

A Class of Mean-Risk Decisions for Noncooperative Games and Distributed Controls

Khanh D. Pham¹ and Alena Gubar²

Abstract—The paper gives a comprehensive presentation of the broad and still developing area of risk-averse decision making in the stochastic linear-quadratic class of multiperson noncooperative differential games and/or distributed controls. The fundamental set of design principles is developed for self-directed decision makers, who are imitatively cautious and privately capable of incorporating summary performance-measure statistics into their distributed output-feedback Nash decision policies for competitive interactions and engagements.

I. INTRODUCTION

In noncooperative stochastic games and distributed controls, there are more than two capable decision makers who optimize different goals and utilities. Each decision maker wishes to influence to his advantage a shared interaction process by exerting his control decisions. To the best knowledge of the authors, most studies, e.g., [1] and [2] have mainly concentrated on the selection of open and/or closed-loop Nash strategy equilibria in accordance of expected utilities under the structural constraints of linear system dynamics, quadratic cost functionals, and additive independent white Gaussian noises corrupting the system dynamics and measurements. Very little work, if any, has been published on the subject of higher-order assessment of performance uncertainty and risks beyond expected performance.

For this reason attention in the research investigation that follows, is directed primarily towards a linear-quadratic class of noncooperative stochastic games and/or distributed controls, which in turn has linear system dynamics, quadratic rewards and/or costs and independent white zero-mean Gaussian noises additively corrupting the system dynamics and output measurements. Notice that, under these conditions, the quadratic rewards or costs are random variables with the generalized chi-squared probability distributions. If a measure of uncertainty such as the variance of the possible rewards or costs was used in addition to the expected reward or costs, the decision makers should be able to correctly order preferences for alternatives. This claim seems plausible, but it is not always correct. Various investigations have indicated that any evaluation scheme based on just the expected reward or cost and reward/cost variance would necessarily imply indifference between some courses of action; therefore, no criterion based solely on the two attributes of means and variances can correctly represent their preferences. See [3] and [4] for further details.

¹ The Air Force Research Laboratory, Space Vehicles Directorate, 3550 Aberdeen Ave, S.E., Kirtland Air Force Base, New Mexico 87117 U.S.A. Email: AFRL.RVSV@kirtland.af.mil

² St. Petersburg State University, Faculty of Applied Mathematics and Control Processes, Universitetskii prospekt 35, Petergof, Saint-Petersburg, Russia, 198504. Email: alyona.gubar@gmail.com

The present research contributions include significant extensions of the existing results [5] toward some completely unexplored areas as such: i) the design of distributed filtering via private observations for self-directed decision makers and/or autonomous controllers with distributed noisy information structures about the uncertain interaction process; ii) an efficiently computational procedure for all the mathematical statistics associated with the generalized chi-squared rewards/costs when respective mean-risk aware utilities are formed; and iii) the synthesis of distributed risk-sensitive decision policies with output feedback for distributed noncooperative solutions of Nash type that now guarantee performance robustness with certainty much stronger than a stochastically averaging measure of performance.

II. PERFORMANCE-MEASURE STATISTICS

In this section, some preliminaries are in order. Some spaces of random variables and stochastic processes are introduced: $L^2_{\mathcal{F}_t}(\Omega; \mathbb{R}^n) \triangleq \{\varphi : \Omega \mapsto \mathbb{R}^n \text{ such that } \varphi \text{ is } \mathcal{F}_t\text{-measurable and } E\{|\varphi|^2\} < \infty\}$. In addition, $L^2_{\mathcal{F}}([t_0, t_f]; \mathbb{R}^k) \triangleq \{\tilde{h} : [t_0, t_f] \times \Omega \mapsto \mathbb{R}^k \text{ such that } \tilde{h}(\cdot) \text{ is } \mathbb{F} = \{\mathcal{F}_t\}_{t \geq t_0} \geq 0\text{-adapted and } E\{\int_{t_0}^{t_f} \|\tilde{h}(t)\|^2 dt\} < \infty\}$.

As for a model specification, a noncooperative stochastic game or distributed control problem is considered with a set $\bar{N} \triangleq \{1, 2, \dots, N\}$ of self-directed decision makers, who are risk averse and controlling decision policies u^1, \dots, u^N . The length of horizon, on which the interaction dynamics evolve, is specified by $[t_0, t_f]$. The interaction process together with environmental uncertainty is governed by the controlled stochastic differential equation modeled on $[t_0, t_f]$

$$dx(t) = (A(t)x(t) + \sum_{i=1}^N B^i(t)u^i(t))dt + G(t)dw(t), \quad (1)$$

where the continuous coefficients $A \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{n \times n})$, $B^i \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{n \times m_i})$, and $G \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{n \times p})$ are deterministic matrix-valued functions. The initial interaction state at time $t = t_0$, $x(t_0) \triangleq x_0 \in \mathbb{R}^n$ is fixed.

Practical situations, where self-autonomy is possible require that each decision maker be able to possess the common knowledge of the parameters associated with the noncooperative interactions (1). Viewed from the mutual influence of one decision maker to those of others, self-autonomy preferred by decision maker i and $i \in \bar{N}$ is therefore approximated by a surrogate model with $x_0^i = x_0$

$$dx^i(t) = (A(t)x^i(t) + B^i(t)u^i(t) + \sum_{j=1, j \neq i}^N B^j(t)u^j(t))dt + G(t)dw(t), \quad x^i(t_0) = x_0^i, \quad (2)$$

from which each decision maker i can presumably observe a noise corrupted version of all best responses $\sum_{j=1, j \neq i}^N B^j(t)u^j(t)$ from the immediate neighbors. More specifically, the private observations are locally available

$$du^{-i}(t) \triangleq u^{-i}(t)dt = \sum_{j=1, j \neq i}^N B^j(t)u^j(t)dt + d\eta^i(t). \quad (3)$$

Furthermore, information accrued at decision maker i about the coupling interactions (2) is assumed to take on the form of a linear transformation $C^i \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{q_i \times n})$ of the states $x^i(t)$ through the local data $\{y^i(\tau) : \tau \in [t_0, t]\}$

$$dy^i(t) = C^i(t)x^i(t)dt + dv^i(t). \quad (4)$$

Notice that decision makers i operate within their own noisy environments modeled by the filtered probability spaces that are now defined with the uncorrelated p , n and q_i -dimensional stationary Wiener processes adapted for $[t_0, t_f]$

$$\begin{aligned} E\{[w(\tau_1) - w(\tau_2)][w(\tau_1) - w(\tau_2)]^T\} &= W|\tau_1 - \tau_2| \\ E\{[\eta^i(\tau_1) - \eta^i(\tau_2)][\eta^i(\tau_1) - \eta^i(\tau_2)]^T\} &= N^i|\tau_1 - \tau_2| \\ E\{[v^i(\tau_1) - v^i(\tau_2)][v^i(\tau_1) - v^i(\tau_2)]^T\} &= V^i|\tau_1 - \tau_2| \end{aligned}$$

whose a-priori second-order statistics $W > 0$, $N^i > 0$ and $V^i > 0$ for $i \in \bar{N}$ are also assumed known.

At this point, each of the N distributed filters whose the outputs are the state estimates $\hat{x}^i(t)$ of (2) has the form

$$\begin{aligned} d\hat{x}^i(t) &= (A(t)\hat{x}^i(t) + B^i(t)u^i(t) + u^{-i}(t))dt \\ &+ L^i(t)(dy^i(t) - C^i(t)\hat{x}^i(t)dt), \quad \hat{x}^i(t_0) = x_0, \end{aligned} \quad (5)$$

where the local filter gain $L^i(t)$ is given by

$$L^i(t) = \Sigma^i(t)(C^i)^T(t)(V^i)^{-1} \quad (6)$$

and

$$\begin{aligned} \frac{d}{dt}\Sigma^i(t) &= A(t)\Sigma^i(t) + \Sigma^i(t)A^T(t) + G(t)WG^T(t) + N^i \\ &- \Sigma^i(t)(C^i)^T(t)(V^i)^{-1}C^i(t)\Sigma^i(t), \quad \Sigma^i(t_0) = 0. \end{aligned} \quad (7)$$

Under the definition of the estimate errors, e.g., $\tilde{x}^i(t) \triangleq x^i(t) - \hat{x}^i(t)$, it can be shown that

$$\begin{aligned} d\tilde{x}^i(t) &= (A(t) - L^i(t)C^i(t))\tilde{x}^i(t)dt \\ &+ G(t)dw(t) - L^i(t)dv^i(t) - d\eta^i(t), \quad \tilde{x}^i(t_0) = 0. \end{aligned} \quad (8)$$

Recall that decision maker i is assumed to act purely on the basis of his own information. It is now necessary to specify the set of admissible feedback policies, $\mathcal{U}^i \subset L^2_{\mathcal{F}_t^i}([t_0, t_f]; \mathbb{R}^{m_i})$ —the subset of Hilbert space of \mathbb{R}^{m_i} -valued square integrable processes on $[t_0, t_f]$ that are adapted to the σ -algebra \mathcal{F}_t^i generated by $\{y^i(\tau) : \tau \in [t_0, t]\}$ in which risk-bearing decisions u^i are sought. Associated with each admissible 2-tuple $(u^i(\cdot), u^{-i}(\cdot))$ is the performance measure with a generalized chi-squared random variable

$$\begin{aligned} J^i(u^i, u^{-i}) &= (x^i)^T(t_f)Q_f^i x^i(t_f) \\ &+ \int_{t_0}^{t_f} [(x^i)^T(\tau)Q^i(\tau)x^i(\tau) + (u^i)^T(\tau)R^i(\tau)u^i(\tau) \\ &- (u^{-i})^T(\tau)M^i(\tau)u^{-i}(\tau)]d\tau, \end{aligned} \quad (9)$$

where the design parameters $Q_f^i \in \mathbb{R}^{n \times n}$, $Q^i \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{n \times n})$, $M^i \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{n \times n})$ and $R^i \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{m_i \times m_i})$ allow to weight differently the various terminal and transient tradeoffs between the regulatory of x^i , the effectiveness of the control and/or decision policy u^i and observable variations in the control and/or decision policies of all other neighbors u^{-i} . For technical reasons, Q_f^i , $Q^i(t)$ and $R^i(t)$ are positive semidefinite with $R^i(t)$ invertible.

The notion of admissible feedback policy sets is discussed next. In the case of incomplete information, an admissible feedback policy u^i for local best response to all other decision makers u^{-i} must be of the form, for some $\bar{\partial}^i(\cdot, \cdot)$

$$u^i(t) = \bar{\partial}^i(t, y^i(\tau)), \quad \tau \in [t_0, t]. \quad (10)$$

In general, the conditional density $p^i(x^i(t)|\mathcal{F}_t^i)$, which is the density of $x^i(t)$ conditioned on \mathcal{F}_t^i (i.e., induced by the observation $\{y^i(\tau) : \tau \in [t_0, t]\}$) represents the sufficient statistics for describing the conditional stochastic effects of future feedback policy u^i . Under the linear-Gaussian assumption the conditional density $p^i(x^i(t)|\mathcal{F}_t^i)$ is parameterized by the locally available estimate $\hat{x}^i(t) \triangleq E\{x^i(t)|\mathcal{F}_t^i\}$ and estimate error covariance $\Sigma^i(t) \triangleq E\{[x^i(t) - \hat{x}^i(t)][x^i(t) - \hat{x}^i(t)]^T|\mathcal{F}_t^i\}$. In addition, $\Sigma^i(t)$ is independent of feedback policy $u^i(t)$ and observations $\{y^i(\tau) : \tau \in [t_0, t]\}$. Henceforth, to look for an optimal control and/or decision policy $u^i(t)$ of the form (10), it is only required that

$$u^i(t) = \gamma^i(t, \hat{x}^i(t)), \quad t \in [t_0, t_f].$$

Given the linear-quadratic properties of the surrogate system (2)-(9), the search for an optimal feedback solution is productively restricted to a linear time-varying feedback policy generated from the locally accessible state $\hat{x}^i(t)$ by

$$u^i(t) = K^i(t)\hat{x}^i(t), \quad t \in [t_0, t_f] \quad (11)$$

with $K^i \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{m_i \times n})$ an admissible feedback form whose further defining properties will be stated shortly.

For the admissible pair (t_0, x_0^i) , the a-priori knowledge about neighboring disturbances $u^{-i}(\cdot)$ and the admissible feedback policy (11), the aggregation of the dynamics (5) and (8) associated with decision maker i is described by the controlled stochastic differential equation having $z^i(t_0) = z_0^i$

$$dz^i(t) = (F^i(t)z^i(t) + E^i(t)u^{-i}(t))dt + G^i(t)dw^i(t) \quad (12)$$

and the performance measure (9) is rewritten as follows

$$\begin{aligned} J^i(u^i, u^{-i}) &= (z^i)^T(t_f)N_f^i z^i(t_f) \\ &+ \int_{t_0}^{t_f} [(z^i)^T(\tau)N^i(\tau)z^i(\tau) - (u^{-i})^T(\tau)M^i(\tau)u^{-i}(\tau)]d\tau \end{aligned} \quad (13)$$

where the aggregate stationary Wiener process noise $w^i \triangleq [w^T \ (\eta^i)^T \ (v^i)^T]^T$ with the correlation of independent increments defined as $E\{[w^i(\tau_1) - w^i(\tau_2)][w^i(\tau_1) - w^i(\tau_2)]^T\} = W^i|\tau_1 - \tau_2|$ for all $\tau_1, \tau_2 \in [t_0, t_f]$ and the aggregate system coefficients

and state variables are given by

$$\begin{aligned}
F^i(t) &\triangleq \begin{bmatrix} A(t) + B^i(t)K^i(t) & L^i(t)C^i(t) \\ 0 & A(t) - L^i(t)C^i(t) \end{bmatrix} \\
G^i(t) &\triangleq \begin{bmatrix} 0 & 0 & L^i(t) \\ G(t) & -I_{n \times n} & -L^i(t) \end{bmatrix}, \quad z^i(t) \triangleq \begin{bmatrix} \hat{x}^i(t) \\ \hat{z}^i(t) \end{bmatrix} \\
E^i(t) &\triangleq \begin{bmatrix} I_{n \times n} \\ 0 \end{bmatrix}, \quad N_f^i \triangleq \begin{bmatrix} Q_f^i & Q_f^i \\ Q_f^i & Q_f^i \end{bmatrix} \\
W^i &\triangleq \begin{bmatrix} W & 0 & 0 \\ 0 & N^i & 0 \\ 0 & 0 & V^i \end{bmatrix}, \quad z_0^i \triangleq \begin{bmatrix} x_0 \\ 0 \end{bmatrix} \\
N^i(t) &\triangleq \begin{bmatrix} Q^i(t) + (K^i)^T(t)R^i(t)K^i(t) & Q^i(t) \\ Q^i(t) & Q^i(t) \end{bmatrix}.
\end{aligned}$$

In the sequel, moving from the background of the general chi-squared random performance (13), one productive step involved in extracting information from complex behavior of (13) is modeling and management of all the mathematical statistics (also known as semi-invariants) associated with it. To this end, the results that follow highlight the rather crucial role played by the endeavor of extracting higher-order statistics pertaining to random distributions of (13).

Theorem 1: Cumulant-Generating Function.

Let the states $z^i(\cdot)$ of the coupling constraint dynamics (12) subject to the performance measure (13) be associated with risk-averse decision maker i . Further, let initial states $z^i(\tau) \equiv z_\tau^i$ and $\tau \in [t_0, t_f]$ and moment-generating functions with risk sensitive parameter θ^i be denoted by

$$\begin{aligned}
\varphi^i(\tau, z_\tau^i, \theta^i) &= \\
\varrho^i(\tau, \theta^i) \exp\{(z_\tau^i)^T \Upsilon^i(\tau, \theta^i) z_\tau^i + 2(z_\tau^i)^T \ell^i(\tau, \theta^i)\} & \quad (14) \\
v^i(\tau, \theta^i) = \ln\{\varrho^i(\tau, \theta^i)\}, \quad \theta^i \in \mathbb{R}^+ & \quad (15)
\end{aligned}$$

Then, the cumulant-generating function is quadratic affine

$$\begin{aligned}
\psi^i(\tau, z_\tau^i, \theta^i) &= \\
= (z_\tau^i)^T \Upsilon^i(\tau, \theta^i) z_\tau^i + 2(z_\tau^i)^T \ell^i(\tau, \theta^i) + v^i(\tau, \theta^i) & \quad (16)
\end{aligned}$$

where the backward-in-time scalar valued $v^i(\tau, \theta^i)$ satisfies

$$\begin{aligned}
\frac{d}{d\tau} v^i(\tau, \theta^i) = -\text{Tr}\{\Upsilon^i(\tau, \theta^i) G^i(\tau) W^i (G^i)^T(\tau)\} \\
+ \theta^i (u^{-i})^T(\tau) M^i(\tau) u^{-i}(\tau), \quad v^i(t_f, \theta^i) = 0 & \quad (17)
\end{aligned}$$

whereas the backward-in-time matrix valued $\Upsilon^i(\tau, \theta^i)$ with $\Upsilon^i(t_f, \theta^i) = \theta^i N_f^i$ and vector valued $\ell^i(\tau, \theta^i)$ satisfy

$$\begin{aligned}
\frac{d}{d\tau} \Upsilon^i(\tau, \theta^i) = -(F^i)^T(\tau) \Upsilon^i(\tau, \theta^i) - \Upsilon^i(\tau, \theta^i) F^i(\tau) \\
- 2\Upsilon^i(\tau, \theta^i) G^i(\tau) W^i (G^i)^T(\tau) \Upsilon^i(\tau, \theta^i) - \theta^i N^i(\tau) & \quad (18)
\end{aligned}$$

$$\frac{d}{d\tau} \ell^i(\tau, \theta^i) = -\Upsilon^i(\tau, \theta^i) E^i(t) u^{-i}(\tau), \quad \ell^i(t_f, \theta^i) = 0. \quad (19)$$

Proof: The proof requires immediate results [6]. ■

Specifically, a MacLaurin series expansion of the cumulant-generating function (16) is employed to infer behaviors regarding probabilistic distributions of (13) through the knowledge representation of the mathematical statistics

$$\psi^i(\tau, z_\tau^i, \theta^i) = \sum_{r=1}^{\infty} \frac{\partial^{(r)}}{\partial (\theta^i)^{(r)}} \psi^i(\tau, z_\tau^i, \theta^i) \Big|_{\theta^i=0} \frac{(\theta^i)^r}{r!} \quad (20)$$

where all $\kappa_r^i \triangleq \frac{\partial^{(r)}}{\partial (\theta^i)^{(r)}} \psi^i(\tau, z_\tau^i, \theta^i) \Big|_{\theta^i=0}$ are performance-measure statistics available at risk-averse decision maker i

$$\begin{aligned}
\kappa_r^i = (z_\tau^i)^T \frac{\partial^{(r)}}{\partial (\theta^i)^{(r)}} \Upsilon^i(\tau, \theta^i) \Big|_{\theta^i=0} z_\tau^i & \quad (21) \\
+ 2(z_\tau^i)^T \frac{\partial^{(r)}}{\partial (\theta^i)^{(r)}} \ell^i(\tau, \theta^i) \Big|_{\theta^i=0} + \frac{\partial^{(r)}}{\partial (\theta^i)^{(r)}} v^i(\tau, \theta^i) \Big|_{\theta^i=0}.
\end{aligned}$$

For notational convenience, the change of variables

$$\begin{aligned}
H_r^i(\tau) &\triangleq \frac{\partial^{(r)} \Upsilon^i(\tau, \theta^i)}{\partial (\theta^i)^{(r)}} \Big|_{\theta^i=0}; \quad \check{D}_r^i(\tau) \triangleq \frac{\partial^{(r)} \ell^i(\tau, \theta^i)}{\partial (\theta^i)^{(r)}} \Big|_{\theta^i=0} \\
D_r^i(\tau) &\triangleq \frac{\partial^{(r)} v^i(\tau, \theta^i)}{\partial (\theta^i)^{(r)}} \Big|_{\theta^i=0}; \quad \tau \in [t_0, t_f]; \quad r \in \mathbb{N} & \quad (22)
\end{aligned}$$

is introduced so that the next result will provide an effective and accurate capability for forecasting all the higher-order characteristics associated with performance uncertainty (13).

Theorem 2: Performance-Measure Statistics.

Let (A, B^i) and (A, C^i) associated with the coupling constraint (12) and the goal function (13) be uniformly stabilizable and detectable. For $k^i \in \mathbb{N}$, the k^i th performance-measure statistic of (13) concerned by risk-averse decision maker i and $i \in \bar{N}$ is given by

$$\kappa_k^i = (z_0^i)^T H_{k^i}^i(t_0) z_0^i + 2(z_0^i)^T \check{D}_{k^i}^i(t_0) + D_{k^i}^i(t_0) \quad (23)$$

where the supporting variables $\{H_r^i(\tau)\}_{r=1}^{k^i}$, $\{\check{D}_r^i(\tau)\}_{r=1}^{k^i}$ and $\{D_r^i(\tau)\}_{r=1}^{k^i}$ evaluated at $\tau = t_0$ satisfy the differential equations (with the dependence of $H_r^i(\tau)$, $\check{D}_r^i(\tau)$ and $D_r^i(\tau)$ upon the admissible feedback policy gain $K^i(\tau)$ and other observable policies $u^{-i}(\tau)$ suppressed)

$$\frac{d}{d\tau} H_1^i(\tau) = -(F^i)^T(\tau) H_1^i(\tau) - H_1^i(\tau) F^i(\tau) - N^i(\tau) \quad (24)$$

$$\begin{aligned}
\frac{d}{d\tau} H_r^i(\tau) = -(F^i)^T(\tau) H_r^i(\tau) - H_r^i(\tau) F^i(\tau) \\
- \sum_{s=1}^{r-1} \frac{2r!}{s!(r-s)!} H_s^i(\tau) G^i(\tau) W^i (G^i)^T(\tau) H_{r-s}^i(\tau) & \quad (25)
\end{aligned}$$

and

$$\frac{d}{d\tau} \check{D}_r^i(\tau) = -H_r^i(\tau) E^i(\tau) u^{-i}(\tau), \quad 1 \leq r \leq k^i \quad (26)$$

and, finally,

$$\begin{aligned}
\frac{d}{d\tau} D_1^i(\tau) = -\text{Tr}\{H_1^i(\tau) G^i(\tau) W^i (G^i)^T(\tau)\} \\
+ (u^{-i})^T(\tau) M^i(\tau) u^{-i}(\tau) & \quad (27)
\end{aligned}$$

$$\frac{d}{d\tau} D_r^i(\tau) = -\text{Tr}\{H_r^i(\tau) G^i(\tau) W^i (G^i)^T(\tau)\}, \quad r \geq 2 \quad (28)$$

whereby the terminal-value conditions $H_1^i(t_f) = N_f^i$, $H_r^i(t_f) = 0$ for $2 \leq r \leq k^i$, $\check{D}_r^i(t_f) = 0$ for $1 \leq r \leq k^i$ and $D_r^i(t_f) = 0$ for $1 \leq r \leq k^i$.

Proof: Refer to [6] for the analytical proof. ■

III. PROBLEM STATEMENTS

In the context of risk-averse decision making, cautious decision makers who realize performance risk akin to a costly preference for certainty will have to leverage higher-order statistics of the probability distribution (13) for optimizing risk-averse decisions. For such a problem it is important to have a compact statement of the risk-averse decision and control optimization so as to aid mathematical manipulations. Precisely, one may think of the k^i -tuple state variables $\mathcal{H}^i(\cdot) \triangleq (\mathcal{H}_1^i(\cdot), \dots, \mathcal{H}_{k^i}^i(\cdot))$, $\check{\mathcal{D}}^i(\cdot) \triangleq (\check{\mathcal{D}}_1^i(\cdot), \dots, \check{\mathcal{D}}_{k^i}^i(\cdot))$ and $\mathcal{D}^i(\cdot) \triangleq (\mathcal{D}_1^i(\cdot), \dots, \mathcal{D}_{k^i}^i(\cdot))$ whose continuously differentiable states $\mathcal{H}_r^i \in \mathcal{C}^1([t_0, t_f]; \mathbb{R}^{2n \times 2n})$, $\check{\mathcal{D}}_r^i \in \mathcal{C}^1([t_0, t_f]; \mathbb{R}^{2n \times 1})$ and $\mathcal{D}_r^i \in \mathcal{C}^1([t_0, t_f]; \mathbb{R})$ having the representations $\mathcal{H}_r^i(\cdot) \triangleq H_r^i(\cdot)$, $\check{\mathcal{D}}_r^i(\cdot) \triangleq \check{D}_r^i(\cdot)$ and $\mathcal{D}_r^i(\cdot) \triangleq D_r^i(\cdot)$ with the right members satisfying the dynamics (24)-(28) are defined on $[t_0, t_f]$. In the remainder of the development, the convenient mappings are introduced as

$$\begin{aligned} \mathcal{F}_r^i &: [t_0, t_f] \times (\mathbb{R}^{2n \times 2n})^{k^i} \times \mathbb{R}^{m_i \times n} \mapsto \mathbb{R}^{2n \times 2n} \\ \check{\mathcal{G}}_r^i &: [t_0, t_f] \times (\mathbb{R}^{2n \times 1})^{k^i} \mapsto \mathbb{R}^{2n \times 1} \\ \mathcal{G}_r^i &: [t_0, t_f] \times (\mathbb{R}^{2n \times 2n})^{k^i} \mapsto \mathbb{R}, \end{aligned}$$

where the rules of action are given by

$$\begin{aligned} \mathcal{F}_1^i(\tau, \mathcal{H}^i, K^i) &\triangleq -(F^i)^T(\tau) \mathcal{H}_1^i(\tau) - \mathcal{H}_1^i(\tau) F^i(\tau) - N^i(\tau) \\ \mathcal{F}_r^i(\tau, \mathcal{H}^i, K^i) &\triangleq -(F^i)^T(\tau) \mathcal{H}_r^i(\tau) - \mathcal{H}_r^i(\tau) F^i(\tau) \\ &- \sum_{s=1}^{r-1} \frac{2r!}{s!(r-s)!} \mathcal{H}_s^i(\tau) G^i(\tau) W^i (G^i)^T(\tau) \mathcal{H}_{r-s}^i(\tau) \\ \check{\mathcal{G}}_r^i(\tau, \mathcal{H}^i) &\triangleq -\mathcal{H}_r^i(\tau) E^i(\tau) u^{-i}(\tau), \quad 1 \leq r \leq k^i \\ \mathcal{G}_1^i(\tau, \mathcal{H}^i) &\triangleq -\text{Tr} \{ \mathcal{H}_1^i(\tau) G^i(\tau) W^i (G^i)^T(\tau) \} \\ &+ (u^{-i})^T(\tau) M^i(\tau) u^{-i}(\tau) \\ \mathcal{G}_r^i(\tau, \mathcal{H}^i) &\triangleq -\text{Tr} \{ \mathcal{H}_r^i(\tau) G^i(\tau) W^i (G^i)^T(\tau) \}, \quad 2 \leq r \leq k^i. \end{aligned}$$

The product mappings that follow are necessary for a compact formulation; for example

$$\begin{aligned} \mathcal{F}_1^i \times \dots \times \mathcal{F}_{k^i}^i &: [t_0, t_f] \times (\mathbb{R}^{2n \times 2n})^{k^i} \times \mathbb{R}^{m_i \times n} \mapsto (\mathbb{R}^{2n \times 2n})^{k^i} \\ \check{\mathcal{G}}_1^i \times \dots \times \check{\mathcal{G}}_{k^i}^i &: [t_0, t_f] \times (\mathbb{R}^{2n \times 1})^{k^i} \mapsto (\mathbb{R}^{2n \times 1})^{k^i} \\ \mathcal{G}_1^i \times \dots \times \mathcal{G}_{k^i}^i &: [t_0, t_f] \times (\mathbb{R}^{2n \times 2n})^{k^i} \mapsto \mathbb{R}^{k^i} \end{aligned}$$

where the corresponding notations $\mathcal{F}^i \triangleq \mathcal{F}_1^i \times \dots \times \mathcal{F}_{k^i}^i$, $\check{\mathcal{G}}^i \triangleq \check{\mathcal{G}}_1^i \times \dots \times \check{\mathcal{G}}_{k^i}^i$ and $\mathcal{G}^i \triangleq \mathcal{G}_1^i \times \dots \times \mathcal{G}_{k^i}^i$ are used. Thus, the dynamical equations (24)-(28) can be rewritten as

$$\frac{d}{d\tau} \mathcal{H}^i(\tau) = \mathcal{F}^i(\tau, \mathcal{H}^i(\tau), K^i(\tau)), \quad \mathcal{H}^i(t_f) \equiv \mathcal{H}_f^i \quad (29)$$

$$\frac{d}{d\tau} \check{\mathcal{D}}^i(\tau) = \check{\mathcal{G}}^i(\tau, \mathcal{H}^i(\tau)), \quad \check{\mathcal{D}}^i(t_f) \equiv \check{\mathcal{D}}_f^i \quad (30)$$

$$\frac{d}{d\tau} \mathcal{D}^i(\tau) = \mathcal{G}^i(\tau, \mathcal{H}^i(\tau)), \quad \mathcal{D}^i(t_f) \equiv \mathcal{D}_f^i \quad (31)$$

where the k^i -tuple terminal-value conditions $\mathcal{H}_f^i \triangleq (N_f^i, 0, \dots, 0)$, $\check{\mathcal{D}}_f^i \triangleq (0, \dots, 0)$ and $\mathcal{D}_f^i \triangleq (0, \dots, 0)$.

Notice that the product system (29)-(31) uniquely determines the state matrices \mathcal{H}^i , $\check{\mathcal{D}}^i$ and \mathcal{D}^i once the admissible

feedback policy gain K^i and observable policies u^{-i} by decision maker i are specified. Henceforth, these state variables are considered as $\mathcal{H}^i \equiv \mathcal{H}^i(\cdot, K^i, u^{-i})$, $\check{\mathcal{D}}^i \equiv \check{\mathcal{D}}^i(\cdot, K^i, u^{-i})$ and $\mathcal{D}^i \equiv \mathcal{D}^i(\cdot, K^i, u^{-i})$. For the given terminal data $(t_f, \mathcal{H}_f^i, \check{\mathcal{D}}_f^i, \mathcal{D}_f^i)$, the class of admissible feedback gains employed by risk-averse decision maker i is next defined.

Definition 1: Admissible Feedback Policy Gains.

Let compact subset $\bar{K}^i \subset \mathbb{R}^{m_i \times n}$ be the set of allowable feedback form values. For the given $k^i \in \mathbb{N}$ and sequence $\mu^i = \{\mu_r^i \geq 0\}_{r=1}^{k^i}$ with $\mu_1^i > 0$, the set of feedback gains $\mathcal{K}_{t_f, \mathcal{H}_f^i, \check{\mathcal{D}}_f^i, \mathcal{D}_f^i; \mu^i}^i$ is assumed to be the class of $\mathcal{C}([t_0, t_f]; \mathbb{R}^{m_i \times n})$ with values $K^i(\cdot) \in \bar{K}^i$, for which the solutions to the dynamic equations (29)-(31) with the terminal-value conditions $\mathcal{H}^i(t_f) = \mathcal{H}_f^i$, $\check{\mathcal{D}}^i(t_f) = \check{\mathcal{D}}_f^i$ and $\mathcal{D}^i(t_f) = \mathcal{D}_f^i$ exist on the interval of optimization $[t_0, t_f]$.

An obvious fact about the private set of design freedom $\mu^i = \{\mu_r^i \geq 0\}_{r=1}^{k^i}$ with $\mu_1^i > 0$ is that risk sensitivity entails the lack of certainty equivalence, in the sense that any performance index formed only by the first statistic of (13) does not lead to optimal decisions. In addition, it is important to recognized that this finite set of custom weights is quite different from those of infinite sets of series expansion coefficients as in [7], [8], [9], just to name a few.

On $\mathcal{K}_{t_f, \mathcal{H}_f^i, \check{\mathcal{D}}_f^i, \mathcal{D}_f^i; \mu^i}^i$ the performance index with mean-risk considerations is subsequently defined as follows.

Definition 2: Mean-Risk Aware Performance Index.

Let cautious decision maker i select $k^i \in \mathbb{N}$ and the set of custom weights $\mu^i = \{\mu_r^i \geq 0\}_{r=1}^{k^i}$ with $\mu_1^i > 0$. Then for the given z_0^i , the mean-risk aware performance index

$$\phi_0^i : \{t_0\} \times (\mathbb{R}^{2n \times 2n})^{k^i} \times (\mathbb{R}^{2n})^{k^i} \times \mathbb{R}^{k^i} \mapsto \mathbb{R}^+$$

pertaining to risk-averse decision making over $[t_0, t_f]$ is

$$\begin{aligned} \phi_0^i(t_0, \mathcal{H}^i, \check{\mathcal{D}}^i, \mathcal{D}^i) &\triangleq \underbrace{\mu_1^i \kappa_1^i}_{\text{Mean}} + \underbrace{\mu_2^i \kappa_2^i + \dots + \mu_{k^i}^i \kappa_{k^i}^i}_{\text{Risk}} \\ &= \sum_{r=1}^{k^i} \mu_r^i [(z_0^i)^T \mathcal{H}_r^i(t_0) z_0^i + 2(z_0^i)^T \check{\mathcal{D}}_r^i(t_0) + \mathcal{D}_r^i(t_0)] \quad (32) \end{aligned}$$

where additional design freedom μ_r^i 's utilized by cautious and defensive decision maker i are tailored to meet different levels of performance-based reliability requirements, e.g., mean, variance, anti-symmetry, heavy tails of the reward/cost density (13), etc., pertaining to closed-loop performance uncertainties and whereas the supporting solutions $\{\mathcal{H}_r^i(\tau)\}_{r=1}^{k^i}$, $\{\check{\mathcal{D}}_r^i(\tau)\}_{r=1}^{k^i}$ and $\{\mathcal{D}_r^i(\tau)\}_{r=1}^{k^i}$ evaluated at $\tau = t_0$ satisfy the dynamical equations (29)-(31).

The technical challenge faced by cautious decision maker i and $i \in \bar{N}$ is that his mean-risk aware performance index (32) depends on the observable policies from neighboring decision makers u^{-i} . The basic question he faces is whether or not a sort of noncooperative equilibrium or Nash solution is possible at all. Defensive decision maker i 's rationale for choosing K_*^i is to force his neighbors to hold u_*^{-i} , so as to secure himself the Nash payoff. Thus, it is precisely in this sense that a Nash solution is risk-averse by nature.

Definition 3: Feedback Nash Equilibrium.

Let K_*^i constitute a feedback Nash strategy such that

$$\phi_0^i(K_*^i, u_*^{-i}) \leq \phi_0^i(K^i, u_*^{-i}), \quad i \in \bar{N} \quad (33)$$

for all admissible $K^i \in \mathcal{K}_{t_f, \mathcal{H}_f^i, \check{\mathcal{D}}_f^i, \mathcal{D}_f^i; \mu^i}^i$, upon which the solutions to the dynamical systems (29)-(31) exist on $[t_0, t_f]$.

Then, (K_*^1, \dots, K_*^N) when restricted to $[t_0, \tau]$ is still a N -tuple feedback Nash equilibrium solution for the multiperson Nash decision problem with the appropriate terminal-value condition $(\tau, \mathcal{H}_*^i(\tau), \check{\mathcal{D}}_*^i(\tau), \mathcal{D}_*^i(\tau))$ for all $\tau \in [t_0, t_f]$.

Now, the objective of cautious decision maker i is to minimize (32) over $K^i = K^i(\cdot)$ in $\mathcal{K}_{t_f, \mathcal{H}_f^i, \check{\mathcal{D}}_f^i, \mathcal{D}_f^i; \mu^i}^i$ and subject to the neighboring feedback Nash policies u_*^{-i} .

Definition 4: Optimization of Mayer Problem.

Given the profile of risk-averse attitudes $\mu^i = \{\mu_r^i \geq 0\}_{r=1}^{k^i}$ with $\mu_1^i > 0$, the decision optimization problem defined by

$$K^i(\cdot) \in \mathcal{K}_{t_f, \mathcal{H}_f^i, \check{\mathcal{D}}_f^i, \mathcal{D}_f^i; \mu^i}^i \quad \min \quad \phi_0^i(K^i, u_*^{-i}) \quad (34)$$

is subject to the dynamical equations (29)-(31) on $[t_0, t_f]$.

In conformity with the dynamic programming approach, the terminal time and states $(t_f, \mathcal{H}_f^i, \check{\mathcal{D}}_f^i, \mathcal{D}_f^i)$ are parameterized as $(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i)$ whereby $\mathcal{Y}^i \triangleq \mathcal{H}^i(\varepsilon)$, $\check{\mathcal{Z}}^i \triangleq \check{\mathcal{D}}^i(\varepsilon)$ and $\mathcal{Z}^i \triangleq \mathcal{D}^i(\varepsilon)$. Thus, the value function of (34) now depends on the parameterization of terminal-value conditions.

Definition 5: Value Function.

Let $(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i) \in [t_0, t_f] \times (\mathbb{R}^{2n \times 2n})^{k^i} \times (\mathbb{R}^{2n \times 1})^{k^i} \times \mathbb{R}^{k^i}$. Then, $\mathcal{V}^i(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i)$ is defined by $\mathcal{V}^i(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i) \triangleq \inf$

$$K^i(\cdot) \in \mathcal{K}_{\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i; \mu^i}^i \quad \phi_0^i(K^i, u_*^{-i}).$$

For convention, $\mathcal{V}^i(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i) \triangleq \infty$ when $\mathcal{K}_{\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i; \mu^i}^i$ is empty. Some candidates for the value function are also constructed with the help of the concept of reachable set.

Definition 6: Reachable Sets.

Let a reachable set of decision maker i be defined by $\mathcal{Q}^i \triangleq \{(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i) \in [t_0, t_f] \times (\mathbb{R}^{2n \times 2n})^{k^i} \times (\mathbb{R}^{2n \times 1})^{k^i} \times \mathbb{R}^{k^i} : \mathcal{K}_{\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i; \mu^i}^i \neq \emptyset\}$.

Formally, it can be shown that the value function associated with decision maker i is satisfying partial differential equation (e.g., Hamilton-Jacobi-Bellman (HJB) equation) at interior points of \mathcal{Q}^i , at which it is differentiable.

Theorem 3: HJB Equation-Mayer Problem.

Let $(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i)$ be any interior point of \mathcal{Q}^i , at which the value function $\mathcal{V}^i(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i)$ is differentiable. If there exists a feedback Nash strategy $K_*^i \in \mathcal{K}_{t_f, \mathcal{H}_f^i, \check{\mathcal{D}}_f^i, \mathcal{D}_f^i; \mu^i}^i$, then the partial differential equation is satisfied

$$\begin{aligned} 0 = & \min_{K^i \in \bar{K}^i} \left\{ \frac{\partial}{\partial \varepsilon} \mathcal{V}^i(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i) \right. \\ & + \frac{\partial}{\partial \text{vec}(\mathcal{Y}^i)} \mathcal{V}^i(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i) \text{vec}(\mathcal{F}^i(\varepsilon, \mathcal{Y}^i, K^i)) \\ & + \frac{\partial}{\partial \text{vec}(\check{\mathcal{Z}}^i)} \mathcal{V}^i(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i) \text{vec}(\check{\mathcal{G}}^i(\varepsilon, \mathcal{Y}^i)) \\ & \left. + \frac{\partial}{\partial \text{vec}(\mathcal{Z}^i)} \mathcal{V}^i(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i) \text{vec}(\mathcal{G}^i(\varepsilon, \mathcal{Y}^i)) \right\} \quad (35) \end{aligned}$$

and $\mathcal{V}^i(t_0, \mathcal{Y}^i(t_0), \check{\mathcal{Z}}^i(t_0), \mathcal{Z}^i(t_0)) = \phi_0^i(\mathcal{H}^i(t_0), \check{\mathcal{D}}^i(t_0), \mathcal{D}^i(t_0))$.

Proof: Similar to that of [10] and hence is omitted. ■

Finally, the sufficient condition used to verify a feedback Nash strategy for cautious and defensive decision maker i and $i \in \bar{N}$ is given next.

Theorem 4: Verification Theorem.

Let $\mathcal{W}^i(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i)$ be continuously differentiable solution of (35) with the boundary $\mathcal{W}^i(t_0, \mathcal{H}^i(t_0), \check{\mathcal{D}}^i(t_0), \mathcal{D}^i(t_0)) = \phi_0^i(t_0, \mathcal{H}^i(t_0), \check{\mathcal{D}}^i(t_0), \mathcal{D}^i(t_0))$. Let $(t_f, \mathcal{H}_f^i, \check{\mathcal{D}}_f^i, \mathcal{D}_f^i) \in \mathcal{Q}^i$, $K^i \in \mathcal{K}_{t_f, \mathcal{H}_f^i, \check{\mathcal{D}}_f^i, \mathcal{D}_f^i; \mu^i}^i$, $(\mathcal{H}^i(\cdot), \check{\mathcal{D}}^i(\cdot), \mathcal{D}^i(\cdot))$ be the trajectory solutions of (29)-(31). Then, $\mathcal{W}^i(\tau, \mathcal{H}^i(\tau), \check{\mathcal{D}}^i(\tau), \mathcal{D}^i(\tau))$ is time-backward increasing function of $\tau \in [t_0, t_f]$.

If K_*^i is in $\mathcal{K}_{t_f, \mathcal{H}_f^i, \check{\mathcal{D}}_f^i, \mathcal{D}_f^i; \mu^i}^i$ with the corresponding trajectory solutions $(\mathcal{H}_*^i(\cdot), \check{\mathcal{D}}_*^i(\cdot), \mathcal{D}_*^i(\cdot))$ of the equations (29)-(31) such that

$$0 = \frac{\partial}{\partial \varepsilon} \mathcal{W}^i(\tau, \mathcal{H}_*^i(\tau), \check{\mathcal{D}}_*^i(\tau), \mathcal{D}_*^i(\tau)) \quad (36)$$

$$\begin{aligned} & + \frac{\partial}{\partial \text{vec}(\mathcal{Y}^i)} \mathcal{W}^i(\tau, \mathcal{H}_*^i(\tau), \check{\mathcal{D}}_*^i(\tau), \mathcal{D}_*^i(\tau)) \\ & \text{vec}(\mathcal{F}^i(\tau, \mathcal{H}_*^i(\tau), K_*^i(\tau))) \\ & + \frac{\partial}{\partial \text{vec}(\check{\mathcal{Z}}^i)} \mathcal{W}^i(\tau, \mathcal{H}_*^i(\tau), \check{\mathcal{D}}_*^i(\tau), \mathcal{D}_*^i(\tau)) \text{vec}(\check{\mathcal{G}}^i(\tau, \mathcal{H}_*^i(\tau))) \\ & + \frac{\partial}{\partial \text{vec}(\mathcal{Z}^i)} \mathcal{W}^i(\tau, \mathcal{H}_*^i(\tau), \check{\mathcal{D}}_*^i(\tau), \mathcal{D}_*^i(\tau)) \text{vec}(\mathcal{G}^i(\tau, \mathcal{H}_*^i(\tau))) \end{aligned}$$

then, K_*^i is a feedback Nash strategy in $\mathcal{K}_{t_f, \mathcal{H}_f^i, \check{\mathcal{D}}_f^i, \mathcal{D}_f^i; \mu^i}^i$,

$$\mathcal{W}^i(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i) = \mathcal{V}^i(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i). \quad (37)$$

Proof: The proof follows the same manner as [10]. ■

IV. DISTRIBUTED RISK-AVERSE DECISION STRATEGIES

To this end, the initial state z_0^i is recognized to contribute linearly and quadratically to the mean-risk performance index (32). Henceforth, it is beneficial to infer that a candidate function for the value function is expected to take the form

$$\begin{aligned} \mathcal{W}^i(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i) = & (z_0^i)^T \sum_{r=1}^{k^i} \mu_r^i (\mathcal{Y}_r^i + \mathcal{E}_r^i(\varepsilon)) z_0^i \\ & + 2(z_0^i)^T \sum_{r=1}^{k^i} \mu_r^i (\check{\mathcal{Z}}_r^i + \check{\mathcal{T}}_r^i(\varepsilon)) + \sum_{r=1}^{k^i} \mu_r^i (\mathcal{Z}_r^i + \mathcal{T}_r^i(\varepsilon)) \quad (38) \end{aligned}$$

where the parametric functions of time $\mathcal{E}_r^i \in \mathcal{C}^1([t_0, t_f]; \mathbb{R}^{2n \times 2n})$, $\check{\mathcal{T}}_r^i \in \mathcal{C}^1([t_0, t_f]; \mathbb{R}^{2n \times 1})$, and $\mathcal{T}_r^i \in \mathcal{C}^1([t_0, t_f]; \mathbb{R})$ are yet to be determined.

In addition, the time derivative of $\mathcal{W}^i(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i)$ can be shown as follows

$$\begin{aligned} \frac{d}{d\varepsilon} \mathcal{W}^i(\varepsilon, \mathcal{Y}^i, \check{\mathcal{Z}}^i, \mathcal{Z}^i) = & \sum_{r=1}^{k^i} \mu_r^i [\mathcal{G}_r^i(\varepsilon, \mathcal{Y}^i) + \frac{d}{d\varepsilon} \mathcal{T}_r^i(\varepsilon)] \\ & + (z_0^i)^T \sum_{r=1}^{k^i} \mu_r^i [\mathcal{F}_r^i(\varepsilon, \mathcal{Y}^i, K^i) + \frac{d}{d\varepsilon} \mathcal{E}_r^i(\varepsilon)] z_0^i \\ & + 2(z_0^i)^T \sum_{r=1}^{k^i} \mu_r^i [\check{\mathcal{G}}_r^i(\varepsilon, \mathcal{Y}^i) + \frac{d}{d\varepsilon} \check{\mathcal{T}}_r^i(\varepsilon)]. \quad (39) \end{aligned}$$

The substitution of this candidate (38) for the value function into the HJB equation (35) and making use of (39) yield

$$0 = \min_{K^i \in \bar{K}^i} \left\{ (z_0^i)^T \sum_{r=1}^{k^i} \mu_r^i [\mathcal{F}_r^i(\varepsilon, \mathcal{Y}^i, K^i) + \frac{d}{d\varepsilon} \mathcal{E}_r^i(\varepsilon)] z_0^i \right. \\ \left. + 2(z_0^i)^T \sum_{r=1}^{k^i} \mu_r^i [\check{\mathcal{G}}_r^i(\varepsilon, \mathcal{Y}^i) + \frac{d}{d\varepsilon} \check{\mathcal{T}}_r^i(\varepsilon)] \right. \\ \left. + \sum_{r=1}^{k^i} \mu_r^i [\mathcal{G}_r^i(\varepsilon, \mathcal{Y}^i) + \frac{d}{d\varepsilon} \mathcal{T}_r^i(\varepsilon)] \right\}. \quad (40)$$

Taking the gradient with respect to K^i of the expression within the bracket of (40) yields the necessary conditions for an extremum of (32) on $[t_0, \varepsilon]$ where $I_0^T \triangleq [I_{n \times n} \quad 0]$

$$K^i = -(R^i)^{-1}(\varepsilon)(B^i)^T(\varepsilon)I_0^T \sum_{r=1}^{k^i} \hat{\mu}_r^i \mathcal{Y}_r^i I_0 ((I_0^T I_0)^{-1})^T \quad (41)$$

whereas $\hat{\mu}_r^i \triangleq \mu_r^i / \mu_1^i$ for $\mu_1^i > 0$. With the feedback Nash strategy (41) replaced in the expression of the bracket (40) and having $\{\mathcal{Y}_r^i\}_{r=1}^{k^i}$ evaluated on the optimal solution trajectories (29)-(31), $\mathcal{E}_r^i(\varepsilon)$, $\check{\mathcal{T}}_r^i(\varepsilon)$ and $\mathcal{T}_r^i(\varepsilon)$ are thus chosen so that the sufficient condition (36) in the verification theorem is satisfied despite of the arbitrary values z_0^i ; for example

$$\begin{aligned} \dot{\mathcal{E}}_1^i(\varepsilon) &= (F_*^i)^T(\varepsilon)\mathcal{H}_{1*}^i(\varepsilon) + \mathcal{H}_{1*}^i(\varepsilon)F_*^i(\varepsilon) + N_*^i(\varepsilon) \\ \dot{\mathcal{E}}_r^i(\varepsilon) &= (F_*^i)^T(\varepsilon)\mathcal{H}_{r*}^i(\varepsilon) + \mathcal{H}_{r*}^i(\varepsilon)F_*^i(\varepsilon) \\ &+ \sum_{s=1}^{r-1} \frac{2r!}{s!(r-s)!} \mathcal{H}_{s*}^i(\varepsilon)G^i(\varepsilon)W^i(G^i)^T(\varepsilon)\mathcal{H}_{r-s*}^i(\varepsilon), \quad r \geq 2 \\ \dot{\check{\mathcal{T}}}_r^i(\varepsilon) &= \mathcal{H}_{r*}^i(\varepsilon)E^i(\varepsilon)u_*^{-i}(\varepsilon), \quad \check{\mathcal{T}}_r^i(t_0) = 0, \quad 1 \leq r \leq k^i \\ \dot{\mathcal{T}}_1^i(\varepsilon) &= \text{Tr} \{ \mathcal{H}_{1*}^i(\varepsilon)G^i(\varepsilon)W^iG^{iT}(\varepsilon) \} - (u_*^{-i})^T(\varepsilon)M^i(\varepsilon)u_*^{-i}(\varepsilon) \\ \dot{\mathcal{T}}_r^i(\varepsilon) &= \text{Tr} \{ \mathcal{H}_{r*}^i(\varepsilon)G^i(\varepsilon)W^i(G^i)^T(\varepsilon) \}, \quad 2 \leq r \leq k^i \end{aligned}$$

where initial conditions $\mathcal{E}_r^i(t_0) = \mathcal{T}_r^i(t_0) = 0$ for $1 \leq r \leq k$.

Before closing the section, it is important to note that the sufficient condition (36) of the verification theorem is satisfied. Hence, the extremizing feedback strategy (41) associated with cautious decision maker i becomes optimal.

Theorem 5: Distributed Risk-Averse Decision Strategies. Consider noncooperative stochastic games and/or distributed controls wherein self-directed decision makers/autonomous controllers i and $i \in \bar{N}$ have complete knowledge of the coupling constraints (2)-(9) and mean-risk aware performance index (32). When all decision makers i have the similar risk attitudes, an imitative or Nash equilibrium exists and is enabled by the risk-averse class of feedback strategies

$$u_*^i(t) = K_*^i(t)\hat{x}_*^i(t), \quad t \triangleq t_0 + t_f - \tau \quad (42)$$

$$K_*^i(\tau) = -(R^i)^{-1}(\tau)(B^i)^T(\tau)I_0^T \sum_{r=1}^{k^i} \hat{\mu}_r^i \mathcal{H}_{r*}^i(\tau)I_0 ((I_0^T I_0)^{-1})^T$$

where all the parametric design freedom through $\hat{\mu}_r^i$ represent the risk-averse preferences toward performance distributions; the optimal trajectory solutions $\mathcal{H}_{r*}^i(\cdot)$ are satisfying (29).

Notice that, to have the distributed feedback Nash policy (42) of decision maker i be defined and continuous for all $\tau \in [t_0, t_f]$, the solutions $\mathcal{H}_{r*}^i(\tau)$ to the equations (29) when evaluated at $\tau = t_0$ must also exist. Therefore, it is necessary that $\mathcal{H}_{r*}^i(\tau)$ are finite for all $\tau \in [t_0, t_f]$. Moreover, the solutions of (29) exist and are continuously differentiable in a neighborhood of t_f . Under the assumption of (A, B^i) and (C^i, A) uniformly stabilizable and detectable, the result from [11] concludes that these solutions can further be extended to the left of t_f as long as $\mathcal{H}_{r*}^i(\tau)$ remain finite. Hence, the existence of unique and continuously differentiable solutions to the equations (29) is certain if $\mathcal{H}_{r*}^i(\tau)$ are bounded for all $\tau \in [t_0, t_f]$. As the result, the candidate value functions $\mathcal{W}^i(\tau, \mathcal{H}^i, \mathcal{D}^i, \mathcal{D}^i)$ are continuously differentiable as well.

V. CONCLUSIONS

Recent advances have generalized the results known for linear-quadratic classes of noncooperative stochastic games and distributed controls. Under risk attitudes toward performance uncertainties, the risk-averse feedback decision laws are not only the functions of higher-order statistics of the chi-squared rewards or costs but also dependent of a priori knowledge of common process noises as well as subjective observation noises. Thus, both certainty equivalence and separation principle do not hold. The existence of the Nash equilibrium as proposed herein is conditional upon the custom sets of selective weights, which in turn relate to risk parameters residing at cautious decision makers or controllers. An extension of the results obtained in this exposition may be worthy of investigation, when there are presence of mistrust and excessive risk aversion; such results could constitute fundamentals and principles in adversarial systems sciences and flexibly survivable decision making.

REFERENCES

- [1] T. Basar and G. J. Olsder. *Dynamic Noncooperative Game Theory*. 2nd Edition, Society for Industrial and Applied Mathematics, 1999
- [2] J. C. Engwerda. *LQ Dynamic Optimization and Differential Games*. Wiley 2005
- [3] A. Pollatsek and A. Tversky, "Theory of Risk," *Journal of Mathematical Psychology*, Vol. 7, No. 3, pp. 540-53, 1970
- [4] R. D. Luce, "Several Possible Measures of Risk," *Theory and Decision*, Vol. 12, pp. 217-228, 1980
- [5] K. D. Pham, "Cooperative Outcomes for Stochastic Nash Games: Decision Strategies towards Multi-Attribute Performance Robustness," *The 17th International Federation of Automatic Control World Congress*, pp. 11750-11756, Seoul, Korea, 2008
- [6] K. D. Pham, *Linear-Quadratic Controls in Risk-Averse Decision Making: Performance-Measure Statistics and Control Decision Optimization*, Springer Briefs in Optimization, ISBN 978-1-4614-5078-8, 2012
- [7] D. H. Jacobson, "Optimal Stochastic Linear Systems with Exponential Performance Criteria and Their Relation to Deterministic Games," *IEEE Transactions on Automatic Control*, Vol. 18, pp. 124-131, 1973
- [8] P. Whittle. *Risk Sensitive Optimal Control*. John Wiley & Sons, 1990
- [9] T. Basar, "Nash Equilibria of Risk-Sensitive Nonlinear Stochastic Differential Games," *Journal of Optimization Theory and Applications*, Vol. 100, No. 3, pp. 479-498, 1999
- [10] K. D. Pham, "Performance-Reliability-Aided Decision-Making in Multiperson Quadratic Decision Games against Jamming and Estimation Confrontations," *Journal of Optimization Theory and Applications*, Edited by F. Giannessi, Vol. 149, No. 1, pp. 599-629, 2011
- [11] J. Dieudonne. *Foundations of Modern Analysis*. Academic Press, New York and London, 1960