

Multi-Cumulant and Pareto Strategies for Stochastic Multi-Player Pursuit-Evasion

Khanh D. Pham

Space Vehicles Directorate
 Air Force Research Laboratory
 Kirtland AFB, NM 87117

Seth Lacy

Space Vehicles Directorate
 Air Force Research Laboratory
 Kirtland AFB, NM 87117

Lawrence Robertson

Space Vehicles Directorate
 Air Force Research Laboratory
 Kirtland AFB, NM 87117

Abstract—The paper presents an extension of cumulant-based control theory over a finite horizon for a class of multi-player pursuit-evasion wherein the evolution of the states of the game in response to adversarial strategies selected by pursuit and evasion teams from the efficient Pareto sets of admissible strategies is described by a stochastic linear differential equation and an integral-quadratic performance-measure. Both cooperation within each team and competition between the teams presumably exist. A direct dynamic programming approach for the Mayer optimization problem is used to solve for a multi-cumulant and Pareto-based solution when the members in each team optimally implement collective strategies and effectively shape the distribution of their Chi-squared random measures of performance associated with this special class of stochastic multi-player pursuit-evasion games.

I. INTRODUCTION

Since the 1950s, the work of Issacs [3] in deterministic pursuit-evasion game of a single pursuer and a single evader with perfect information and common knowledge has been greatly extended to pursuit-evasion with multiple pursuers and multiple evaders. Recent developments [4], [7] and important applications therein respectively treat probabilistic discrete-time and continuous-time problems with averaged performance indices. To the best knowledge of the authors, there hasn't yet been any work done for stochastic multi-player pursuit-evasion wherein the members in each team implement collective strategies to optimize their realized performance distribution beyond the traditional long-run average performance while in the competition with the other rival team and Nature's mixed random strategies. To address this completely unexplored research area, the present work presumes a cognitive model of Nature actions against which pursuit and evasion teams evaluate how they are doing. In effect pursuit and evasion teams apply criteria of performance to their coalitive decision strategies in response to almost all realizations generated by Nature. Such criteria are concerned with the actions imposed by Nature. Both rival teams try to anticipate what Nature is going to do in the future and they choose their actions to optimize performances given their anticipations. In other words, pursuit and evasion teams are conscious that their respective criteria of performance are dependent on Nature's mixed random realizations in the course of pursuit and evasion. Therefore, the minimization of the average measures of performance guarantees a priori nothing about the actual performance distributions because

This work was sponsored in part by the Air Force Office of Scientific Research under LRI 00VSI7COR. Correspondences should be directed to the email: AFRL.RVSV@kirtland.af.mil

Nature ultimately decides all the realizations of environmental interferences for the underlying stochastic process.

This paper is organized to provide innovative answers to the emerging unresolved questions: 1) What if pursuit and evasion teams are not reasonably content with such a limited average measure of performance as often seen in all existing state-of-the-art literatures? 2) Is there an efficient and tractable method that calculates exactly higher-order characteristics of the performance distribution? and 3) What new team strategies may be defined to directly influence the shared and conflicting performance distribution and to guarantee its robustness?

II. PROBLEM FORMULATION

For analytical tractability, a class of differential games against Nature whose interaction dynamics of pursuers and evaders are linear and performance measures are quadratic functions of the states and controls is considered. Specifically, a pursuit-evasion game with team P with m_P pursuers, identified as m_1, \dots, m_P , and team E with m_E evaders, identified as m_1, \dots, m_E , in an open subset of Hilbert space. Denote by $x_i^X(t) \triangleq x_i^X(t, \omega_i^X) : [t_0, t_f] \times \Omega_i^X \mapsto \mathbb{R}^{n_i^X}$ belonging to the Hilbert space $L^2_{\mathcal{F}_{it}^X}(\Omega_i^X; \mathcal{C}([t_0, t_f]; \mathbb{R}^{n_i^X}))$ of $\mathbb{R}^{n_i^X}$ -valued, square integrable processes on $[t_0, t_f]$ that are adapted to the σ -field \mathcal{F}_{it}^X generated by $w_i^X(t)$ with $E \left\{ \int_{t_0}^{t_f} (x_i^X)^T(\tau) x_i^X(\tau) d\tau \right\} < \infty$ the state variables for the members $i = 1, \dots, m_X$ in each non-cooperative team $X = P, E$ whose interactive relations are described by

$$dx_i^X(t) = (A_i^X(t)x_i^X(t) + B_{ui}^X(t)u_i^X(t))dt + G_i^X(t)dw_i^X(t), \quad x_i^X(t_0) = x_{i0}^X \quad (1)$$

where the initial states x_{i0}^X are known. Nature with action space of $w_i^X(t) \triangleq w_i^X(t, \omega_i^X) : [t_0, t_f] \times \Omega_i^X \mapsto \mathbb{R}^{p_i^X}$ are the p_i^X -dimensional stationary Wiener process defined with $\{\mathcal{F}_{it}^X\}_{t \geq 0}$ being its natural filtration on complete filtered probability spaces $(\Omega_i^X, \mathcal{F}_{it}^X, \{\mathcal{F}_{it}^X\}_{t \geq 0}, \mathcal{P}_i^X)$ over $[t_0, t_f]$ with the correlations of independent increments

$$E \{ [w_i^X(\tau) - w_i^X(\xi)][w_i^X(\tau) - w_i^X(\xi)]^T \} = W_i^X |\tau - \xi|,$$

and continuous-time coefficients $A_i^X \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{n_i^X \times n_i^X})$, $B_{ui}^X \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{n_i^X \times m_i^X})$, and $G_i^X \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{n_i^X \times p_i^X})$. In (1), $u_i^X \in \mathcal{U}_i^X$ are the admissible controls for the members in each team where $\mathcal{U}_i^X \in L^2_{\mathcal{F}_{it}^X}(\Omega_i^X; \mathcal{C}([t_0, t_f]; \mathbb{R}^{m_i^X}))$ are the subsets of Hilbert space of $\mathbb{R}^{m_i^X}$ -valued, square integrable processes on $[t_0, t_f]$ that are adapted to the σ -field

\mathcal{F}_{it}^X generated by $w_i^X(t)$. For simplicity of notation, let $x^X \triangleq [(x_1^X)^T, \dots, (x_{m_X}^X)^T]^T$, $u^X \triangleq [(u_1^X)^T, \dots, (u_{m_X}^X)^T]^T$, $A^X \triangleq \text{diag}(A_1^X, \dots, A_{m_X}^X)$, $B_u^X \triangleq \text{diag}(B_{u_1}^X, \dots, B_{u_{m_X}}^X)$, and $G^X \triangleq \text{diag}(G_1^X, \dots, G_{m_X}^X)$. Then, the homogeneous interactions within each rival team can be rewritten as

$$dx^X(t) = (A^X(t)x^X(t) + B_u^X(t)u^X(t))dt + G^X(t)dw^X(t), \quad x^X(t_0) = x_0^X \quad (2)$$

and the heterogenous interactions for the multi-player stochastic pursuit-evasion modeled on $[t_0, t_f]$ is given by

$$dx(t) = (A(t)x(t) + B^P(t)u^P(t) + B^E(t)u^E(t))dt + G(t)dw(t), \quad x(t_0) = x_0 \quad (3)$$

where the aggregate coefficients $A \triangleq \text{diag}(A^P, A^E)$, $B^P \triangleq [(B_u^P)^T, 0]^T$, $B^E \triangleq [0, (B_u^E)^T]^T$, $G \triangleq \text{diag}(G^P, G^E)$, $x \triangleq [(x^P)^T, (x^E)^T]^T$, $dw \triangleq [(dw^P)^T, (dw^E)^T]^T$, and $W \triangleq \text{diag}(W_1^P, \dots, W_{m_P}^P, W_1^E, \dots, W_{m_E}^E)$. Suppose further that $\mathcal{U}^P \triangleq \prod_{i=1}^{m_P} \mathcal{U}_i^P$, $\mathcal{U}^E \triangleq \prod_{j=1}^{m_E} \mathcal{U}_j^E$, $\mathcal{X} \triangleq \prod_{i=1}^{m_P} \mathbb{R}^{n_i^P} \times \prod_{j=1}^{m_E} \mathbb{R}^{n_j^E}$. Then, associated with admissible 3-tuple $(x(\cdot), u^P(\cdot), u^E(\cdot))$ and a sample path realization from Nature is a traditional performance-measure $J_i^X : \mathcal{X} \times \mathcal{U}^P \times \mathcal{U}^E \mapsto \mathbb{R}^+$ for which member i in team X attempts to optimize in presence of the other rival team's confrontations

$$J_i^X(x_0; u^P(\cdot), u^E(\cdot)) = x^T(t_f)Q_{if}^X x(t_f) + \int_{t_0}^{t_f} \left[x^T(\tau)Q_i^X(\tau)x(\tau) + \sum_{j=1}^{m_P} (u_j^P)^T(\tau)R_{ij}^{XP}(\tau)u_j^P(\tau) + \sum_{j=1}^{m_E} (u_j^E)^T(\tau)R_{ij}^{XE}(\tau)u_j^E(\tau) \right] d\tau \quad (4)$$

subject to the dynamics of the differential game (3) where $Q_{if}^X \in \mathbb{R}^{(\sum_{i=1}^{m_P} n_i^P + \sum_{j=1}^{m_E} n_j^E) \times (\sum_{i=1}^{m_P} n_i^P + \sum_{j=1}^{m_E} n_j^E)}$, $Q_i^X \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{(\sum_{i=1}^{m_P} n_i^P + \sum_{j=1}^{m_E} n_j^E) \times (\sum_{i=1}^{m_P} n_i^P + \sum_{j=1}^{m_E} n_j^E)})$, cross-coupling control inputs $R_{ij}^{XP} \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{m_j^P \times m_j^P})$ and $R_{ij}^{XE} \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{m_j^E \times m_j^E})$ are symmetric and positive semidefinite with $R_{ij}^{XP}(t)$ and $R_{ij}^{XE}(t)$ invertible.

Within team X , all the members decide to act cooperatively via a negotiating solution that comes from a coalitive Pareto decision subset \mathcal{U}_C^X of the class of permissible decisions \mathcal{U}^X . This solution is particularly attractive with the property that if any other solution is used at least one of the team members is penalized in the sense that his realized performance is worse, or all the members do the same.

Definition 1: Efficient Pareto Control Decisions.

For any random realization $\omega^X \in \Omega^X$ drawn by Nature, the coalitive strategies \hat{u}_C^P and \hat{u}_C^E are Pareto efficient if

$$(\hat{u}_C^P, \hat{u}_C^E) \in \arg \min_{u_C^P \in \mathcal{U}_C^P, u_C^E \in \mathcal{U}_C^E} \{J_i^X(x_0; u_C^P, u_C^E)\}_{i=1}^{m_X}. \quad (5)$$

The point $(J_1^X(x_0; u_C^P, u_C^E), \dots, J_{m_X}^X(x_0; u_C^P, u_C^E)) \in \mathbb{R}^{m_X}$ is called a Pareto solution. The set of all Pareto solutions is called the Pareto frontier.

In the subsequent analysis, the set of parameters, W^X represents for team cooperative profiles $\xi^X \in W^X$

$$W^X \triangleq \left\{ \xi^X \in \mathbb{R}^{m_X} : \sum_{i=1}^{m_X} \xi_i^X = 1; 0 < \xi_i^X \leq 1 \right\}. \quad (6)$$

Proposition 1: Efficient Pareto Parameterizations.

Let $\xi^X \in W^X$ and $\omega^X \in \Omega^X$. If

$$(\hat{u}_C^P, \hat{u}_C^E) \in \arg \min_{u_C^P \in \mathcal{U}_C^P, u_C^E \in \mathcal{U}_C^E} \left\{ \sum_{i=1}^{m_X} \xi_i^X J_i^X(x_0; u_C^P, u_C^E) \right\} \quad (7)$$

then $(\hat{u}_C^P, \hat{u}_C^E)$ is Pareto efficient.

Since the performance-measure (4) is a convex function on the convex set $\mathcal{U}_C^P \times \mathcal{U}_C^E$ with the convex constraint (3), the problem of solving for a set of efficient Pareto decisions within each team X with a vector-valued performance-measure criterion is equivalent to the problem of solving an $m_X - 1$ parameter family of optimal control problems with scalar cost criteria [5].

Proposition 2: Necessary and Sufficient Conditions.

Suppose $\mathcal{U}_C^P \times \mathcal{U}_C^E$ is convex and realized performance measure J_i^X is convex and concave with respect to u_C^P and u_C^E . The 2-tuple $(\hat{u}_C^P, \hat{u}_C^E)$ is efficient if and only if there exists $\xi^X \in W^X$ such that \hat{u}_C^P and \hat{u}_C^E are the corresponding Pareto-efficient strategies obtained as the argument of

$$\min_{u_C^P \in \mathcal{U}_C^P, u_C^E \in \mathcal{U}_C^E} \left\{ J_{\xi^X}^X(x_0; u_C^P, u_C^E) \triangleq \sum_{i=1}^{m_X} \xi_i^X J_i^X(x_0; u_C^P, u_C^E) \right\}$$

However, there are, in general, many Pareto solutions which depend on parameters ξ^X and Nature's actions ω^X . This observation raises the question as to which one is the best. To address this question, the present work therefore suggests an emerging arena of what is called performance robustness against Nature which is now explored by the cumulant-based control theory, in conjunction with the Pareto efficiency.

Let $Q_f^X \triangleq \sum_{i=1}^{m_X} \xi_i^X Q_{if}^X$, $Q^X \triangleq \sum_{i=1}^{m_X} \xi_i^X Q_i^X$, $R_j^{XP} \triangleq \sum_{i=1}^{m_X} \xi_i^X R_{ij}^{XP}$, and $R_j^{XE} \triangleq \sum_{i=1}^{m_X} \xi_i^X R_{ij}^{XE}$. Then, the aggregate performance on a realization can be rewritten as

$$J_{\xi^X}^X(x_0; \hat{u}_C^P(\cdot), \hat{u}_C^E(\cdot)) = x^T(t_f)Q_f^X x(t_f) + \int_{t_0}^{t_f} \left[x^T(\tau)Q^X(\tau)x(\tau) + \sum_{j=1}^{m_P} (\hat{u}_{cj}^P)^T(\tau)R_j^{XP}(\tau)(\hat{u}_{cj}^P)^P(\tau) + \sum_{j=1}^{m_E} (\hat{u}_{cj}^E)^T(\tau)R_j^{XE}(\tau)(\hat{u}_{cj}^E)^E(\tau) \right] d\tau. \quad (8)$$

For a compact notation, let $R^{XP} \triangleq \text{diag}(R_1^{XP}, \dots, R_{m_X}^{XP})$, and $R^{XE} \triangleq \text{diag}(R_1^{XE}, \dots, R_{m_X}^{XE})$. The negotiating performance-measure (8) associated with team X becomes

$$J_{\xi^X}^X(x_0; \hat{u}_C^P(\cdot), \hat{u}_C^E(\cdot)) = x^T(t_f)Q_f^X x(t_f) + \int_{t_0}^{t_f} \left[x^T(\tau)Q^X(\tau)x(\tau) + (\hat{u}_C^P)^T(\tau)R^{XP}(\tau)\hat{u}_C^P(\tau) + (\hat{u}_C^E)^T(\tau)R^{XE}(\tau)\hat{u}_C^E(\tau) \right] d\tau. \quad (9)$$

In fact, the game is zero-sum only if $R^{PP} = -R^{EP} \triangleq R_P$, $R^{EE} = -R^{PE} \triangleq R_E$, $Q_f^P = -Q_f^E \triangleq Q_f$, and $Q^P =$

$-Q^E \triangleq Q$. Replacing these insights into (9), one obtains the realized performance-measure for the pursuit-evasion

$$J_{\xi}(x_0; \hat{u}_C^P(\cdot), \hat{u}_C^E(\cdot)) = x^T(t_f)Q_f x(t_f) + \int_{t_0}^{t_f} \left[x^T(\tau)Q(\tau)x(\tau) + (\hat{u}_C^P)^T(\tau)R^P(\tau)\hat{u}_C^P(\tau) - (\hat{u}_C^E)^T(\tau)R^E(\tau)\hat{u}_C^E(\tau) \right] d\tau. \quad (10)$$

In addition, the aggregate system (3) in the absence of uncertain disturbances is assumed to be uniformly exponentially stable. For instance, there exist positive constants η_1 and η_2 such that the pointwise matrix norm of the closed-loop state transition matrix for pursuit-evasion satisfies the inequality

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

The pair $(A(t), [B^P(t), B^E(t)])$ is pointwise stabilizable if there exist bounded matrix-valued functions $K^P(t)$ and $K^E(t)$ so that the closed-loop system $dx(t) = (A(t) + B^P(t)K^P(t) + B^E(t)K^E(t))x(t)dt$ is uniformly exponentially stable.

As shown in [6], all cumulants of the IQF random measure of performance have the same quadratic-affine functional form given the linear-quadratic structure of the system considered therein. This common form of the cumulants facilitates the definition of a cumulant-based performance index and the formulation of an optimization problem involving a finite number of the cumulants of J_{ξ} . Therefore, it is reasonable to restrict the search for decision strategies to linear time-varying decision laws by the rules of action

$$\hat{u}_C^P(t) = \hat{\gamma}_C^P(t, x(t)) \triangleq K^P(t)x(t), \quad (11)$$

$$\hat{u}_C^E(t) = \hat{\gamma}_C^E(t, x(t)) \triangleq K^E(t)x(t) \quad (12)$$

where $K^P \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{\sum_{i=1}^{m_P} m_i^P \times (\sum_{i=1}^{m_P} n_i^P + \sum_{j=1}^{m_E} n_j^E)})$ and $K^E \in \mathcal{C}([t_0, t_f]; \mathbb{R}^{\sum_{j=1}^{m_E} m_j^E \times (\sum_{i=1}^{m_P} n_i^P + \sum_{j=1}^{m_E} n_j^E)})$ are admissible gains for teams P and E . With the decision strategies (11)-(12), the dynamics of pursuit-evasion (3) becomes

$$dx(t) = [A(t) + B^P(t)K^P(t) + B^E(t)K^E(t)]x(t)dt + G(t)dw(t), \quad x(t_0) = x_0 \quad (13)$$

together with its realized measure of performance

$$J_{\xi}(x_0; K^P(\cdot), K^E(\cdot)) = x^T(t_f)Q_f x(t_f) + \int_{t_0}^{t_f} x^T(\tau)[Q(\tau) + (K^P)^T(\tau)R^P(\tau)K^P(\tau) - (K^E)^T(\tau)R^E(\tau)K^E(\tau)]x(\tau)d\tau \quad (14)$$

Clearly then, the performance-measure (14) is now a random variable with Chi-squared type. Hence, the uncertainty of shared performance distribution for conflicting teams must be assessed via a complete set of higher-order statistics beyond the statistical averaging. Next compact and robust results not only offer a tractable method for higher-order characteristics of (14) but also allow the incorporation of a class of linear feedback decision syntheses as being utilized in the cumulant-based control optimization.

Theorem 1: Cumulants in Multi-Player Pursuit-Evasion. Suppose the multi-player pursuit-evasion is described by (13)-(14) where (A, B^P) and (A, B^E) are uniformly stabilizable. Two rival teams respectively choose coalitive decisions $(\hat{u}_C^P(t), \hat{u}_C^E(t)) = (K^P(t)x(t), K^E(t)x(t))$. For any given $k \in \mathbb{Z}^+$, $\xi^P \in W^P$, and $\xi^E \in W^E$, the k^{th} cumulant in multi-player pursuit-evasion is given by

$$\kappa_k(t_0, x_0) = x_0^T H(t_0, k)x_0 + D(t_0, k) \quad (15)$$

where the cumulant-generating components $\{H(\alpha, i)\}_{i=1}^k$ and $\{D(\alpha, i)\}_{i=1}^k$ evaluated at $\alpha = t_0$ satisfy the following differential equations (with the dependence of $H(\alpha, i)$ and $D(\alpha, i)$ upon the admissible gains K^P and K^E suppressed)

$$\begin{aligned} \frac{d}{d\alpha} H(\alpha, 1) = & \\ & - [A(\alpha) + B^P(\alpha)K^P(\alpha) + B^E(\alpha)K^E(\alpha)]^T H(\alpha, 1) \\ & - H(\alpha, 1)[A(\alpha) + B^P(\alpha)K^P(\alpha) + B^E(\alpha)K^E(\alpha)] - Q(\alpha) \\ & - (K^P)^T(\alpha)R^P(\alpha)K^P(\alpha) + (K^E)^T(\alpha)R^E(\alpha)K^E(\alpha) \end{aligned} \quad (16)$$

and, for $2 \leq i \leq k$

$$\begin{aligned} \frac{d}{d\alpha} H(\alpha, i) = & \\ & - [A(\alpha) + B^P(\alpha)K^P(\alpha) + B^E(\alpha)K^E(\alpha)]^T H(\alpha, i) \\ & - H(\alpha, i)[A(\alpha) + B^P(\alpha)K^P(\alpha) + B^E(\alpha)K^E(\alpha)] \\ & - \sum_{j=1}^{i-1} \frac{2i!}{j!(i-j)!} H(\alpha, j)G(\alpha)WG^T(\alpha)H(\alpha, i-j), \end{aligned} \quad (17)$$

together with $1 \leq i \leq k$

$$\frac{d}{d\alpha} D(\alpha, i) = -\text{Tr} \{H(\alpha, i)G(\alpha)WG^T(\alpha)\} \quad (18)$$

where the terminal-value conditions $H(t_f, 1) = Q_f$, $H(t_f, i) = 0$ for $2 \leq i \leq k$ and $D(t_f, i) = 0$ for $1 \leq i \leq k$.

III. PROBLEM STATEMENTS

Suffice it to say here that all the performance-measure cumulant values (15) depend in part of the known initial condition $x(t_0)$. Although different states $x(t)$ will result in different values for the ‘‘cost-to-go’’ where (14) is redefined by replacing the lower integration limit t_0 with the running time variable α , the cumulant values are however, functions of time-backward evolutions of the cumulant-generating components $H(\alpha, i)$ and $D(\alpha, i)$ that totally ignore all the intermediate values $x(t)$. This fact therefore makes the new optimization problem particularly unique as compared with the more traditional dynamic programming class of investigations. Consequently, the time-backward trajectories (16)-(18) should be considered as the ‘‘new’’ dynamical equations from which the resulting Mayer optimization [2] and associated value function in the framework of dynamic programming therefore depend on these ‘‘new’’ states $H(\alpha, i)$ and $D(\alpha, i)$, not the classical states $x(t)$ as the people may often expect.

In the subsequent development, the subset of symmetric matrices of the vector space of all $n \times n$ matrices with real elements is denoted by \mathbb{S}^n where $n \triangleq \sum_{i=1}^{m_P} n_i^P + \sum_{j=1}^{m_E} n_j^E$.

Now let k -tuple variables \mathcal{H} and \mathcal{D} be defined as follows $\mathcal{H}(\cdot) \triangleq (\mathcal{H}_1(\cdot), \dots, \mathcal{H}_k(\cdot))$ and $\mathcal{D}(\cdot) \triangleq (\mathcal{D}_1(\cdot), \dots, \mathcal{D}_k(\cdot))$ for each element $\mathcal{H}_i \in \mathcal{C}^1([t_0, t_f]; \mathbb{S}^n)$ of \mathcal{H} and $\mathcal{D}_i \in \mathcal{C}^1([t_0, t_f]; \mathbb{R})$ of \mathcal{D} having the representations $\mathcal{H}_i(\cdot) = H(\cdot, i)$ and $\mathcal{D}_i(\cdot) = D(\cdot, i)$ with the right members satisfying the dynamic equations (16)-(18) on the horizon $[t_0, t_f]$. For notational tractability, the following mappings are introduced

$$\begin{aligned} \mathcal{F}_i &: [t_0, t_f] \times (\mathbb{S}^n)^k \times \mathbb{R}^{m^P \times n} \times \mathbb{R}^{m^E \times n} \mapsto \mathbb{S}^n \\ \mathcal{G}_i &: [t_0, t_f] \times (\mathbb{S}^n)^k \mapsto \mathbb{R} \end{aligned}$$

where $m^P \triangleq \sum_{i=1}^{m^P} m_i^P$, $m^E \triangleq \sum_{j=1}^{m^E} m_j^E$, and the actions are given by

$$\begin{aligned} \mathcal{F}_1(\alpha, \mathcal{H}, K^P, K^E) &\triangleq -Q(\alpha) \\ &- [A(\alpha) + B^P(\alpha)K^P(\alpha) + B^E(\alpha)K^E(\alpha)]^T \mathcal{H}_1(\alpha) \\ &- \mathcal{H}_1(\alpha) [A(\alpha) + B^P(\alpha)K^P(\alpha) + B^E(\alpha)K^E(\alpha)] \\ &- (K^P)^T(\alpha)R^P(\alpha)K^P(\alpha) + (K^E)^T(\alpha)R^E(\alpha)K^E(\alpha), \\ \mathcal{F}_i(\alpha, \mathcal{H}, K^P, K^E) &\triangleq \\ &- [A(\alpha) + B^P(\alpha)K^P(\alpha) + B^E(\alpha)K^E(\alpha)]^T \mathcal{H}_i(\alpha) \\ &- \mathcal{H}_i(\alpha) [A(\alpha) + B^P(\alpha)K^P(\alpha) + B^E(\alpha)K^E(\alpha)] \\ &- \sum_{j=1}^{i-1} \frac{2i!}{j!(i-j)!} \mathcal{H}_j(\alpha)G(\alpha)WG^T(\alpha)\mathcal{H}_{i-j}(\alpha), \quad 2 \leq i \leq k \\ \mathcal{G}_i(\alpha, \mathcal{H}) &\triangleq -\text{Tr} \{ \mathcal{H}_i(\alpha)G(\alpha)WG^T(\alpha) \}, \quad 1 \leq i \leq k. \end{aligned}$$

For a compact formulation, the product mappings follow

$$\begin{aligned} \mathcal{F}_1 \times \dots \times \mathcal{F}_k &: [t_0, t_f] \times (\mathbb{S}^n)^k \times \mathbb{R}^{m^P \times n} \times \mathbb{R}^{m^E \times n} \mapsto (\mathbb{S}^n)^k \\ \mathcal{G}_1 \times \dots \times \mathcal{G}_k &: [t_0, t_f] \times (\mathbb{S}^n)^k \mapsto \mathbb{R}^k \end{aligned}$$

along with the corresponding notations $\mathcal{F} \triangleq \mathcal{F}_1 \times \dots \times \mathcal{F}_k$ and $\mathcal{G} \triangleq \mathcal{G}_1 \times \dots \times \mathcal{G}_k$. Thus, the dynamic equations of motion (16)-(18) can be rewritten as follows

$$\frac{d}{d\alpha} \mathcal{H}(\alpha) = \mathcal{F}(\alpha, \mathcal{H}(\alpha), K^P(\alpha), K^E(\alpha)), \quad \mathcal{H}(t_f) \quad (19)$$

$$\frac{d}{d\alpha} \mathcal{D}(\alpha) = \mathcal{G}(\alpha, \mathcal{H}(\alpha)), \quad \mathcal{D}(t_f) \quad (20)$$

where $\mathcal{H}(t_f) = (Q_f, 0, \dots, 0)$ and $\mathcal{D}(t_f) = (0, \dots, 0)$.

Note that the product system uniquely determines \mathcal{H} and \mathcal{D} once the admissible decision gains K^P and K^E are specified. Hence, \mathcal{H} and \mathcal{D} are considered as $\mathcal{H}(\cdot, K^P, K^E)$ and $\mathcal{D}(\cdot, K^P, K^E)$, respectively. The performance index in cumulant-based control optimization can now be formulated in the admissible decision gains K^P and K^E .

Definition 2: Performance Index.

Fix $k \in \mathbb{Z}^+$ and $\mu = \{\mu_i \geq 0\}_{i=1}^k$ with $\mu_1 > 0$. Then for given (t_0, x_0) , $\xi^P \in W^P$, and $\xi^E \in W^E$, the performance index $\phi_0 : \{t_0\} \times (\mathbb{S}^n)^k \times \mathbb{R}^k \mapsto \mathbb{R}^+$ for the cumulant-based optimization is defined as follows

$$\begin{aligned} \phi_0(t_0, \mathcal{H}(t_0, K^P, K^E), \mathcal{D}(t_0, K^P, K^E)) &\triangleq \sum_{i=1}^k \mu_i \kappa_i(K^P, K^E) \\ &= \sum_{i=1}^k \mu_i [x_0^T \mathcal{H}_i(t_0, K^P, K^E)x_0 + \mathcal{D}_i(t_0, K^P, K^E)] \quad (21) \end{aligned}$$

where the parametric design freedom μ_i mutually chosen by adversarial teams represent different levels of influence as they deem important to the realized performance distribution and cumulant-supporting solutions $\{\mathcal{H}_i(t_0, K^P, K^E) \geq 0\}_{i=1}^k$ and $\{\mathcal{D}_i(t_0, K^P, K^E) \geq 0\}_{i=1}^k$ evaluated at $\alpha = t_0$ satisfy the dynamical equations of motion (19)-(20).

For the given terminal data $(t_f, \mathcal{H}_f, \mathcal{D}_f)$, the classes $\mathcal{K}_{t_f, \mathcal{H}_f, \mathcal{D}_f; \xi^P, \xi^E; \mu}^P$ and $\mathcal{K}_{t_f, \mathcal{H}_f, \mathcal{D}_f; \xi^P, \xi^E; \mu}^E$ of admissible decision gains are then defined.

Definition 3: Admissible Decision Gain Parameters.

Let the compact subsets $\bar{K}^P \subset \mathbb{R}^{m^P \times n}$ and $\bar{K}^E \subset \mathbb{R}^{m^E \times n}$ be the sets of allowable gain values. For given $\xi^P \in W^P$, $\xi^E \in W^E$, $k \in \mathbb{Z}^+$, and $\mu = \{\mu_i \geq 0\}_{i=1}^k$ with $\mu_1 > 0$, the sets of admissible decision strategies $\mathcal{K}_{t_f, \mathcal{H}_f, \mathcal{D}_f; \xi^P, \xi^E; \mu}^P$ and $\mathcal{K}_{t_f, \mathcal{H}_f, \mathcal{D}_f; \xi^P, \xi^E; \mu}^E$ are assumed to be the classes of $\mathcal{C}([t_0, t_f]; \mathbb{R}^{m^P \times n})$ and $\mathcal{C}([t_0, t_f]; \mathbb{R}^{m^E \times n})$ with values $K^P(\cdot) \in \bar{K}^P$ and $K^E(\cdot) \in \bar{K}^E$ for which solutions to the dynamic equations (19)-(20) exist on $[t_0, t_f]$. If the game admits a saddle-point equilibrium in pure strategies, there exist $K^{P*} \in \mathcal{K}_{t_f, \mathcal{H}_f, \mathcal{D}_f; \xi^P, \xi^E; \mu}^P$ and $K^{E*} \in \mathcal{K}_{t_f, \mathcal{H}_f, \mathcal{D}_f; \xi^P, \xi^E; \mu}^E$, such that the lower bound $\phi_0(t_0, \mathcal{H}(t_0, K^{P*}, K^{E*}), \mathcal{D}(t_0, K^{P*}, K^{E*}))$ is less than the saddle-point $\phi_0(t_0, \mathcal{H}(t_0, K^{P*}, K^{E*}), \mathcal{D}(t_0, K^{P*}, K^{E*}))$ which is also bounded from above by the upper bound $\phi_0(t_0, \mathcal{H}(t_0, K^P, K^{E*}), \mathcal{D}(t_0, K^P, K^{E*}))$ for any arbitrary $K^P \in \mathcal{K}_{t_f, \mathcal{H}_f, \mathcal{D}_f; \xi^P, \xi^E; \mu}^P$ and $K^E \in \mathcal{K}_{t_f, \mathcal{H}_f, \mathcal{D}_f; \xi^P, \xi^E; \mu}^E$.

Theorem 2: Existence of Saddle-Point Equilibrium.

If $\mathcal{K}_{t_f, \mathcal{H}_f, \mathcal{D}_f; \xi^P, \xi^E; \mu}^P$ and $\mathcal{K}_{t_f, \mathcal{H}_f, \mathcal{D}_f; \xi^P, \xi^E; \mu}^E$ are nonempty, compact and convex and the continuous performance index $\phi_0(t_0, \mathcal{H}(t_0, K^P, K^E), \mathcal{D}(t_0, K^P, K^E))$ is strictly convex in K^P and strictly concave in K^E , the pursuit-evasion admits a saddle-point equilibrium in pure strategies.

Proof: The sets of pure strategies $\mathcal{K}_{t_f, \mathcal{H}_f, \mathcal{D}_f; \xi^P, \xi^E; \mu}^P$ and $\mathcal{K}_{t_f, \mathcal{H}_f, \mathcal{D}_f; \xi^P, \xi^E; \mu}^E$ are nonempty, compact and convex. What remains is to show the continuous function $\phi_0(t_0, \mathcal{H}(t_0, K^P, K^E), \mathcal{D}(t_0, K^P, K^E))$ is strictly convex in K^P and strictly concave in K^E . It can be done by aggregating the supporting equations (16)-(17) as follows

$$\begin{aligned} \frac{d}{d\alpha} \Lambda(\alpha) &= -\mu_1 Q(\alpha) \\ &- [A(\alpha) + B^P(\alpha)K^P(\alpha) + B^E(\alpha)K^E(\alpha)]^T \Lambda(\alpha) \\ &- \Lambda(\alpha) [A(\alpha) + B^P(\alpha)K^P(\alpha) + B^E(\alpha)K^E(\alpha)] \\ &- \mu_1 (K^P)^T(\alpha)R^P(\alpha)K^P(\alpha) + \mu_1 (K^E)^T(\alpha)R^E(\alpha)K^E(\alpha) \\ &- \sum_{i=2}^k \mu_i \sum_{j=1}^{i-1} \frac{2i!}{j!(i-j)!} \mathcal{H}_j(\alpha)G(\alpha)WG^T(\alpha)\mathcal{H}_{i-j}(\alpha) \quad (22) \end{aligned}$$

where $\Lambda(\alpha) \triangleq \sum_{i=1}^k \mu_i \mathcal{H}_i(\alpha)$ and $\Lambda(t_f) = \mu_1 Q_f$. The fundamental theorem of calculus and stochastic differential rule applied to $x^T(\tau)\Lambda(\tau)x(\tau)$ yield the result

$$\begin{aligned} E \{ x^T(t_f)Q_f x(t_f) \} - x_0^T \Lambda(t_0)x_0 &= \\ E \left\{ \int_{t_0}^{t_f} d [x^T(\tau)\Lambda(\tau)x(\tau)] \right\} \end{aligned}$$

$$\begin{aligned}
&= E \left\{ \int_{t_0}^{t_f} [dx^T(\tau)\Lambda(\tau)x(\tau) + x^T(\tau)\Lambda(\tau)dx(\tau)] \right\} \\
&\quad + E \left\{ \int_{t_0}^{t_f} x^T(\tau) \frac{d}{d\tau} \Lambda(\tau) x(\tau) d\tau \right\} \\
&\quad + E \left\{ \int_{t_0}^{t_f} dx^T(\tau)\Lambda(\tau)dx(\tau) \right\}.
\end{aligned}$$

After some manipulations, it follows that

$$\begin{aligned}
&E \{ x^T(t_f)Q_f x(t_f) \} - x_0^T \Lambda(t_0)x_0 \\
&= E \left\{ \int_{t_0}^{t_f} x^T(\tau) [A(\tau) + B^P(\tau)K^P(\tau) + B^E(\tau)K^E(\tau)]^T \right. \\
&\quad \cdot \Lambda(\tau)x(\tau)d\tau \left. \right\} + E \left\{ \int_{t_0}^{t_f} x^T(\tau) \frac{d}{d\tau} \Lambda(\tau)x(\tau)d\tau \right\} \\
&+ E \left\{ \int_{t_0}^{t_f} x^T(\tau)\Lambda(\tau) [A(\tau) + B^P(\tau)K^P(\tau) + B^E(\tau)K^E(\tau)] \right. \\
&\quad \left. x(\tau)d\tau \right\} + \int_{t_0}^{t_f} \text{Tr} \{ \Lambda(\tau)G(\tau)WG^T(\tau) \} d\tau. \quad (23)
\end{aligned}$$

Note that an integral form of the solution of (18) is given by

$$\mathcal{D}_i(t_0) = \int_{t_0}^{t_f} \text{Tr} \{ \mathcal{H}_i(\tau)G(\tau)WG^T(\tau) \} d\tau.$$

In view of the definition of $\Lambda(\cdot)$, it is then easy to see that

$$\sum_{i=1}^k \mu_i \mathcal{D}_i(t_0) = \int_{t_0}^{t_f} \text{Tr} \{ \Lambda(\tau)G(\tau)WG^T(\tau) \} d\tau.$$

Hence, the performance index (21) is now rewritten as

$$\begin{aligned}
&\phi_0(t_0, \mathcal{H}(t_0, K^P, K^E), \mathcal{D}(t_0, K^P, K^E)) \\
&= x_0^T \Lambda(t_0)x_0 + \int_{t_0}^{t_f} \text{Tr} \{ \Lambda(\tau)G(\tau)WG^T(\tau) \} d\tau. \quad (24)
\end{aligned}$$

Replacing the results (22) and (24) into (23), it yields

$$\begin{aligned}
&\phi_0(t_0, \mathcal{H}(t_0, K^P, K^E), \mathcal{D}(t_0, K^P, K^E)) = \quad (25) \\
&E \{ x^T(t_f)Q_f x(t_f) \} + E \left\{ \int_{t_0}^{t_f} x^T(\tau) [\mu_1 Q(\tau) \right. \\
&\quad + \mu_1 (K^P)^T(\tau)R^P(\tau)K^P(\tau) - \mu_1 (K^E)^T(\tau)R^E(\tau)K^E(\tau) \\
&\quad \left. + \sum_{i=2}^k \mu_i \sum_{j=1}^{i-1} \frac{2i!}{j!(i-j)!} \mathcal{H}_j(\tau)G(\tau)WG^T(\tau)\mathcal{H}_{i-j}(\tau)] x(\tau)d\tau \right\}
\end{aligned}$$

which leads to

$$\begin{aligned}
&\phi_0(t_0, \mathcal{H}(t_0, K^P, K^E), \mathcal{D}(t_0, K^P, K^E)) = \quad (26) \\
&\text{Tr} \{ Q_f P(t_f) \} + \text{Tr} \left\{ \int_{t_0}^{t_f} [\mu_1 Q(\tau) \right. \\
&\quad + \mu_1 (K^P)^T(\tau)R^P(\tau)K^P(\tau) - \mu_1 (K^E)^T(\tau)R^E(\tau)K^E(\tau) \\
&\quad \left. + \sum_{i=2}^k \mu_i \sum_{j=1}^{i-1} \frac{2i!}{j!(i-j)!} \mathcal{H}_j(\tau)G(\tau)WG^T(\tau)\mathcal{H}_{i-j}(\tau)] P(\tau)d\tau \right\}
\end{aligned}$$

where $P(s) \triangleq E\{x(s)x^T(s)\}$ with $P(t_0) = x_0x_0^T$ satisfies

$$\begin{aligned}
\frac{d}{ds}P(s) &= P(s)[A(s) + B^P(s)K^P(s) + B^E(s)K^E(s)]^T \quad (27) \\
&+ [A(s) + B^P(s)K^P(s) + B^E(s)K^E(s)]P(s) + G(s)WG^T(s)
\end{aligned}$$

Hence, the continuous function (21) is strictly convex in K^P and strictly concave in K^E . Therefore, the pursuit-evasion game considered here admits a saddle-point equilibrium in accordance of Proposition 3.3 in [1]. ■

Next, one may state the optimization problem for the stochastic multi-player pursuit-evasion.

Definition 4: Optimization Problem.

Fix $\xi^P \in W^P$, $\xi^E \in W^E$, $k \in \mathbb{Z}^+$, and $\mu = \{\mu_i \geq 0\}_{i=1}^k$ with $\mu_1 > 0$. Then, the optimization problem for multi-player pursuit-evasion over $[t_0, t_f]$ is given by

$$\begin{aligned}
&\min_{K^P(\cdot) \in \mathcal{K}_{t_f, \mathcal{H}_f, \mathcal{D}_f; \xi^P, \xi^E, \mu}^P} \max_{K^E(\cdot) \in \mathcal{K}_{t_f, \mathcal{H}_f, \mathcal{D}_f; \xi^P, \xi^E, \mu}^E} \\
&\phi_0(t_0, \mathcal{H}(t_0, K^P, K^E), \mathcal{D}(t_0, K^P, K^E)) \quad (28)
\end{aligned}$$

subject to the dynamic equations (19)-(20) for $\alpha \in [t_0, t_f]$.

As a tenet of transition from the principle of optimality, a family of games based on different starting points is now of concerned. With an intermission of time, ε in mid-play considered here, the path has reached some definitive point at its commencement. There exist some potential $(\mathcal{H}, \mathcal{D})$ which may be reached at the end of the intermission for all possible choices of (K^P, K^E) .

Suppose that for each endpoint, the game beginning there has already been solved. Then, the value function $\mathcal{V}(\varepsilon, \mathcal{H}, \mathcal{D})$ resulting from each choice of (K^P, K^E) is known, and they are to be so chosen as to render it minimax. As the duration of the intermission approaches t_f together with the existence of a saddle-point equilibrium, it leads to the necessary and sufficient Hamilton-Jacobi-Isaacs (HJI) condition.

Definition 5: Playable Set.

Let playable set \mathcal{Q} be defined as $\mathcal{Q} \triangleq \{(\varepsilon, \mathcal{Y}, \mathcal{Z}) \in [t_0, t_f] \times (\mathbb{S}^n)^k \times \mathbb{R}^k \text{ such that } \mathcal{K}_{\varepsilon, \mathcal{Y}, \mathcal{Z}; \xi^P, \xi^E; \mu}^P \times \mathcal{K}_{\varepsilon, \mathcal{Y}, \mathcal{Z}; \xi^P, \xi^E; \mu}^E \neq \emptyset\}$.

Theorem 3: HJI Equation-Mayer Problem.

Let $(\varepsilon, \mathcal{Y}, \mathcal{Z})$ be any interior point of the playable set \mathcal{Q} at which the value function $\mathcal{V}(\varepsilon, \mathcal{Y}, \mathcal{Z})$ is differentiable. If there exists a saddle point $(K^{P*}, K^{E*}) \in \mathcal{K}_{\varepsilon, \mathcal{Y}, \mathcal{Z}; \xi^P, \xi^E; \mu}^P \times \mathcal{K}_{\varepsilon, \mathcal{Y}, \mathcal{Z}; \xi^P, \xi^E; \mu}^E$, then the partial differential equation

$$\begin{aligned}
0 &= \min_{K^P \in \bar{K}^P} \max_{K^E \in \bar{K}^E} \left\{ \frac{\partial}{\partial \varepsilon} \mathcal{V}(\varepsilon, \mathcal{Y}, \mathcal{Z}) \right. \\
&\quad + \frac{\partial}{\partial \text{vec}(\mathcal{Y})} \mathcal{V}(\varepsilon, \mathcal{Y}, \mathcal{Z}) \cdot \text{vec}(\mathcal{F}(\varepsilon, \mathcal{Y}, K^P, K^E)) \\
&\quad \left. + \frac{\partial}{\partial \text{vec}(\mathcal{Z})} \mathcal{V}(\varepsilon, \mathcal{Y}, \mathcal{Z}) \cdot \text{vec}(\mathcal{G}(\varepsilon, \mathcal{Y})) \right\} \quad (29)
\end{aligned}$$

is satisfied together with $\mathcal{V}(t_0, \mathcal{H}_0, \mathcal{D}_0) = \phi_0(t_0, \mathcal{H}_0, \mathcal{D}_0)$ and $\text{vec}(\cdot)$ the vectorizing operator of enclosed entities.

IV. SADDLE-POINT 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 [2]. Often, it is required to denote the terminal time and states of a family of optimization problems as $(\varepsilon, \mathcal{Y}, \mathcal{Z})$ rather than $(t_f, \mathcal{H}_f, \mathcal{D}_f)$. That is, for $\varepsilon \in [t_0, t_f]$, the states of (19)-(20) defined on $[t_0, \varepsilon]$ have the terminal values denoted by $\mathcal{H}(\varepsilon) \equiv \mathcal{Y}$ and $\mathcal{D}(\varepsilon) \equiv \mathcal{Z}$. Since the performance index (21) is quadratic affine in terms of arbitrarily fixed x_0 , this observation suggests a solution to the HJI equation (29) may be of the form as follows.

For any given interior point $(\varepsilon, \mathcal{Y}, \mathcal{Z})$ of the reachable set \mathcal{Q} at which the real-valued function $\mathcal{W}(\varepsilon, \mathcal{Y}, \mathcal{Z}) \triangleq x_0^T \sum_{i=1}^k \mu_i (\mathcal{Y}_i + \mathcal{E}_i(\varepsilon)) x_0 + \sum_{i=1}^k \mu_i (\mathcal{Z}_i + \mathcal{T}_i(\varepsilon))$ is differentiable and time parametric functions $\mathcal{E}_i \in \mathcal{C}^1([t_0, t_f]; \mathbb{S}^n)$ and $\mathcal{T}_i \in \mathcal{C}^1([t_0, t_f]; \mathbb{R})$ are yet to be determined. Moreover, the time derivative of $\mathcal{W}(\varepsilon, \mathcal{Y}, \mathcal{Z})$ can be obtained as below

$$\begin{aligned} \frac{d}{d\varepsilon} \mathcal{W}(\varepsilon, \mathcal{Y}, \mathcal{Z}) &= \sum_{i=1}^k \mu_i \left(\mathcal{G}_i(\varepsilon, \mathcal{Y}) + \frac{d}{d\varepsilon} \mathcal{T}_i(\varepsilon) \right) \\ &+ x_0^T \sum_{i=1}^k \mu_i \left(\mathcal{F}_i(\varepsilon, \mathcal{Y}, K^P, K^E) + \frac{d}{d\varepsilon} \mathcal{E}_i(\varepsilon) \right) x_0. \end{aligned} \quad (30)$$

The substitution of this candidate for the value function into the HJI equation (29) and making use of the result (30) yield

$$\begin{aligned} 0 \equiv \min_{K^P \in \bar{K}^P} \max_{K^E \in \bar{K}^E} \left\{ x_0^T \sum_{i=1}^k \mu_i \frac{d}{d\varepsilon} \mathcal{E}_i(\varepsilon) x_0 + \sum_{i=1}^k \mu_i \frac{d}{d\varepsilon} \mathcal{T}_i(\varepsilon) \right. \\ \left. + x_0^T \sum_{i=1}^k \mu_i \mathcal{F}_i(\varepsilon, \mathcal{Y}, K^P, K^E) x_0 + \sum_{i=1}^k \mu_i \mathcal{G}_i(\varepsilon, \mathcal{Y}) \right\} \end{aligned} \quad (31)$$

Differentiating the expression within the bracket of (31) with respect to K^P and K^E yield the necessary conditions for an extremum of the performance index (21) on $[t_0, \varepsilon]$

$$K^P(\varepsilon, \mathcal{Y}) = -(R^P)^{-1}(\varepsilon)(B^P)^T(\varepsilon) \sum_{r=1}^k \hat{\mu}_r \mathcal{Y}_r, \quad (32)$$

$$K^E(\varepsilon, \mathcal{Y}) = (R^E)^{-1}(\varepsilon)(B^E)^T(\varepsilon) \sum_{r=1}^k \hat{\mu}_r \mathcal{Y}_r \quad (33)$$

where the normalized weights $\hat{\mu}_r \triangleq \mu_r / \mu_1$ with $\mu_1 > 0$. Substituting the gain expressions (32) and (33) into the right member of the HJI equation (31) yields the value of minimax

$$\begin{aligned} x_0^T \left[\sum_{i=1}^k \mu_i \frac{d}{d\varepsilon} \mathcal{E}_i(\varepsilon) - A^T(\varepsilon) \sum_{i=1}^k \mu_i \mathcal{Y}_i - \sum_{i=1}^k \mu_i \mathcal{Y}_i A(\varepsilon) \right. \\ \left. - \mu_1 Q(\varepsilon) + \sum_{r=1}^k \hat{\mu}_r \mathcal{Y}_r B^P(\varepsilon)(R^P)^{-1}(\varepsilon)(B^P)^T(\varepsilon) \sum_{i=1}^k \mu_i \mathcal{Y}_i \right. \\ \left. + \sum_{i=1}^k \mu_i \mathcal{Y}_i B^P(\varepsilon)(R^P)^{-1}(\varepsilon)(B^P)^T(\varepsilon) \sum_{s=1}^k \hat{\mu}_s \mathcal{Y}_s \right. \\ \left. - \sum_{r=1}^k \hat{\mu}_r \mathcal{Y}_r B^E(\varepsilon)(R^E)^{-1}(\varepsilon)(B^E)^T(\varepsilon) \sum_{i=1}^k \mu_i \mathcal{Y}_i \right. \end{aligned}$$

$$\begin{aligned} \left. - \sum_{i=1}^k \mu_i \mathcal{Y}_i B^E(\varepsilon)(R^E)^{-1}(\varepsilon)(B^E)^T(\varepsilon) \sum_{s=1}^k \hat{\mu}_s \mathcal{Y}_s \right. \\ \left. - \mu_1 \sum_{r=1}^k \hat{\mu}_r \mathcal{Y}_r B^P(\varepsilon)(R^P)^{-1}(\varepsilon)(B^P)^T(\varepsilon) \sum_{s=1}^k \hat{\mu}_s \mathcal{Y}_s \right. \\ \left. + \mu_1 \sum_{r=1}^k \hat{\mu}_r \mathcal{Y}_r B^E(\varepsilon)(R^E)^{-1}(\varepsilon)(B^E)^T(\varepsilon) \sum_{s=1}^k \hat{\mu}_s \mathcal{Y}_s \right. \\ \left. - \sum_{i=2}^k \mu_i \sum_{j=1}^{i-1} \frac{2i!}{j!(i-j)!} \mathcal{Y}_j G(\varepsilon) W G^T(\varepsilon) \mathcal{Y}_{i-j} \right] x_0 \\ + \sum_{i=1}^k \mu_i \frac{d}{d\varepsilon} \mathcal{T}_i(\varepsilon) - \sum_{i=1}^k \mu_i \text{Tr} \{ \mathcal{Y}_i G(\varepsilon) W G^T(\varepsilon) \}. \end{aligned} \quad (34)$$

It is now necessary to exhibit functions $\{\mathcal{E}_i(\cdot)\}_{i=1}^k$ and $\{\mathcal{T}_i(\cdot)\}_{i=1}^k$ which will render the left side of (34) equal to zero for $\varepsilon \in [t_0, t_f]$, when $\{\mathcal{Y}_i\}_{i=1}^k$ are evaluated along solution trajectories of the cumulant-generating equations. Studying the expression (34) reveals the following results

$$\frac{d}{d\varepsilon} \mathcal{E}_1(\varepsilon) = A^T(\varepsilon) \mathcal{H}_1(\varepsilon) + \mathcal{H}_1(\varepsilon) A(\varepsilon) + Q(\varepsilon) \quad (35)$$

$$\begin{aligned} - \mathcal{H}_1(\varepsilon) B^P(\varepsilon)(R^P)^{-1}(\varepsilon)(B^P)^T(\varepsilon) \sum_{s=1}^k \hat{\mu}_s \mathcal{H}_s(\varepsilon) \\ - \sum_{r=1}^k \hat{\mu}_r \mathcal{H}_r(\varepsilon) B^P(\varepsilon)(R^P)^{-1}(\varepsilon)(B^P)^T(\varepsilon) \mathcal{H}_1(\varepsilon) \\ + \mathcal{H}_1(\varepsilon) B^E(\varepsilon)(R^E)^{-1}(\varepsilon)(B^E)^T(\varepsilon) \sum_{s=1}^k \hat{\mu}_s \mathcal{H}_s(\varepsilon) \\ + \sum_{r=1}^k \hat{\mu}_r \mathcal{H}_r(\varepsilon) B^E(\varepsilon)(R^E)^{-1}(\varepsilon)(B^E)^T(\varepsilon) \mathcal{H}_1(\varepsilon) \\ + \sum_{r=1}^k \hat{\mu}_r \mathcal{H}_r(\varepsilon) B^P(\varepsilon)(R^P)^{-1}(\varepsilon)(B^P)^T(\varepsilon) \sum_{s=1}^k \hat{\mu}_s \mathcal{H}_s(\varepsilon) \\ - \sum_{r=1}^k \hat{\mu}_r \mathcal{H}_r(\varepsilon) B^E(\varepsilon)(R^E)^{-1}(\varepsilon)(B^E)^T(\varepsilon) \sum_{s=1}^k \hat{\mu}_s \mathcal{H}_s(\varepsilon) \end{aligned}$$

and, for $2 \leq i \leq k$

$$\begin{aligned} \frac{d}{d\varepsilon} \mathcal{E}_i(\varepsilon) &= A^T(\varepsilon) \mathcal{H}_i(\varepsilon) + \mathcal{H}_i(\varepsilon) A(\varepsilon) \\ &- \mathcal{H}_i(\varepsilon) B^P(\varepsilon)(R^P)^{-1}(\varepsilon)(B^P)^T(\varepsilon) \sum_{s=1}^k \hat{\mu}_s \mathcal{H}_s(\varepsilon) \\ &- \sum_{r=1}^k \hat{\mu}_r \mathcal{H}_r(\varepsilon) B^P(\varepsilon)(R^P)^{-1}(\varepsilon)(B^P)^T(\varepsilon) \mathcal{H}_i(\varepsilon) \\ &+ \mathcal{H}_i(\varepsilon) B^E(\varepsilon)(R^E)^{-1}(\varepsilon)(B^E)^T(\varepsilon) \sum_{s=1}^k \hat{\mu}_s \mathcal{H}_s(\varepsilon) \\ &+ \sum_{r=1}^k \hat{\mu}_r \mathcal{H}_r(\varepsilon) B^E(\varepsilon)(R^E)^{-1}(\varepsilon)(B^E)^T(\varepsilon) \mathcal{H}_i(\varepsilon) \\ &+ \sum_{j=1}^{i-1} \frac{2i!}{j!(i-j)!} \mathcal{H}_j(\varepsilon) G(\varepsilon) W G^T(\varepsilon) \mathcal{H}_{i-j}(\varepsilon) \end{aligned} \quad (36)$$

together, for $1 \leq i \leq k$

$$\frac{d}{d\varepsilon} \mathcal{T}_i(\varepsilon) = \text{Tr} \{ \mathcal{H}_i(\varepsilon) G(\varepsilon) W G^T(\varepsilon) \}, \quad 1 \leq i \leq k \quad (37)$$

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

$$\begin{aligned} x_0^T \sum_{i=1}^k \mu_i (\mathcal{H}_{i0} + \mathcal{E}_i(t_0)) x_0 + \sum_{i=1}^k \mu_i (\mathcal{D}_{i0} + \mathcal{T}_i(t_0)) \\ = x_0^T \sum_{i=1}^k \mu_i \mathcal{H}_{i0} x_0 + \sum_{i=1}^k \mu_i \mathcal{D}_{i0}. \end{aligned}$$

Thus, matching the boundary condition yields the corresponding initial value conditions $\mathcal{E}_i(t_0) = 0$ and $\mathcal{T}_i(t_0) = 0$ for the equations (35)-(37). Applying the decision gains specified in (32) and (33) along the solution trajectories of the equations (19)-(20), these equations become Riccati-type. Thus, whenever these corresponding equations admit solutions $\{\mathcal{H}_i(\cdot)\}_{i=1}^k$ and $\{\mathcal{D}_i(\cdot)\}_{i=1}^k$, the existence of $\{\mathcal{E}_i(\cdot)\}_{i=1}^k$ and $\{\mathcal{T}_i(\cdot)\}_{i=1}^k$ satisfying the equations (35)-(37) are then assured. By comparing the parametric equations (35)-(37) to those Riccati-type equations, one recognizes that these equations are related to one another by

$$\frac{d}{d\varepsilon} \mathcal{E}_i(\varepsilon) = -\frac{d}{d\varepsilon} \mathcal{H}_i(\varepsilon) \quad \text{and} \quad \frac{d}{d\varepsilon} \mathcal{T}_i(\varepsilon) = -\frac{d}{d\varepsilon} \mathcal{D}_i(\varepsilon)$$

for $1 \leq i \leq k$. Enforcing the initial value conditions of $\mathcal{E}_i(t_0) = 0$ and $\mathcal{T}_i(t_0) = 0$ uniquely implies that $\mathcal{E}_i(\varepsilon) = \mathcal{H}_i(t_0) - \mathcal{H}_i(\varepsilon)$ and $\mathcal{T}_i(\varepsilon) = \mathcal{D}_i(t_0) - \mathcal{D}_i(\varepsilon)$ for all $\varepsilon \in [t_0, t_f]$ and yields a value function $\mathcal{W}(\varepsilon, \mathcal{Y}, \mathcal{Z}) \equiv \mathcal{V}(\varepsilon, \mathcal{Y}, \mathcal{Z}) = x_0^T \sum_{i=1}^k \mu_i \mathcal{H}_i(t_0) x_0 + \sum_{i=1}^k \mu_i \mathcal{D}_i(t_0)$ for which the sufficient condition (29) of the verification theorem is satisfied. Therefore, the feedback decision gains for pursuit team P , (32) and evader team E , (33) optimizing the performance index stated in (21) become optimal

$$K^{P*}(\varepsilon) = -(R^P)^{-1}(\varepsilon) (B^P)^T(\varepsilon) \sum_{r=1}^k \hat{\mu}_r \mathcal{H}_r^*(\varepsilon), \quad (38)$$

$$K^{E*}(\varepsilon) = (R^E)^{-1}(\varepsilon) (B^E)^T(\varepsilon) \sum_{r=1}^k \hat{\mu}_r \mathcal{H}_r^*(\varepsilon). \quad (39)$$

Theorem 4: Strategies for Multi-Player Pursuit-Evasion.

Consider the multi-player pursuit-evasion game as described by (13)-(14) where the pairs (A, B^P) and (A, B^E) are uniformly stabilizable. Fix $\xi^P \in W^P$, $\xi^E \in W^E$, $k \in \mathbb{Z}^+$, and $\mu = \{\mu_i \geq 0\}_{i=1}^k$ with $\mu_1 > 0$. Then, the saddle-point equilibrium is achieved by non-cooperative decision gains

$$K^{P*}(\alpha) = -(R^P)^{-1}(\alpha) (B^P)^T(\alpha) \sum_{r=1}^k \hat{\mu}_r \mathcal{H}_r^*(\alpha) \quad (40)$$

$$K^{E*}(\alpha) = (R^E)^{-1}(\alpha) (B^E)^T(\alpha) \sum_{r=1}^k \hat{\mu}_r \mathcal{H}_r^*(\alpha) \quad (41)$$

where $\hat{\mu}_r$ mutually chosen by rival teams represent various robustness levels of the common performance distribution

and the optimal solutions $\{\mathcal{H}_r^*(\alpha) \geq 0\}_{r=1}^k$ are satisfying

$$\begin{aligned} \frac{d}{d\alpha} \mathcal{H}_1^*(\alpha) = & -Q(\alpha) - (K^{P*})^T(\alpha) R^P(\alpha) K^{P*}(\alpha) \\ & - [A(\alpha) + B^P(\alpha) K^{P*}(\alpha) + B^E(\alpha) K^{E*}(\alpha)]^T \mathcal{H}_1^*(\alpha) \\ & - \mathcal{H}_1^*(\alpha) [A(\alpha) + B^P(\alpha) K^{P*}(\alpha) + B^E(\alpha) K^{E*}(\alpha)] \\ & + (K^{E*})^T(\alpha) R^E(\alpha) K^{E*}(\alpha), \quad \mathcal{H}_1^*(t_f) = Q_f \end{aligned} \quad (42)$$

and, for $2 \leq r \leq k$ with $\mathcal{H}_r^*(t_f) = 0$

$$\begin{aligned} \frac{d}{d\alpha} \mathcal{H}_r^*(\alpha) = & - [A(\alpha) + B^P(\alpha) K^{P*}(\alpha) + B^E(\alpha) K^{E*}(\alpha)]^T \mathcal{H}_r^*(\alpha) \\ & - \mathcal{H}_r^*(\alpha) [A(\alpha) + B^P(\alpha) K^{P*}(\alpha) + B^E(\alpha) K^{E*}(\alpha)] \\ & - \sum_{s=1}^{r-1} \frac{2r!}{s!(r-s)!} \mathcal{H}_s^*(\alpha) G(\alpha) W G^T(\alpha) \mathcal{H}_{r-s}^*(\alpha). \end{aligned} \quad (43)$$

With the saddle-point equilibrium applied to the dynamics (13), the coupling Pareto-efficient laws (11)-(12) are partitioned consistently by the states $(x^*)^T = [(x^{P*})^T, (x^{E*})^T]$

$$\hat{u}_C^{P*}(t) = (K^{PP})^*(t) x^{P*}(t) + (K^{PE})^*(t) x^{E*}(t), \quad (44)$$

$$\hat{u}_C^{E*}(t) = (K^{EP})^*(t) x^{P*}(t) + (K^{EE})^*(t) x^{E*}(t) \quad (45)$$

where the components of Pareto-efficient decision gains $(K^{PP})^*$, $(K^{PE})^*$, $(K^{EP})^*$, and $(K^{EE})^*$ are determined by the results (40)-(41) and the partition of cumulant variables

$$\mathcal{H}_r^* = \begin{bmatrix} (\mathcal{H}_r^{PP})^* & (\mathcal{H}_r^{PE})^* \\ (\mathcal{H}_r^{EP})^* & (\mathcal{H}_r^{EE})^* \end{bmatrix}, \quad 1 \leq r \leq k$$

according to the division of the optimal states x^* .

V. CONCLUSIONS

This paper dealt with a multi-player pursuit-evasion differential game modeled in stochastic and adverse environments. Matrix differential equations for generating statistics of the standard integral-quadratic performance-measure used in this game were derived. A direct dynamic programming approach was used to solve for a novel saddle-point solution that can address both decision strategy selection and performance robustness. These results are expected to make new theoretical contributions and performance assessment tools for statistical learning and stochastic game-theoretic communities.

REFERENCES

- [1] T. Basar and G. J. Olsder, *Dynamic Non-cooperative Game Theory*, 2nd Edition, Society for Industrial and Applied Mathematics, 1998.
- [2] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*. New York: Springer-Verlag, 1975.
- [3] Isaacs, *Differential Games*, John Wiley & Sons, Inc. New York, 1965.
- [4] D. Li, J. B. Cruz, G. Chen, C. Kwan, and M.-H. Chang, "A Hierarchical Approach to Multi-Player Pursuit-Evasion Differential Games," *Proceedings of Joint IEEE Conference on CDC-ECC05*, Spain, 2005.
- [5] A. Klinger, "Vector-Valued Performance Criteria," *IEEE Transactions on Automatic Control*, Vol. AC-9, No. 1, pp. 117-118, 1964.
- [6] K. D. Pham, "Multi-Cumulant Control for Zero-Sum Differential Games: Performance-Measure Statistics and State-Feedback Paradigm," *Advances in Cooperative Control and Optimization*. Series Lecture Notes in Control and Information Sciences, Vol. 369 Hirsch, M. J.; Pardalos, P. M.; Murphey, R.; Grunzel, D. (Eds.), 2007.
- [7] R. Vidal, O. Shakernia, H. J. Kim, D. H. Shim and S. Sastry, "Probabilistic Pursuit-Evasion Games: Theory, Implementation, and Experimental Evaluation," *IEEE Transactions on Robotics and Automation*, Vol. 18, pp. 662-669, 2002.