) have purposely traded the property of the certainty equivalence principle as they would obtain from the special case of LQ decision problems, for the adaptability to deal with uncertain environments. Future work will be a next attainment of cumulant-based solutions in multi-person decision problems for complex situations: (i) performance robustness via a desired statistical description, (ii) non-cooperative decision selection via output feedback information patterns, and (iii) confrontations among rational decision makers.

REFERENCES

T. Basar and G.J. Olsder. *Dynamic Noncooperative Game Theory.* Academic Press, London, 1982.

J.B. Cruz, Jr., M.A. Simaan, A. Gacic, H. Jiang, B. Letellier, M. Li, and Y. Liu. Game-theoretic modeling and control of military operations. *IEEE Transactions on Aerospace & Electronic Systems,* 37:1393-1405, 2001.

J. Dieudonne. *Foundations of Modern Analysis.* Academic Press, New York and London, 1960.

W.H. Fleming and R.W. Rishel. *Deterministic and Stochastic Optimal Control.* Springer-Verlag, 1975.

L.P. Hansen, T.J. Sargent and T.D. Tallarini, Jr. Robust permanent income and pricing. *Review of Economic Studies,* 66:873-907, 1999.

K.D. Pham, M.K. Sain, and S.R. Liberty. Statistical control for smart base-isolated buildings via cost cumulants and output feedback paradigm. *Proceedings of the American Control Conference,* pages 3090-3095, 2005.

K.D. Pham. Cost cumulant-based control for a class of linear-quadratic tracking problems. *Proceedings of the American Control Conference,* pages 335-340, 2007.

K.D. Pham. Cooperative solutions in multi-person quadratic decision problems: performance-measure statistics and cost-cumulant control paradigm. *The 46th Conference on Decisions and Control,* pages 2484-2490.

K.D. Pham. New results in stochastic cooperative games: strategic coordination for multi-resolution performance robustness. Series Lecture Notes in Control and Information Sciences. M.J. Hirsch; P.M. Pardalos; R. Murphey; D. Grundel, editors, *Advances in Cooperative Control and Optimization*, 2008.