Near-Optimal Mixed Strategy for Zero-Sum Linear-Quadratic Differential Games

Tao Xu, Wang Xi, and Jianping He

***Abstract*—Deriving analytic solutions for optimal mixed strategies in zero-sum linear-quadratic differential games (ZSLQDGs) remains an open problem. In this paper, we analytically synthesize near-optimal mixed strategies for ZSLQDGs and establish rigorous performance certifications. Specifically, we construct a surrogate pure-strategy stochastic differential game (SDG) by matching the first two moments of the mixed strategies. This method achieves an $\mathcal{O}(\bar{\pi}^2)$ weak approximation of state distributions and expected costs with respect to the maximum commitment delay $\bar{\pi}$. By analytically resolving the surrogate SDG, we derive closed-form optimal control laws for the matched moments. Crucially, we reveal that the surrogate game is governed by a Generalized Riccati Differential Equation (GRDE), which explicitly dictates a dynamic energy allocation law for variance injection. Building on these solutions, we propose a robust dual-routing architecture to execute the near-optimal mixed strategies. Furthermore, we certify that both the global value approximation error and the strategy suboptimality gaps are bounded by $\mathcal{O}(\bar{\pi}^{\frac{1}{2}})$. Finally, numerical experiments on a double-integrator pursuit-evasion game illustrate the induced physical behaviors and validate the theoretical bounds.**

***Index Terms*—Zero-Sum Linear-Quadratic Differential Game, Mixed Strategy, Stochastic Differential Game, Weak Approximation.**

I. Introduction

Zero-sum differential games, first formalized by Isaacs [2], provide a fundamental framework for modeling antagonistic interactions in continuous time, where one player's gain is exactly the opponent's loss. However, the generality of nonlinear dynamics and cost functionals in zero-sum differential games often leads to intractable Hamilton-Jacobi-Isaacs (HJI) equations. Restricting to linear dynamics and quadratic costs, a zero-sum linear-quadratic differential game (ZSLQDG) admits a remarkable simplification: the game value and optimal pure strategies can be analytically characterized via coupled Riccati differential equations [3]–[5]. This analytic solution not only guarantees global optimality but also facilitates real-time implementation. As a result, ZSLQDGs have become a cornerstone in applications ranging from pursuit–evasion and missile guidance to secure networked control [6]–[11].

While classical ZSLQDG theory focuses on pure strategies, where deterministic feedback rules map states to controls, a mixed-strategy formulation naturally extends this paradigm by allowing players to randomize over a family of feedback laws. Mixed strategies have been extensively studied in various fields such as economics [12], generative adversarial networks [13], reinforcement learning [14], robotics [15], and prominently in pursuit-evasion games [16]–[19]. In continuous-time games, however, naively defining mixed strategies encounters the fundamental pathology of instantaneous randomization: without a rigorous temporal definition of strategic commitment, each player might revise their randomized choice at the same instant, undermining any meaningful strategic commitment [20]–[22]. To resolve this, one must explicitly introduce a *commitment delay* into the definition of mixed strategies [23, p.114], ensuring that once a player samples a control action, it remains fixed for a prescribed time interval before the opponent can observe and respond. The commitment delay pattern is therefore essential both for the mathematical well-posedness of mixed strategy designs and for capturing realistic digital control scenarios where strategic interventions cannot be renegotiated infinitely fast.

A. Motivations

Classical pure-strategy equilibria are fully characterized by the elegant Riccati-equation framework. However, extending this methodology to randomized controls introduces formidable obstacles. Specifically, one must handle infinite-dimensional measure-valued spaces, enforce non-trivial commitment delays, and reconcile continuous-time dynamics with discrete-time probability laws. To date, closed-form analytic descriptions for optimal mixed strategies remain open. It leaves a critical gap in our understanding of how to systematically exploit randomization in continuous-time adversarial control.

Deriving analytic solutions for optimal mixed strategies in ZSLQDGs is essential for advancing the broader theory of differential games. First, closed-form mixed-strategy characterizations would explicitly illuminate the fundamental mechanisms of randomization, clarifying how players dynamically balance their expected control costs against the benefits of adversarial uncertainty. Second, establishing exact analytical benchmarks would dramatically accelerate the development of numerical methods for general mixed-strategy differential games, providing rigorous targets for algorithmic convergence and performance validation.

B. Challenges

The primary challenge in solving optimal mixed strategies for differential games stems from the inherent intractability of optimizing over infinite-dimensional probability measure spaces. As summarized in Table I, our prior works initiated a weak approximation framework to circumvent this issue. Here, "weak" refers to the convergence in distribution of the state-cost processes. We provide a rigorous mathematical treatment

This work was supported by the National Natural Science Foundation of China under Grant 62373247, 625B2117, 62633014. *(Corresponding author: Jianping He.)* The authors are with the Department of Automation, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240, China. E-mail: {Zerken, bddwyx, jphe}@sjtu.edu.cn.Preliminary results of this work were presented at IEEE Conference on Decision and Control (CDC), 2023 [1].

of this framework in Section III. However, these methods either lack explicit strategy synthesis [1] or rely heavily on numerical control-space discretization [24]. Consequently, these existing frameworks inherently yield numerical approximations rather than exact analytic laws, leaving a critical gap for solving ZSLQDGs analytically.

To derive analytic solutions for ZSLQDGs, we pivot to a fundamentally different surrogate construction: parameterizing Markovian mixed strategies via their statistical moments over unbounded spaces. This shift introduces a distinct theoretical hurdle in performance certification. Although the general bounding framework established in [24] remains applicable, it provides implicit, term-heavy upper bounds that obscure the exact scaling laws. The crux of the challenge lies in fully exploiting the linear-quadratic structure to derive a refined error analysis, which is essential to collapse these implicit execution penalties into clean, explicit $\mathcal{O}(\bar{\pi}^{1/2})$ error bounds.

C. Contributions

This paper presents the first closed-form analytical characterization of near-optimal mixed strategies for ZSLQDGs. As contrasted in Table I, our main contributions are three-fold:

- **Markovian formulation and high-order weak approximation:** We construct a surrogate pure-strategy SDG via moment matching. By incorporating state-dependent energy bounds, this formulation natively accommodates unbounded Markovian feedbacks. We prove that this surrogate SDG guarantees the weak approximation accuracy of state distributions and expected costs by $\mathcal{O}(\bar{\pi}^2)$.
- **Analytic solutions and dynamic energy allocation:** By analytically resolving the surrogate SDG, we bypass intractable HJI partial differential equations. Instead, we reveal that optimal moment trajectories are governed by a Generalized Riccati Differential Equation (GRDE). This uncovers a novel *dynamic energy allocation law* that explicitly dictates variance injection. Based on these closed-form moments, we propose a robust dual-routing architecture for practical strategy execution.
- **Explicit suboptimality certification:** By fully exploiting the linear-quadratic problem structure to design a refined error-decoupling analysis, we successfully replace the implicit implementation penalties of previous works with clean, explicit metrics. We certify that both the value approximation error and the strategy suboptimality gaps are bounded by $\mathcal{O}(\bar{\pi}^{1/2})$.

D. Related Works

The instantaneous-randomization issue in continuous-time mixed strategies was raised in the seminal work [22]. Subsequent research on mixed strategies in differential games has proceeded along two complementary lines, distinguished by how controls are modeled. The first line retains real-valued feedback controls and enforces commitment delays in the strategy definition to prevent pathological instantaneous randomization [24]–[29]. In these works, commitment delays are typically assumed to vanish asymptotically to recover the classical game value, which is characterized as the unique viscosity solution of an HJI equation. The second line builds on the relaxed-control framework [30], treating admissible controls directly as measure-valued functions and leveraging martingale or weak-formulation techniques to define equilibria [31]–[33]. In this paper, we adopt the real-valued control perspective with explicit, non-vanishing commitment delays, aligning with the first line of research to preserve direct physical implementability.

Despite the solid theoretical foundation provided by these studies, analytic solutions for optimal mixed strategies in ZSDGs remain critically limited. Prior efforts have focused primarily on proving the existence of, and characterizing, the value function under vanishing delays, rather than yielding explicit strategy laws [25]–[28]. First, solving the associated HJI equations in infinite-dimensional measure spaces is notoriously intractable: even pure-strategy solutions lack closed-form expressions except in very special cases [34], and classical methods such as the Pontryagin maximum principle [35] or viscosity-solution techniques [36] do not readily generalize. Second, only a handful of studies have tackled specific instances: Kumar's generalized bomber-and-battleship game yields an explicit mixed-strategy solution [37], and Fleming and Hernández established the existence of approximately optimal strategies [29]. While more recent works often resort to neural-network-based policy learning [38], [39], these approaches typically sacrifice analytical transparency. Third, forcing commitment delays toward zero induces chattering controls that switch infinitely fast, undermining both physical implementability and system robustness [40].

More recently, our prior works initiated an SDG weak approximation framework to circumvent the infinite-dimensional measure optimization. While an open-loop moment-matching surrogate was explored in [1], a systematic near-optimal design was later established in [24] utilizing control-space discretization. Although the latter provides rigorous $\mathcal{O}(\bar{\pi})$ performance guarantees for general nonlinear ZSDGs, its structural reliance on spatial discretization fundamentally restricts the output to numerical approximations. This methodological gap directly motivates our current work. By specializing the weak approximation framework to ZSLQDGs and parameterizing Markovian mixed strategies via their statistical moments, we completely bypass numerical spatial discretization. This paradigm shift enables us to derive exact analytic mixed strategies governed by GRDEs, elevate the weak approximation accuracy to $\mathcal{O}(\bar{\pi}^2)$, and rigorously certify the suboptimality gaps, ultimately yielding closed-form strategies that are directly implementable in digital systems.

Organization. The remainder of this paper is organized as follows. Section II formally formulates the ZSLQDG problem under mixed strategies. Section III introduces the SDG weak approximation framework and proves that the moment-matching surrogate game achieves a second-order approximation accuracy. In Section IV, we analytically resolve the surrogate game to synthesize the near-optimal mixed strategies and rigorously certify the bounds for both the value approximation error and the strategy suboptimality gaps. Section V validates our theoretical results and explores the physical behaviors of

TABLE I
COMPARISON OF GAME FORMULATIONS AND SOLUTION METHODS ACROSS DIFFERENT WORKS

	Game Formulation					Solution Methods					
						(Step 1)	(Step 2)	(Step 3)	(Step 4)	(Step 5)	
	Dynamics	Control Space	Cost Function	Information Structure	Strategy Constraints	Design $\tilde{G}$	Weak Approx.	Solve $\tilde{G}$	Strategy Synthesis	Value Approx.	Suboptimality Gap
This work	linear	unbounded	quadratic	Markovian feedback	state-dep. 2nd-moment bound	moment matching	$\mathcal{O}(\bar{\pi}^2)$	GRDE	ZOH	$\mathcal{O}(\bar{\pi}^{\frac{1}{2}})$	$\mathcal{O}(\bar{\pi}^{\frac{1}{2}})$
[24]	bounded	compact	bounded	NAD	×	control space discretization	$\mathcal{O}(\bar{\pi})$	HJI	ZOH	$\max\{\epsilon_1,\epsilon_2\}$ $+\mathcal{O}(\bar{\pi})$	$\max\{\epsilon_4,\epsilon_5\}$ $+\epsilon_3+\mathcal{O}(\bar{\pi})$
[1]	linear	unbounded	polynomial growth	open-loop	×	moment matching	$\mathcal{O}(\bar{\pi})$	HJI	×	×	×

mixed strategies via a 2-D double-integrator pursuit-evasion game. Finally, Section VI concludes the paper.

Notation. Let $\pi = \{t_0 < t_1 < \ldots < t_N = T\}$ with a finite positive integer N be a partition of a time interval $[t_0, T]$, and $\Pi_{[t_0,T]}$ be the set of all partitions on $[t_0, T]$. We define the lower bound, upper bound, and order of π as follows:

$$\underline{\pi} \triangleq \min_{0\le k\le N-1}(t_{k+1}-t_k),\ \bar{\pi} \triangleq \max_{0\le k\le N-1}(t_{k+1}-t_k),\ |\pi| \triangleq N.$$

A partition π is called equispaced if $\underline{\pi} = \bar{\pi}$. The k-th time interval, $[t_{k-1}, t_k)$ is denoted as Δ_k, with length $|\Delta_k| \triangleq t_k - t_{k-1}$. The indicator function $\mathbb{I}_{\Delta_k}(t)$ equals 1 if $t \in \Delta_k$, otherwise equals 0. $\zeta_\pi : [t_0, T] \to \{1, \ldots, |\pi|\}$ is the index function of π, and $\delta_\pi : [t_0, T] \to [\underline{\pi}, \bar{\pi}]$ is the interval length function of π, i.e.,

$$\zeta_\pi(t) \triangleq \sum_{k=1}^{|\pi|} k\,\mathbb{I}_{\Delta_k}(t), \quad \delta_\pi(t) \triangleq \sum_{k=1}^{|\pi|} |\Delta_k|\,\mathbb{I}_{\Delta_k}(t).$$

For a U-valued random variable x and a map $h : U \to \mathbb{R}^d$, the expectation and covariance of $h(x)$ are denoted as $\mathbb{E}[h(x)]$ and $\mathbb{D}[h(x)]$, respectively. Let $\mathcal{P}(U)$ denote the set of probability measures on U. We denote the set of positive semi-definite matrices of dimension n as S_+^n. Given $A, B \in \mathbb{R}^{n\times n}$, we use $A \succeq B$ to indicate $A - B \in S_+^n$.

II. Problem Formulation

In this section, we first present the system dynamics and the associated quadratic cost functional. Then, we define the admissible mixed strategies as conditionally independent Markovian randomized feedbacks and characterize the state-dependent energy constraints. Finally, we formalize the game value and outline the main problems of interest.

A. Dynamics and Cost

Consider a continuous-time ZSLQDG defined on the time horizon $[t_0, T]$. In contrast to the classical pure-strategy paradigm where control inputs are deterministic functions, we consider a broader class of games where players may randomize their actions. Consequently, the control inputs $u(t)$ and $v(t)$ are generally modeled as stochastic processes. The dynamics and cost of the game are as follows:

$$(\mathbf{G}_{lq}):\begin{cases}\dot{x}(t) = Ax(t)+B_1u(t)+B_2v(t), \quad x(t_0)=x_0,\\ J(t_0,x_0,u,v) = \mathbb{E}\Big\{\int_{t_0}^{T}\big[x^\top(t)Qx(t)+u^\top(t)R_1u(t)\\ \qquad\qquad -v^\top(t)R_2v(t)\big]dt + x^\top(T)Q_Tx(T)\Big\},\end{cases}$$

where $x(t) \in \mathbb{R}^d$ is the state vector. Player 1 (the minimizer) and Player 2 (the maximizer) take their control inputs in unbounded continuous spaces $U = \mathbb{R}^{d_1}$ and $V = \mathbb{R}^{d_2}$, respectively. The expectation operator $\mathbb{E}[\cdot]$ encapsulates the stochasticity intentionally injected via the players' randomized mixed strategies, whose rigorous definition will be formalized in Section II-B.

Assumption 1 (System matrices)**.** *The weighting matrices Q, Q_T, and R_i for $i = 1, 2$, are positive definite. Additionally, the pair (A, B_1) is stabilizable, and the pair $(A, Q^{\frac{1}{2}})$ is detectable.*

For an individual realized trajectory, we define the accumulated cost function evaluated at time $t \in [t_0, T]$ as

$$\text{cost}(t) = \begin{cases}\int_{t_0}^{t} h(s,x,u,v)ds, & t < T,\\ \int_{t_0}^{t} h(s,x,u,v)ds + g(x(T)), & t = T,\end{cases} \tag{1}$$

where $h(s,x,u,v) = x^\top(s)Qx(s) + u^\top(s)R_1u(s) - v^\top(s)R_2v(s)$ and $g(x(T)) = x^\top(T)Q_Tx(T)$. By definition, $\mathbb{E}[\text{cost}(T)] = J(t_0, x_0, u, v)$.

B. Admissible Mixed Strategy

In realistic digital control systems, it is standard for players to update their actions at discrete sampling instants due to communication latencies and computational constraints. To align with this practical mechanism, we introduce the commitment delay pattern into the continuous-time game.

Assumption 2 (Information structure and commitment delay)**.** *The control processes are governed by a commitment delay pattern $\pi = \{t_0 < t_1 < \cdots < t_N = T\}$. During each sampling interval $[t_k, t_{k+1})$ for $k \in \{0, \ldots, N-1\}$, both*

players must commit to a constant control action determined at the sampling instant t_k based on the available state $x(t_k)$, i.e., $u(t) = u_k$ and $v(t) = v_k$ for $t \in [t_k, t_{k+1})$.

While this assumption on shared commitment delay pattern is standard in the literature of sampled-data differential games [26], [28], we acknowledge that in practical decentralized scenarios, players may operate on different temporal grids. Analyzing such games with asynchronous delays, though beyond the scope of this paper, remains an important direction for future research.

Furthermore, since the control spaces U and V are unbounded, unrestricted randomization could yield infinite expected costs and destroy the well-posedness of the LQ framework. Therefore, it is necessary to physically bound the randomization. In linear feedback systems, the nominal control effort naturally scales linearly with the state deviation [41]. Concurrently, in networked control environments under Signal-to-Noise Ratio (SNR) constraints, the variance of injected control noise is fundamentally proportional to the signal power (i.e., state magnitude) [42], [43]. Motivated by these physical rationales, we impose state-dependent second-moment constraints on the probability measures, ensuring that the total injected stochastic energy dynamically scales with the system's current deviation.

Definition 1 (Admissible mixed strategy)**.** *An admissible mixed strategy for Player 1, denoted as $\alpha \in \mathcal{A}_m^\pi$, is a sequence of measurable Markovian maps $\{\mu_k\}_{k=0}^{N-1}$ such that:*

$$\mu_k : \mathbb{R}^d \to \mathcal{P}(U), \quad x(t_k) \mapsto \mu_k(\cdot \mid x(t_k)). \tag{2}$$

At each sampling instant t_k, Player 1 samples an action $u_k \sim \mu_k(\cdot \mid x(t_k))$. Additionally, the probability measures are subject to the state-dependent second-moment constraint:

$$\int_U \|u\|^2 \mu_k(du \mid x(t_k)) \le \gamma_1^2 \|x(t_k)\|^2, \tag{3}$$

where $\gamma_1 > 0$ represents the actuation capacity coefficient.

Symmetrically, the admissible strategy space $\mathcal{B}_m^\pi$ for Player 2 is defined as a sequence of measurable Markovian maps $\{\nu_k\}_{k=0}^{N-1}$ subject to the actuation capacity $\gamma_2 > 0$:

$$\int_V \|v\|^2 \nu_k(dv \mid x(t_k)) \le \gamma_2^2 \|x(t_k)\|^2. \tag{4}$$

To preserve fairness and non-anticipativity in this simultaneous-move game, the intrinsic randomization mechanisms (i.e., random seeds) employed by the players must remain strictly private and causally decoupled. This prevents any player from anticipatively observing or exploiting the opponent's action within the same commitment interval.

Assumption 3 (Conditional independence)**.** *Given the commonly observed state $x(t_k)$, the random controls u_k and v_k are sampled independently, i.e., the joint probability measure factorizes as $\mathbb{P}(u_k, v_k \mid x(t_k)) = \mu_k(u_k \mid x(t_k))\nu_k(v_k \mid x(t_k))$.*

Remark 1 (Markovian feedback vs. NAD formulation)**.** *For general ZSDGs with bounded control spaces, existing literature [24], [26], [28] traditionally defines mixed strategies through the framework of non-anticipative strategies with delay (NAD). A core motivation for adopting the NAD functional mapping is to rigorously characterize the game value as the unique viscosity solution to the associated HJI equation. In contrast, for the ZSLQDG addressed in this paper, our weak approximation framework allows the surrogate game to be solved analytically via a GRDE. Since the GRDE admits explicit classical solutions, the NAD framework and viscosity solutions are no longer necessitated, which directly yields an easily implementable Markovian control architecture.*

C. Problems of Interest

The sequential execution of any admissible mixed strategy pair $(\alpha, \beta) \in \mathcal{A}_m^\pi \times \mathcal{B}_m^\pi$, coupled with the initial state x_0, uniquely induces a joint probability measure over the sample paths of the control processes u and v. Consequently, the expected cost under a specific strategy pair is unambiguously defined as $J(t_0, x_0, \alpha, \beta)$, which evaluates the cost functional under this induced measure.

To investigate the equilibrium, we define the upper and lower value functions as:

$$\begin{aligned} V_{m,+}^\pi(t_0, x_0) &\triangleq \inf_{\alpha \in \mathcal{A}_m^\pi} \sup_{\beta \in \mathcal{B}_m^\pi} J(t_0, x_0, \alpha, \beta), \\ V_{m,-}^\pi(t_0, x_0) &\triangleq \sup_{\beta \in \mathcal{B}_m^\pi} \inf_{\alpha \in \mathcal{A}_m^\pi} J(t_0, x_0, \alpha, \beta). \end{aligned} \tag{5}$$

We are interested in the following three core problems:

- **Existence of game value.** Establish whether the saddle-point equilibrium exists for $\mathbf{G}_{lq}$ under the defined spaces $\mathcal{A}_m^\pi \times \mathcal{B}_m^\pi$, i.e., $V_{m,+}^\pi(t_0, x_0) = V_{m,-}^\pi(t_0, x_0)$.
- **State and cost characterization.** Given any pair of mixed strategies, rigorously characterize the continuous-time evolution of the system state distribution and the accumulation of expected costs over the horizon.
- **Near-optimal strategy and certification.** Derive analytic mixed strategies by solving a surrogate pure-strategy SDG via GRDE, and certify both the value approximation error and the strategy suboptimality gap.

To address these problems, our main idea is to weakly approximate the original mixed-strategy game $\mathbf{G}_{lq}$ by a pure-strategy SDG $\tilde{\mathbf{G}}_{lq}$, then derive both approximated game value and near-optimal mixed strategy with performance certifications. The key points are illustrated in Fig. 1.

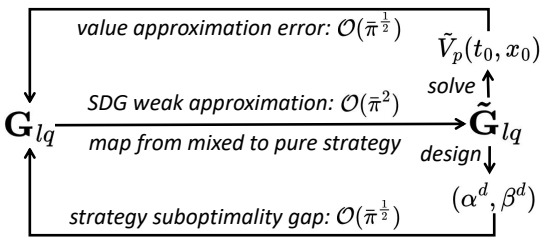


Fig. 1. Illustration of the main ideas.

III. Game Value Existence and SDG Weak Approximation

In this section, we address the first two problems of interest outlined in Section II. First, we establish the mathematical well-posedness of the game by proving the existence of a unique saddle-point game value. Second, to analytically characterize the evolution of the system state and expected costs driven by the players' discrete-time randomizations, we introduce a weak approximation framework. By mapping the random interventions to continuous-time diffusion processes, we construct a mathematically tractable surrogate SDG.

A. Existence of Game Value

Before constructing the surrogate approximation, it is imperative to ensure that the original game $\mathbf{G}_{lq}$ is well-posed under the defined mixed strategy spaces. Benefiting from the ZOH commitment mechanism and the weakly compact probability measure spaces induced by the state-dependent energy constraints, the existence of the game value is guaranteed.

Theorem 1 (Existence of game value)**.** *The linear-quadratic game $\mathbf{G}_{lq}$ under the admissible Markovian mixed strategy spaces $\mathcal{A}_m^\pi \times \mathcal{B}_m^\pi$ has a unique game value, denoted as $V_m^\pi(t_0, x_0)$.*

Proof. We establish the existence of the game value via backward induction, starting from the terminal time T. For each interval $[t_{k-1}, t_k)$, we demonstrate the existence of a local saddle point by invoking Sion's Minimax Theorem. We rigorously verify its requisite conditions by showing that: (i) the energy-constrained mixed strategy spaces are weakly compact, as ensured by the uniform second-moment bounds and Prokhorov's Theorem; and (ii) the local expected cost functional is bilinear and continuous in the weak topology. The recursive coincidence of the local upper and lower values yields the global game value. Full details are provided in Appendix A. □

B. Weak Approximation Framework

Characterizing the continuous-time evolution of the state distribution and expected cost under mixed strategies (our second problem of interest) is analytically intractable. The piecewise-constant random interventions inherently create a complex stochastic process whose exact pathwise behavior cannot be evaluated in closed form. However, since the cost functional of $\mathbf{G}_{lq}$ relies solely on the *expectations* of these trajectories, exact pathwise matching (i.e., strong approximation) is unnecessary. Instead, it is sufficient to capture the macroscopic probability distributions over time.

Let PG denote the set of continuous functions $\mathbb{R}^d \to \mathbb{R}$ of at most polynomial growth, i.e., $\psi(\cdot) \in \mathrm{PG}$ if there exist $\kappa_1, \kappa_2 > 0$ such that $|\psi(x)| \le \kappa_1 (1 + |x|^{\kappa_2})$ for all $x \in \mathbb{R}^d$. If $\kappa_2 = 1$, ψ is of linear growth. Moreover, PG^n denotes the set of n-times continuously differentiable functions whose partial derivatives up to order n belong to PG. Given a partition π on $[t_0, T]$, a continuous-time process $\{\tilde{x}(t)\}_{t\in[t_0,T]}$ and a discrete-time process $\{x(k)\}_{k=0}^{|\pi|}$ weakly approximate each other if their probability distributions at the discrete time steps t_k are uniformly close.

Definition 2 (Weak approximation)**.** *Given a partition π on $[t_0, T]$, a continuous-time stochastic process $\{\tilde{x}(t)\}_{t\in[t_0,T]}$ is an order n weak approximation of a discrete-time stochastic process $\{x(k)\}_{k=0}^{|\pi|}$ if, for any polynomial-growth function $\psi \in \mathrm{PG}^{n+1}$, there is a positive constant C independent of $\bar{\pi}$ such that:*

$$\max_{k=0,\dots,|\pi|} |\mathbb{E}\psi(x(k)) - \mathbb{E}\psi(\tilde{x}(t_k))| \le C\bar{\pi}^n.$$

This concept is generalized to the game setting in [24]. A surrogate SDG under pure strategies weakly approximates the original ZSDG under mixed strategies if both the state distributions and the expected accumulated costs are well approximated. The approximation error bounds defined in terms of $\mathcal{O}(\bar{\pi}^n)$ are asymptotic properties valid in the limit $\bar{\pi} \to 0$. In the context of differential games, this corresponds to high-frequency sampling where the commitment delay is sufficiently small. Within this regime ($\bar{\pi} < 1$), a higher order n provides a strictly superior approximation accuracy.

Definition 3 (SDG weak approximation [24])**.** *A zero-sum SDG $(\tilde{\mathbf{G}})$ under pure strategy spaces $\tilde{\mathcal{A}}_p^\pi \times \tilde{\mathcal{B}}_p^\pi$ is an order n weak approximation of a zero-sum differential game $(\mathbf{G})$ under mixed strategy spaces $\mathcal{A}_m^\pi \times \mathcal{B}_m^\pi$ if:*

(i) there exist

$$\phi_1 : \mathcal{A}_m^\pi \to \tilde{\mathcal{A}}_p^\pi \text{ and } \phi_2 : \mathcal{B}_m^\pi \to \tilde{\mathcal{B}}_p^\pi, \tag{6}$$

where for any $(\alpha, \beta) \in \mathcal{A}_m^\pi \times \mathcal{B}_m^\pi$, one has $\tilde{\alpha} = \phi_1(\alpha) \in \tilde{\mathcal{A}}_p^\pi$ and $\tilde{\beta} = \phi_2(\beta) \in \tilde{\mathcal{B}}_p^\pi$.

(ii) the state trajectory of the SDG $\tilde{\mathbf{G}}$ under strategies $(\tilde{\alpha}, \tilde{\beta})$, i.e., $\{\tilde{x}(t)\}_{t\in[t_0,T]}$, is an order n weak approximation of the state trajectory of the game $\mathbf{G}$ under strategies (α, β) on the time steps of π, i.e., $\{x(t_k)\}_{k=0}^{|\pi|}$.

(iii) the accumulated cost trajectory of the SDG $\tilde{\mathbf{G}}$ under strategies $(\tilde{\alpha}, \tilde{\beta})$ approximates the accumulated cost trajectory of the game $\mathbf{G}$ under strategies (α, β) in an expected sense that $\left|\mathbb{E}\left[\mathrm{cost}(t_k)\right] - \mathbb{E}\left[\tilde{\mathrm{cost}}(t_k)\right]\right| = \mathcal{O}(\bar{\pi}^n)$ for $k = 0, \dots, |\pi|$.

Unlike classical weak approximation for SDEs, which typically concerns the distribution of the state process x_t, our definition enforces the joint weak convergence of the state process and the accumulated cost process (x_t, cost_t). This joint convergence is essential for differential games because the value function is defined via the cost functional. Convergence of the state alone is insufficient to guarantee the convergence of the game value. By matching both the state and the cost process, our framework ensures that the saddle point of the surrogate SDG provides a mathematically rigorous approximation of the original game value.

C. Design of the Surrogate SDG $\tilde{\mathbf{G}}_{lq}$

The intuition behind our moment-matching design is that, instead of choosing a distribution μ_k, Player 1 in the surrogate

game directly controls the conditional expectation and conditional covariance of their actions, effectively shaping the drift and diffusion of the system.

Formally, let $\{W_t\}_{t\in[t_0,T]}$ be a standard (d_1+d_2)-dimensional Wiener process defined on a probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P})$ with filtration $\tilde{\mathbb{F}}$. We define the virtual pure controls for the surrogate game as the tuple processes $\tilde{u}(t)=(\tilde{\Gamma}_1(t),\tilde{\Lambda}_1(t))$ and $\tilde{v}(t)=(\tilde{\Gamma}_2(t),\tilde{\Lambda}_2(t))$, where $\tilde{\Gamma}_i(t)$ takes values in $\mathbb{R}^{d_i}$ and $\tilde{\Lambda}_i(t)$ takes values in $\mathbb{R}^{d_i\times d_i}$. The covariance matrix is defined as $\tilde{\Sigma}_i(t)\triangleq\tilde{\Lambda}_i(t)\tilde{\Lambda}_i^\top(t)\succeq\mathbf{0}$.

Driven by these moment-based controls, the continuous-time surrogate ZSLQDG is constructed as follows:

$$(\tilde{\mathbf{G}}_{lq}):\begin{cases} d\tilde{x}(t)=\Big(A\tilde{x}(t)+B_1\tilde{\Gamma}_1(t)+B_2\tilde{\Gamma}_2(t)\Big)\,dt \\ \qquad\quad +\delta_{\bar{\pi}}^{\frac{1}{2}}(t)\Big[B_1\tilde{\Lambda}_1(t),B_2\tilde{\Lambda}_2(t)\Big]\,dW_t,\ \tilde{x}(t_0)=x_0, \\ \tilde{J}(t_0,x_0,\tilde{u},\tilde{v})=\mathbb{E}\displaystyle\int_{t_0}^{T}\Big[\tilde{x}^\top(t)Q\tilde{x}(t)+\tilde{\Gamma}_1^\top(t)R_1\tilde{\Gamma}_1(t) \\ \qquad -\tilde{\Gamma}_2^\top(t)R_2\tilde{\Gamma}_2(t)+\mathrm{tr}(\tilde{\Lambda}_1(t)^\top R_1\tilde{\Lambda}_1(t)) \\ \qquad -\mathrm{tr}(\tilde{\Lambda}_2(t)^\top R_2\tilde{\Lambda}_2(t))\Big]\,dt+\tilde{x}^\top(T)Q_T\tilde{x}(T). \end{cases}$$

A fundamental advantage of this formulation emerges when translating the state-dependent energy constraints. Using the well-known variance identity $\mathbb{E}[\|u\|^2]=\|\mathbb{E}[u]\|^2+\mathrm{tr}(\mathbb{D}[u])$, the original integral constraints over probability measures gracefully translate into deterministic algebraic boundaries. This enables us to directly define the admissible strategies for the surrogate game symmetrically to the original game.

Definition 4 (Admissible pure strategy)**.** *An admissible pure strategy for Player 1 in the surrogate game, denoted as $\tilde{\alpha}\in\tilde{\mathcal{A}}_p^\pi$, is a sequence of measurable Markovian maps $\{\tilde{\mu}_k\}_{k=0}^{N-1}$:*

$$\tilde{\mu}_k:\mathbb{R}^d\to\mathbb{R}^{d_1}\times\mathbb{R}^{d_1\times d_1},\quad \tilde{x}(t_k)\mapsto(\tilde{\Gamma}_1(t_k),\tilde{\Lambda}_1(t_k)).\quad(7)$$

At each sampling instant t_k, the extracted moments are held constant for $t\in[t_k,t_{k+1})$ and must satisfy the state-dependent algebraic boundary:

$$\|\tilde{\Gamma}_1(t_k)\|^2+\mathrm{tr}(\tilde{\Sigma}_1(t_k))\le\gamma_1^2\|\tilde{x}(t_k)\|^2,\qquad(8)$$

where $\tilde{\Sigma}_1(t_k)\triangleq\tilde{\Lambda}_1(t_k)\tilde{\Lambda}_1^\top(t_k)\succeq\mathbf{0}$.

Symmetrically, the admissible pure strategy space $\tilde{\mathcal{B}}_p^\pi$ for Player 2 is defined as a sequence of maps $\{\tilde{\nu}_k\}_{k=0}^{N-1}$ mapping $\tilde{x}(t_k)$ to $(\tilde{\Gamma}_2(t_k),\tilde{\Lambda}_2(t_k))$, subject to $\|\tilde{\Gamma}_2(t_k)\|^2+\mathrm{tr}(\tilde{\Sigma}_2(t_k))\le\gamma_2^2\|\tilde{x}(t_k)\|^2$.

By replacing the infinite-dimensional measure spaces with Euclidean moment matrices, the original game is successfully relaxed into a highly tractable pure-strategy SDG that strictly preserves the LQ structure. Let $\tilde{V}_{p,+}^\pi$ and $\tilde{V}_{p,-}^\pi$ denote its upper and lower value functions, respectively.

D. Strategy Projection and Certification of Order 2

To mathematically validate that $\tilde{\mathbf{G}}_{lq}$ satisfies the weak approximation requirements (Definition 3), we must construct explicit strategy projection maps (ϕ_1,ϕ_2) that connect the two distinct probability spaces.

Definition 5 (Strategy projection maps)**.** *For Player 1 employing an admissible Markovian mixed strategy $\alpha=\{\mu_k\}_{k=0}^{N-1}\in\mathcal{A}_m^\pi$, the projection map ϕ_1 constructs a surrogate pure strategy $\tilde{\alpha}=\{\tilde{\mu}_k\}_{k=0}^{N-1}\in\tilde{\mathcal{A}}_p^\pi$ by extracting the conditional statistical moments. Specifically, the surrogate controls $(\tilde{\Gamma}_1(t_k),\tilde{\Lambda}_1(t_k))$ evaluated at the mapped state $\tilde{x}(t_k)$ are defined by:*

$$\tilde{\Gamma}_1(t_k)=\int_U u\,\mu_k(du\mid\tilde{x}(t_k)),$$
$$\tilde{\Sigma}_1(t_k)=\int_U\big(u-\tilde{\Gamma}_1(t_k)\big)\big(u-\tilde{\Gamma}_1(t_k)\big)^\top\mu_k(du\mid\tilde{x}(t_k)),$$

where $\tilde{\Sigma}_1(t_k)\triangleq\tilde{\Lambda}_1(t_k)\tilde{\Lambda}_1^\top(t_k)\succeq\mathbf{0}$. The map ϕ_2 for Player 2 is symmetrically defined via ν_k.

Remark 2 (Feasibility)**.** *By the variance identity $\int\|u\|^2d\mu=\|\mathbb{E}[u]\|^2+\mathrm{tr}(\mathbb{D}[u])$, the energy constraint of $\mathbf{G}_{lq}$ intrinsically ensures $\|\tilde{\Gamma}_i(t_k)\|^2+\mathrm{tr}(\tilde{\Sigma}_i(t_k))\le\gamma_i^2\|\tilde{x}(t_k)\|^2$. Thus, the maps (ϕ_1,ϕ_2) are strictly feasible for $\tilde{\mathbf{G}}_{lq}$.*

With the projection maps rigorously defined, we establish the approximation performance. By matching both the first and second moments, the surrogate game accurately traces the original game dynamics.

Theorem 2 ($\tilde{\mathbf{G}}_{lq}$ approximation)**.** *Under the designed projection maps (ϕ_1,ϕ_2), the surrogate pure-strategy SDG $\tilde{\mathbf{G}}_{lq}$ is an order 2 weak approximation (i.e., $\mathcal{O}(\bar{\pi}^2)$) of the original mixed-strategy game $\mathbf{G}_{lq}$.*

Proof. We certify the $\mathcal{O}(\bar{\pi}^2)$ weak approximation using Milshtein's framework. First, we augment the state dynamics with the accumulated cost variable to enable joint state-cost approximation. Second, we perform a local one-step error analysis between the exact game increment and the surrogate SDG, proving that their first three moments deviate by at most $\mathcal{O}(\bar{\pi}^3)$. Finally, by aggregating these local deviations over N sampling intervals, we rigorously establish that the global approximation error is bounded by $\mathcal{O}(\bar{\pi}^2)$. Full details are provided in Appendix B. □

In our prior works on general nonlinear ZSDGs [24], the weak approximation error was strictly bounded by $\mathcal{O}(\bar{\pi})$ (order 1). Remarkably, owing to the global linearity of the dynamics and the uncentered moment substitutions in the ZSLQDG framework, the local truncation errors of the third-order moments vanish structurally. This elevates the global weak approximation accuracy to $\mathcal{O}(\bar{\pi}^2)$ (order 2), providing a significantly tighter characterization of both state distributions and expected costs.

IV. Near-Optimal Mixed Strategies

In this section, we analytically resolve the surrogate SDG $\tilde{\mathbf{G}}_{lq}$ to extract closed-form optimal pure strategies. By inversely mapping these continuous-time deterministic moments back to the discrete-time probability spaces, we synthesize near-optimal mixed strategies for the original game $\mathbf{G}_{lq}$. Finally, we develop a rigorous performance certification result, systematically bounding the global value approximation error and the strategy suboptimality gaps.

A. Solve $\tilde{\mathbf{G}}_{lq}$

Solving the surrogate game $\tilde{\mathbf{G}}_{lq}$ directly under the delay-dependent strategy spaces $\tilde{\mathcal{A}}_p^{\pi} \times \tilde{\mathcal{B}}_p^{\pi}$ is analytically challenging. To derive closed-form optimal control laws, we analyze the game under the asymptotic continuous-time regime where the commitment delays approach zero, yielding the delay-free admissible spaces $\tilde{\mathcal{A}}_p \triangleq \cup_{\pi} \tilde{\mathcal{A}}_p^{\pi}$ and $\tilde{\mathcal{B}}_p \triangleq \cup_{\pi} \tilde{\mathcal{B}}_p^{\pi}$. Since the surrogate linear-quadratic SDG rigorously satisfies the Isaacs condition, its saddle-point game value, denoted as $\tilde{V}_p(t_0, x_0)$, is well-defined over $\tilde{\mathcal{A}}_p \times \tilde{\mathcal{B}}_p$. This condition holds naturally because the Hamiltonian is separable in $\tilde{u}$ and $\tilde{v}$, implying that the minimax equality holds [44].

To analytically resolve the surrogate game, we first derive the optimal moments under the continuous-time delay-free spaces $\tilde{\mathcal{A}}_p \times \tilde{\mathcal{B}}_p$. These continuous-time optimal mappings will later be strictly projected or truncated into the piecewise-constant ZOH execution mechanism during practical implementation (Algorithm 1) and ZOH strategy replacement errors (Lemma 1).

Next, we introduce a symmetric matrix process $P(t)$ uniquely satisfying the following backward Generalized Riccati Differential Equation (GRDE):

$$\begin{cases} -\dot{P}(t) = & A^{\top}P(t) + P(t)A + Q - P(t)B_1 R_1^{-1} B_1^{\top} P(t) \\ & + P(t)B_2 \left(R_2 + \lambda_2(t) I\right)^{-1} B_2^{\top} P(t) + \lambda_2(t)\gamma_2^2 I, \\ P(T) = & Q_T. \end{cases} \tag{9}$$

In this formulation, $\lambda_2(t)$ acts as a dynamic penalty factor tracking the local system divergence, defined strictly as:

$$\lambda_2(t) \triangleq \max\left(0, \lambda_{\max}(Q_{2,eq}(t))\right). \tag{10}$$

Here, the equivalent weight matrices $Q_{i,eq}(t)$ for $i \in \{1, 2\}$ are elegantly constructed as:

$$Q_{i,eq}(t) \triangleq \delta_{\pi}(t) B_i^{\top} P(t) B_i - (-1)^i R_i. \tag{11}$$

These equivalent matrices bear a fundamental control-theoretic interpretation that will be mathematically justified in the subsequent optimal synthesis: they explicitly combine the baseline steady-state control costs (R_i) with the dynamic value-to-go penalty ($B_i^{\top} P(t) B_i$) induced by the stochastic diffusion propagating over the delay duration $\delta_{\pi}(t)$.

With $P(t)$ and the equivalent matrices formally defined, we restrict the game to a well-posed parameter regime where the linear-quadratic properties do not collapse due to extreme hardware limitations.

Assumption 4 (Sufficient actuation capacity)**.** *We assume the physical energy capacities γ_1 and γ_2 are sufficiently large to support the unconstrained optimal feedback gains, satisfying the following matrix inequalities for all $t \in [t_0, T]$:*

$$\begin{cases} \gamma_1^2 I - P(t) B_1 R_1^{-2} B_1^{\top} P(t) \succeq \mathbf{0}, \\ \gamma_2^2 I - P(t) B_2 (R_2 + \lambda_2(t) I)^{-2} B_2^{\top} P(t) \succeq \mathbf{0}. \end{cases} \tag{12}$$

Condition (12) ensures that the actuators possess adequate power margins relative to the system's dynamic requirements. If violated, the optimal mean controls would inevitably suffer from strict boundary saturation (e.g., $\|\tilde{\Gamma}_i\|^2 = \gamma_i^2 \|x\|^2$). Such persistent saturation destroys the global quadratic structure, degenerating the LQ game into a highly nonlinear bounded-control problem. Assumption 4 explicitly bounds the parameter regime where the analytical properties of the LQ paradigm are preserved.

Theorem 3 (Game value and optimal surrogate strategies)**.** *Let $v_2(t)$ be the unit principal eigenvector corresponding to $\lambda_{\max}(Q_{2,eq}(t))$. Under Assumption 4, the game value and the optimal pure strategies of $\tilde{\mathbf{G}}_{lq}$ over $\tilde{\mathcal{A}}_p \times \tilde{\mathcal{B}}_p$ admit exact analytical solutions:*

(i) The saddle-point game value is strictly a global quadratic form:

$$\tilde{V}_p(t, x) = x^{\top} P(t) x, \quad \forall t \in [t_0, T].$$

(ii) For Player 1 (Minimizer), the optimal strategy is:

$$\tilde{\Gamma}_1^*(t) = -R_1^{-1} B_1^{\top} P(t) \tilde{x}(t), \quad \tilde{\Sigma}_1^*(t) = \mathbf{0}. \tag{13}$$

(iii) For Player 2 (Maximizer), let $K_2(t) \triangleq (R_2 + \lambda_2(t) I)^{-1} B_2^{\top} P(t)$. The optimal strategy is:

$$\begin{cases} \tilde{\Gamma}_2^*(t) = K_2(t)\tilde{x}(t), \\ \tilde{\Sigma}_2^*(t) = \tilde{x}(t)^{\top} \left(\gamma_2^2 I - K_2(t)^{\top} K_2(t)\right) \tilde{x}(t) \cdot v_2(t) v_2(t)^{\top} \\ \qquad \cdot \mathbb{I}_{\{\lambda_{\max}(Q_{2,eq}(t)) > 0\}}. \end{cases} \tag{14}$$

Proof. We solve the game by analyzing the HJI equation of the surrogate SDG. By postulating a quadratic value function $\tilde{V}(t, x) = x^{\top} P(t) x$, the HJI equation decouples into two independent subproblems. We resolve Player 2's maximizer subproblem using Lagrangian duality, establishing that the optimal mean is a linear feedback and the variance must align with the principal eigenvector of the variance weight matrix $Q_{2,eq}$. Player 1's minimizer subproblem is solved via standard minimization. Substituting the optimal controls back into the HJI equation yields the GRDE (9), characterizing the evolution of $P(t)$. Full details are provided in Appendix C. □

Remark 3 (Violation of actuation capacity)**.** *Theorem 3 relies strictly on Assumption 4 to guarantee that the optimal value function retains a globally explicit quadratic structure. In practical applications where the physical energy budget γ_2 is severely restricted and fails to satisfy condition* (12)*, the residual matrix $\gamma_2^2 I - K_2(t)^{\top} K_2(t)$ may lose its positive semi-definiteness, structurally invalidating the unconstrained Riccati derivation. Under this specific scenario, one must resort to numerically resolving the underlying HJI partial differential equation over state grids, as the true value function inevitably transitions from a global quadratic form into a saturated, highly nonlinear structure.*

Theorem 3 reveals a profound *dynamic energy allocation law*: Player 2 prioritizes their energy budget strictly for the deterministic linear feedback mean $\tilde{\Gamma}_2^*$. Only when the local dynamics approach divergence ($\lambda_{\max}(Q_{2,eq}) > 0$) does Player 2 inject the residual energy budget into the variance along the system's most vulnerable eigendirection $v_2(t)$. Furthermore, as the commitment delay $\bar{\pi} \to 0$, we naturally have $\lambda_2(t) \to 0$. Consequently, the optimal variance vanishes ($\tilde{\Sigma}_2^*(t) = \mathbf{0}$), and the GRDE (9) degenerates to the standard continuous-time Riccati equation, fully recovering classical LQ theory.

Algorithm 1 Near-Optimal Mixed Strategies for $\mathbf{G}_{lq}$

Input: Dynamics A, B_1, B_2; costs Q, R_1, R_2, Q_T; delay π; capacities γ_1, γ_2; optional spatial grids $\mathcal{X}$.

▷ **Phase 1: Offline verification & routing**
Solve GRDE (9) for $P(t)$, $t \in [t_0, T]$.
if Assumption 4 holds via (12) **then**
  Mode $\leftarrow$ Analytical (LQ)
else
  Mode $\leftarrow$ Numerical (HJI Fallback)
  Solve HJI PDE over grids $\mathcal{X}$ for true value $V(t, x)$.
end if
Initialize $x(t_0) = x_0$ and $\text{cost}(t_0) \leftarrow 0$.
for $k = 0, \ldots, N-1$ **do**
  ▷ **Phase 2: Online moment evaluation**
  if Mode == Analytical (LQ) **then**
    Compute $(\Gamma^*_{1,k}, \Sigma^*_{1,k})$ and $(\Gamma^*_{2,k}, \Sigma^*_{2,k})$ via explicit solutions (13) and (14).
  else
    $V_x \leftarrow \nabla_x V(t_k, x(t_k))$, $V_{xx} \leftarrow \nabla^2_{xx} V(t_k, x(t_k))$.
    $M_{1,k} \leftarrow R_1 + \frac{1}{2}\delta_\pi(t_k) B_1^\top V_{xx} B_1$.
    $M_{2,k} \leftarrow \frac{1}{2}\delta_\pi(t_k) B_2^\top V_{xx} B_2 - R_2$.
    Solve Minimizer's optimization:
    $\min_{\Gamma_1, \Sigma_1 \succeq \mathbf{0}} \{\Gamma_1^\top R_1 \Gamma_1 + V_x^\top B_1 \Gamma_1 + \text{tr}(M_{1,k}\Sigma_1)\}$
      s.t. $\|\Gamma_1\|^2 + \text{tr}(\Sigma_1) \le \gamma_1^2 \|x(t_k)\|^2$.
    Solve Maximizer's optimization:
    $\max_{\Gamma_2, \Sigma_2 \succeq \mathbf{0}} \{-\Gamma_2^\top R_2 \Gamma_2 + V_x^\top B_2 \Gamma_2 + \text{tr}(M_{2,k}\Sigma_2)\}$
      s.t. $\|\Gamma_2\|^2 + \text{tr}(\Sigma_2) \le \gamma_2^2 \|x(t_k)\|^2$.
    Let $(\Gamma^*_{i,k}, \Sigma^*_{i,k})$ be the resulting optimal moments.
  end if
  ▷ **Phase 3: ZOH inter-sample execution**
  Set prob. measures: $\mu_k = \mathcal{N}(\Gamma^*_{1,k}, \Sigma^*_{1,k})$, $\nu_k = \mathcal{N}(\Gamma^*_{2,k}, \Sigma^*_{2,k})$.
  Sample independently: $u_k \sim \mu_k$, $v_k \sim \nu_k$.
  for $t \in [t_k, t_{k+1})$ **do**
    Evolve dynamics: $\dot{x}(t) = Ax(t) + B_1 u_k + B_2 v_k$.
  end for
  $J_{step} \leftarrow \int_{t_k}^{t_{k+1}} h(s, x(s), u_k, v_k) ds$.
  $\text{cost}(t_{k+1}) \leftarrow \text{cost}(t_k) + J_{step}$.
end for
Add terminal cost: $\text{cost}(T) \leftarrow \text{cost}(T) + x(T)^\top Q_T x(T)$.

Output: Trajectories $x(t), u(t), v(t)$, and accumulated $\text{cost}(T)$.

B. Synthesis of Near-Optimal Strategies

Theorem 3 provides explicit analytical solutions for the surrogate game $\tilde{\mathbf{G}}_{lq}$. A major computational advantage of this LQ framework is that the synthesis process relies entirely on solving the GRDE (9) offline, fundamentally circumventing the curse of dimensionality inherent in general PDEs.

To seamlessly integrate the analytical GRDE solutions with the strict physical capacity limits addressed in Remark 3, we propose a robust *dual-routing execution architecture*. As detailed in Algorithm 1, this architecture bridges continuous-time theory with discrete-time digital control practice through three distinct operational phases.

Phase 1: Offline verification and routing. Before online deployment, the system evaluates the mathematical well-posedness of the unconstrained LQ framework. By pre-computing the GRDE (9) over $[t_0, T]$, we rigorously verify whether the physical actuation capacities (γ_1, γ_2) are sufficient to support the unconstrained optimal feedback gains (Assumption 4). If the condition holds, the system is routed to the highly efficient analytical mode. Conversely, if the capacities are severely restricted, the optimal moments will inevitably suffer from boundary saturations, destroying the global quadratic structure. In this scenario, the system triggers the numerical fallback mode, resorting to a grid-based HJI PDE solver to pre-compute the true non-quadratic value function over spatial grids $\mathcal{X}$.

Phase 2: Online moment evaluation. During real-time operation at each sampling instant t_k, the players compute their optimal statistical moments based on the observed state $x(t_k)$. In the analytical mode, this requires merely trivial $\mathcal{O}(1)$ matrix-vector multiplications via the explicit closed-form solutions (13) and (14), rendering it extremely computationally efficient. In the numerical fallback mode, the players extract the local gradient and Hessian (V_x, V_{xx}) from the pre-computed HJI value function to construct the equivalent local weight matrices $M_{1,k}$ and $M_{2,k}$. The moments are then explicitly extracted by solving localized, state-dependent trace-optimization problems subject to the strict energy boundaries.

Phase 3: ZOH inter-sample execution. Having uniquely determined the optimal means and variances $(\Gamma^*_{i,k}, \Sigma^*_{i,k})$, the players inversely project these deterministic moments back into the randomized mixed-strategy spaces. Specifically, they construct multivariate Gaussian distributions $\mathcal{N}(\Gamma^*_{i,k}, \Sigma^*_{i,k})$ and independently draw their control actions u_k and v_k. These sampled controls are then applied to the physical system under a Zero-Order Hold (ZOH) mechanism over the commitment interval $[t_k, t_{k+1})$, naturally inducing a macroscopic stochastic diffusion that physically replicates the continuous-time surrogate game dynamics.

C. Performance Certification

Having synthesized the near-optimal mixed strategies, we now rigorously certify the performance of the proposed methodology by evaluating the bounds on the value approximation error and the suboptimality gaps.

Assumption 5 (Lipschitz mixed strategies)**.** *There exists a pair of saddle-point mixed strategies (α^*, β^*) for the original discrete-time game $\mathbf{G}_{lq}$ such that their projected statistical moments $(\phi_1(\alpha^*), \phi_2(\beta^*))$ are uniformly Lipschitz continuous.*

Remark 4 (Regularity of optimal mixed strategies)**.** *Unlike unconstrained pure strategies, the mathematical regularity of the first two moments for optimal mixed strategies has not been extensively explored in existing differential game literature. Furthermore, as revealed in Theorem 3, the analytical variance of the surrogate SDG inherently involves indicator functions, implying that the exact optimal mappings may exhibit non-Lipschitz jumps on certain boundaries. Therefore, Assumption 5 serves as a pragmatic working hypothesis. It effectively isolates the intrinsic theoretical weak approximation performance of our framework from the profound*

analytical complexities of non-smooth controls, rendering the suboptimality certification mathematically tractable.

As established in Theorem 2, the surrogate game $\tilde{\mathbf{G}}_{lq}$ under pure spaces $\tilde{\mathcal{A}}_p^\pi \times \tilde{\mathcal{B}}_p^\pi$ is an order 2 weak approximation of the original game $\mathbf{G}_{lq}$ under mixed spaces $\mathcal{A}_m^\pi \times \mathcal{B}_m^\pi$. Let $(\tilde{\alpha}^*, \tilde{\beta}^*)$ be the optimal pure strategies for $\tilde{\mathbf{G}}_{lq}$ derived in Theorem 3.

We define the continuous-time best responses against the ZOH-executed opponent strategies as $\tilde{\alpha}^{br} = \arg\min_{\tilde{\alpha}\in\tilde{\mathcal{A}}_p^\pi} \tilde{J}(t_0, x_0, \tilde{\alpha}, \mathrm{ZOH}_\pi[\tilde{\beta}^*])$ and $\tilde{\beta}^{br} = \arg\max_{\tilde{\beta}\in\tilde{\mathcal{B}}_p^\pi} \tilde{J}(t_0, x_0, \mathrm{ZOH}_\pi[\tilde{\alpha}^*], \tilde{\beta})$.

To uniquely isolate the theoretical weak approximation accuracy from the physical implementation error, we define the following ZOH execution penalties:

$$
\begin{aligned}
\epsilon_1 =& |\tilde{J}(t_0, x_0, \tilde{\alpha}^*, \phi_2(\beta^*)) - \tilde{J}(t_0, x_0, \mathrm{ZOH}_\pi[\tilde{\alpha}^*], \phi_2(\beta^*))|,\\
\epsilon_2 =& |\tilde{J}(t_0, x_0, \phi_1(\alpha^*), \tilde{\beta}^*) - \tilde{J}(t_0, x_0, \phi_1(\alpha^*), \mathrm{ZOH}_\pi[\tilde{\beta}^*])|,\\
\epsilon_3 =& |\tilde{J}(t_0, x_0, \tilde{\alpha}^*, \tilde{\beta}^*) - \tilde{J}(t_0, x_0, \mathrm{ZOH}_\pi[\tilde{\alpha}^*], \mathrm{ZOH}_\pi[\tilde{\beta}^*])|,\\
\epsilon_4 =& |\tilde{J}(t_0, x_0, \tilde{\alpha}^{br}, \mathrm{ZOH}_\pi[\tilde{\beta}^*]) - \tilde{J}(t_0, x_0, \tilde{\alpha}^{br}, \tilde{\beta}^*)|,\\
\epsilon_5 =& |\tilde{J}(t_0, x_0, \mathrm{ZOH}_\pi[\tilde{\alpha}^*], \tilde{\beta}^{br}) - \tilde{J}(t_0, x_0, \tilde{\alpha}^*, \tilde{\beta}^{br})|.
\end{aligned}
$$

Lemma 1 (ZOH strategy replacement errors)**.** *The cost differences caused by replacing the continuous-time strategies with their piecewise-constant ZOH equivalents are bounded by:*

$$\epsilon_1 = \mathcal{O}(\bar{\pi}), \quad \epsilon_2 = \mathcal{O}(\bar{\pi}^{\frac{1}{2}}), \quad \epsilon_3 = \mathcal{O}(\bar{\pi}^{\frac{1}{2}}). \tag{15}$$

Proof. We bound the replacement errors by analyzing the state tracking error $e(t) = \tilde{x}^d(t) - \tilde{x}^*(t)$ via Itô's formula and Grönwall's inequality. For ϵ_1, the Lipschitz continuity of Player 1's optimal strategy ensures an $\mathcal{O}(\bar{\pi})$ convergence rate. For ϵ_2, the indicator-function-dependent variance injection introduces non-Lipschitz jump discontinuities; we decouple this deviation into a quadratic state-mismatch error and a time-discretization error, showing that the non-smooth jumps degrade the tracking performance to $\mathcal{O}(\bar{\pi}^{1/2})$. The bound for ϵ_3 follows directly from the dominance of this non-Lipschitz variance error in the combined strategy replacement. Full details are provided in Appendix D. □

Lemma 2 (Best response deviation errors)**.** *The cost variations evaluating the continuous-time best responses against the continuous versus ZOH-executed opponent strategies are bounded by:*

$$\epsilon_4 = \mathcal{O}(\bar{\pi}^{\frac{1}{2}}), \quad \epsilon_5 = \mathcal{O}(\bar{\pi}). \tag{16}$$

Proof. We bound the best-response deviation by analyzing the state tracking error between the ideal continuous-time response and the ZOH-constrained implementation. For ϵ_4, the opponent's strategy inherits non-Lipschitz variance jumps (from Theorem 3), forcing a conservative Cauchy-Schwarz bound on the diffusion error, which yields $\mathcal{O}(\bar{\pi}^{1/2})$ convergence. Conversely, for ϵ_5, Player 1's optimal strategy is zero-variance and globally Lipschitz, eliminating the diffusion discontinuities. This structure allows for a tighter $\mathcal{O}(\bar{\pi})$ bound derived via Grönwall's inequality. Full details are provided in Appendix E. □

Building on these bounded execution penalties, we formally establish the global value approximation error and the strategy suboptimality gaps.

Theorem 4 (Value approximation error)**.** *Under Assumption 5, the analytical game value $\tilde{V}_p(t_0, x_0)$ of $\tilde{\mathbf{G}}_{lq}$ approximates the true game value $V_m^\pi(t_0, x_0)$ of $\mathbf{G}_{lq}$ with an error bounded by:*

$$|V_m^\pi(t_0, x_0) - \tilde{V}_p(t_0, x_0)| = \mathcal{O}(\bar{\pi}^{\frac{1}{2}}). \tag{17}$$

Proof. According to the general SDG value certification framework established in our previous work [24, Theorem 2], the value approximation error is structurally bounded by the sum of the weak approximation error and the dominant execution penalty: $|V_m^\pi - \tilde{V}_p| \le \max\{\epsilon_1, \epsilon_2\} + \mathcal{O}(\bar{\pi}^n)$. By Theorem 2, the ZSLQDG weak approximation inherently guarantees an order $n = 2$ accuracy, contributing an $\mathcal{O}(\bar{\pi}^2)$ error. Concurrently, Lemma 1 bounds the dominant execution penalty at $\epsilon_2 = \mathcal{O}(\bar{\pi}^{\frac{1}{2}})$. Summing these composite errors yields $\mathcal{O}(\bar{\pi}^{\frac{1}{2}}) + \mathcal{O}(\bar{\pi}^2) = \mathcal{O}(\bar{\pi}^{\frac{1}{2}})$. The proof is completed. □

Theorem 5 (Strategy suboptimality gap)**.** *The near-optimal mixed strategies (α^d, β^d) synthesized via Algorithm 1 possess bounded suboptimality gaps against the worst-case responses:*

$$
\begin{aligned}
&J(t_0, x_0, \alpha^d, \beta^d) - \min_{\alpha\in\mathcal{A}_m^\pi} J(t_0, x_0, \alpha, \beta^d) \le \mathcal{O}(\bar{\pi}^{\frac{1}{2}}),\\
&\max_{\beta\in\mathcal{B}_m^\pi} J(t_0, x_0, \alpha^d, \beta) - J(t_0, x_0, \alpha^d, \beta^d) \le \mathcal{O}(\bar{\pi}^{\frac{1}{2}}).
\end{aligned}
\tag{18}
$$

Proof. Following the suboptimality certification framework derived in [24, Theorem 3], the performance gaps against the best responses are strictly upper bounded by the algebraic sum of the weak approximation error and the respective execution penalties. Specifically, the gap for Player 1 is bounded by $\epsilon_3 + \epsilon_4 + \mathcal{O}(\bar{\pi}^n)$, and for Player 2 by $\epsilon_3 + \epsilon_5 + \mathcal{O}(\bar{\pi}^n)$. Substituting $n = 2$ from Theorem 2 and incorporating the $\mathcal{O}(\bar{\pi}^{\frac{1}{2}})$ penalty bounds from Lemmas 1 and 2, the lower-order execution penalty $\mathcal{O}(\bar{\pi}^{\frac{1}{2}})$ mathematically dominates the higher-order $\mathcal{O}(\bar{\pi}^2)$ theoretical gap in both dimensions. The suboptimality bounds are thus established. □

V. Experiment

In this section, we investigate a double-integrator pursuit-evasion problem in a 2-D plane. We cast the problem as a 4-dimensional ZSLQDG, which serves as a canonical pursuit-evasion benchmark in differential game literature. This model is chosen for its ability to represent the fundamental trade-off between acceleration-based control and state trajectory regulation, making it a classic representative of the broader class of linear-quadratic differential games. After specifying the parameters for the system dynamics, cost functional, and admissible mixed strategy spaces, we numerically simulate the game under the proposed dual-routing near-optimal mixed strategies synthesized via Algorithm 1. Given varying commitment delay patterns, our empirical study serves a three-fold purpose: i) to investigate the physical influence of mixed strategies on the game states, control inputs, and accumulated costs; ii) to verify that the designed SDG under pure strategies strictly guarantees a weak approximation of the original game

in the sense of both state distributions and expected costs; and iii) to evaluate the strategy suboptimality gap by pitting the proposed mixed strategies against continuous-time best responses (BR).

A. 2-D Double-Integrator Pursuit-Evasion Game

Let the pursuer's state be $(x_p, v_p) \in \mathbb{R}^2 \times \mathbb{R}^2$ and the evader's state $(x_e, v_e) \in \mathbb{R}^2 \times \mathbb{R}^2$. Under planar double-integrator kinematics for time $t \in [t_0, T]$, the dynamics evolve as

$$\begin{cases} \dot{x}_i(t) = v_i(t), & x_i(t_0) = x_{i,0} \\ \dot{v}_i(t) = a_i(t), & v_i(t_0) = v_{i,0} \end{cases} \tag{19}$$

where $i \in \{p, e\}$ represents the pursuer or the evader, and $x_i, v_i, a_i \in \mathbb{R}^2$ denote position, velocity, and acceleration, respectively. By defining the relative state vector

$$z(t) = \begin{bmatrix} x_r(t) \\ v_r(t) \end{bmatrix} = \begin{bmatrix} x_p(t) - x_e(t) \\ v_p(t) - v_e(t) \end{bmatrix} \in \mathbb{R}^4, \tag{20}$$

we formulate the pursuit-evasion scenario as a 4-dimensional ZSLQDG:

$$(\mathbf{G}_{pe}): \begin{cases} \dot{z}(t) = \underbrace{\begin{bmatrix} \mathbf{0}_{2\times2} & \mathbf{I}_2 \\ \mathbf{0}_{2\times2} & \mathbf{0}_{2\times2} \end{bmatrix}}_{A} z(t) + \underbrace{\begin{bmatrix} \mathbf{0}_{2\times2} \\ \mathbf{I}_2 \end{bmatrix}}_{B_1} \underbrace{a_p(t)}_{u(t)} \\ \qquad + \underbrace{\begin{bmatrix} \mathbf{0}_{2\times2} \\ -\mathbf{I}_2 \end{bmatrix}}_{B_2} \underbrace{a_e(t)}_{v(t)}, \; z(t_0) = \underbrace{\begin{bmatrix} x_r(t_0) \\ v_r(t_0) \end{bmatrix}}_{z_0}, \\ J(t_0, z_0, u, v) = \mathbb{E} \int_{t_0}^{T} \left[z^\top(t) Q z(t) + u^\top(t) R_1 u(t) \right. \\ \qquad \left. - v^\top(t) R_2 v(t) \right] dt + z^\top(T) Q_T z(T), \end{cases}$$

where the expectation is taken over the underlying probability space induced by the mixed strategies. The matrix $Q \in \mathbb{R}^{4\times4}$ penalizes the relative separation and closing-rate errors, $R_1, R_2 \in \mathbb{R}^{2\times2}$ penalize the pursuer's and evader's control efforts respectively, and $Q_T \in \mathbb{R}^{4\times4}$ enforces the terminal-state penalty. The pursuer aims to minimize J while the evader aims to maximize J, both adopting strategies from the admissible mixed strategy spaces in Definition 1.

B. Experiment Setup

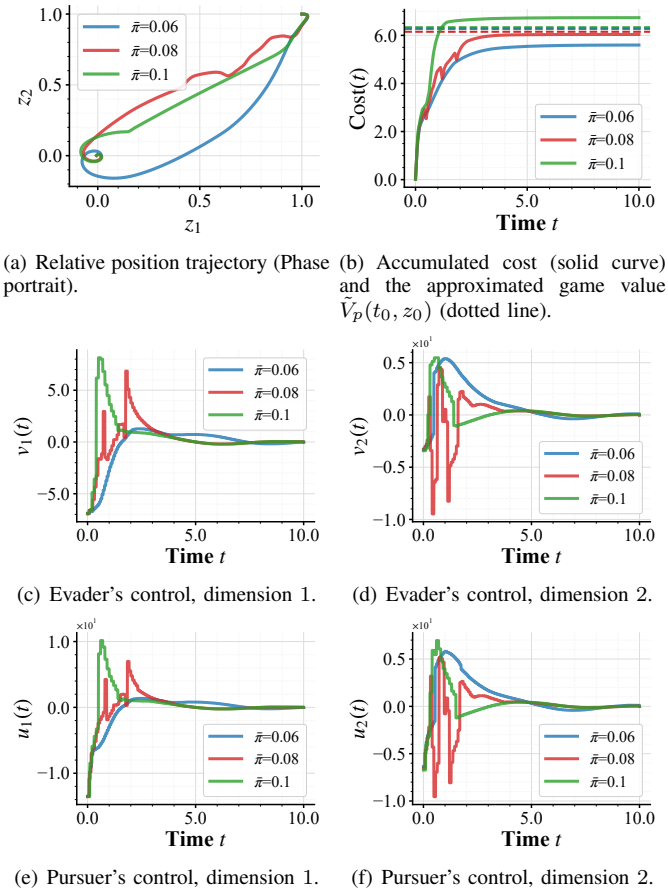
Dynamics and Costs: For the linear dynamics of $\mathbf{G}_{pe}$, we initialize the system at $z_0 = [1, 1, 0.6, 0]^\top$, $t_0 = 0$, with a total horizon $T = 10$. For the quadratic cost functional, the weighting matrices are configured as $Q = \mathbf{I}_4$, $R_1 = 0.1\mathbf{I}_2$, $R_2 = 0.1\mathbf{I}_2$, and $Q_T = \mathbf{I}_4$. Notably, assigning a relatively small control penalty R_2 deliberately intensifies the conflict, effectively incentivizing the evader to activate variance-injection behaviors (i.e., mixed strategies) to exploit the commitment delay.

Admissible Mixed Strategies: We investigate equispaced commitment delay partitions on $[t_0, T]$, such that the interval π satisfies $\bar{\pi} = \underline{\pi}$, and $\bar{\pi}$ is evaluated across $\{0.06, 0.08, 0.1\}$. For the energy constraint, the physical capacities are conservatively bounded by $\gamma_1 = 30$ and $\gamma_2 = 5$.

Numerical Methods: We employ the explicit Runge-Kutta method of order 5(4) to simulate the continuous dynamics of $\mathbf{G}_{pe}$, and the Euler-Maruyama method to discretize the stochastic dynamics of $\tilde{\mathbf{G}}_{pe}$, both operating at a fine time step of $\Delta t = 10^{-3}$. To accurately capture the macroscopic probability distributions of the state trajectories and the accumulated costs, 100 independent Monte Carlo sample paths are simulated concurrently for each $\bar{\pi}$ configuration.

Crucially, to verify the suboptimality gaps, we simulate the game under an asymmetric information structure. While the evaluated player is strictly restricted to the ZOH mixed strategy, the opponent is permitted to continuously observe the state and execute a continuous-time optimal feedback law (acting as the best response). This adversarial proxy mathematically guarantees strict upper and lower empirical bounds for the accumulated costs.

C. Results and Analysis



(a) Relative position trajectory (Phase portrait).

(b) Accumulated cost (solid curve) and the approximated game value $\tilde{V}_p(t_0, z_0)$ (dotted line).

(c) Evader's control, dimension 1.

(d) Evader's control, dimension 2.

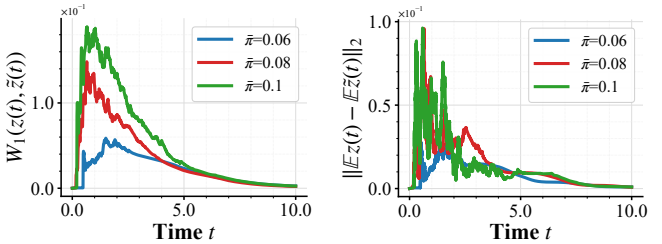
(e) Pursuer's control, dimension 1.

(f) Pursuer's control, dimension 2.

Fig. 2. A realized sample trajectory of $\mathbf{G}_{pe}$, illustrating the physical manifestations of the near-optimal mixed strategies across different commitment frequencies.

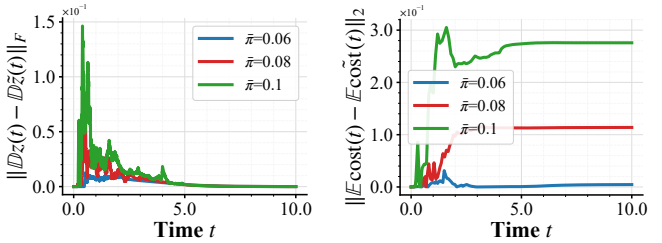
1) Influence of Mixed Strategies: Fig. 2(a) presents the phase trajectory of the relative positions $\{x_r(t)\}_{t\in[t_0,T]}$ under near-optimal mixed strategies across varying commitment frequencies. Noticeably, as $\bar{\pi}$ increases (i.e., the commitment delay widens), the trajectory exhibits intensified stochastic exploration, diverging from the deterministic pure-strategy counterpart.

Fig. 2(b) maps the real-time accumulated costs against the theoretical LQ game values (horizontal dotted lines). The relationship between the commitment delay $\bar{\pi}$ and the baseline game value reveals a profound structural property: the evader actively leverages mixed strategies to exploit the prolonged uncertainty induced by commitment delays, thereby elevating the guaranteed saddle-point value. Furthermore, Fig. 2(c) and Fig. 2(d) plot the realized piecewise-constant control signals of the evader, explicitly demonstrating the artificial variance injected by the normal distributions to disguise their intent. In contrast, as shown in Fig. 2(e) and 2(f), although the pursuer's realized control actions also exhibit fluctuations, this apparent stochasticity is strictly propagated from the state trajectory. Because the pursuer's intrinsic injected variance is zero ($\tilde{\Sigma}_1^* = \mathbf{0}$), their control law remains a purely deterministic mapping of the stochastic state, perfectly corroborating the theoretical degeneration dictated by the dynamic energy allocation law.



(a) 1-Wasserstein distance between the state distributions.

(b) Euclidean distance between the state expectations.



(c) Frobenius distance between the state covariances.

(d) Euclidean distance between the expected accumulated costs.

Fig. 3. Weak approximation verification between the original game $\mathbf{G}_{pe}$ and the surrogate SDG $\tilde{\mathbf{G}}_{pe}$.

2) SDG Weak Approximation Verification: To empirically verify that the constructed surrogate SDG successfully achieves a weak approximation of the original mixed-strategy game, we analyze the distributional divergence between the true state $z(t)$ and the surrogate state $\tilde{z}(t)$. Fig. 3(a) computes the 1-Wasserstein metric $W_1(z(t), \tilde{z}(t)) = \inf_\gamma \mathbb{E}_{(z,\tilde{z})\sim\gamma}\|z - \tilde{z}\|_1$ over the horizon. The approximation errors remain strictly bounded and uniformly diminish as $\bar{\pi}$ decreases, settling at a scale of 10^{-1}.

Delving into the moment evolutions, Fig. 3(b) and 3(c) isolate the approximation errors in the state expectations and covariances, respectively. The empirical results confirm that the first two statistical moments of the discrete ZOH mixed strategy are precisely tracked by the continuous SDG diffusion, with deviations strictly bounded at the 10^{-2} scale. This precise moment-matching practically corroborates the high-order weak approximation accuracy analytically established in Theorem 2. Furthermore, Fig. 3(d) verifies the required consistency in the expected accumulated costs, thereby fulfilling the performance criterion of the SDG weak approximation outlined in Definition 3.

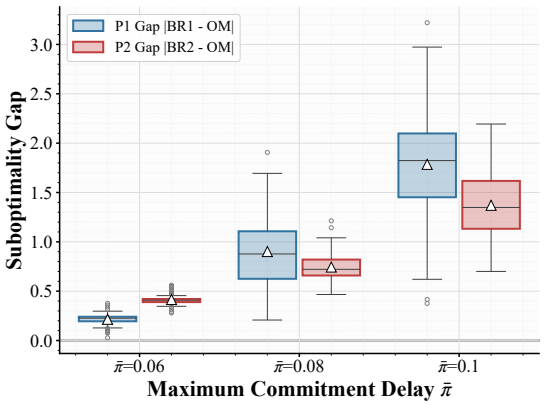


Fig. 4. Empirical evaluation of the strategy suboptimality. By relaxing the ZOH constraint for one player to execute a continuous-time best response, the paired performance gap is robustly bounded.

3) Suboptimality Gap via Continuous Best Responses: The suboptimality gap characterizes the maximum performance degradation a player might suffer when adopting the proposed near-optimal mixed strategy, assuming the opponent executes a worst-case optimal counter-strategy. To empirically quantify this gap and certify the theoretical bounds proposed in Theorem 5, we evaluate our synthesized strategies under an asymmetric information structure. Specifically, while the evaluated player is strictly restricted to the discrete-time ZOH mixed strategy, the adversarial opponent is granted continuous-time state observations to execute the ideal delay-free analytical feedback law. Because this unconstrained continuous-time feedback strictly outperforms any discrete-time counter-strategy, it serves as a conservative, cost-minimizing (or maximizing) proxy for the true best response.

To eliminate the cross-sample variance induced by the stochastic policies, we align the Monte Carlo random seeds, allowing us to evaluate the exact paired deviation over identical Brownian paths. Fig. 4 maps the distributions of the empirical suboptimality gaps for Player 1 and Player 2. The boxplots demonstrate that despite facing a continuously observing and adapting adversary, the performance degradation incurred by the ZOH commitment delay strictly decreases. The gaps progressively vanish toward zero as the $\bar{\pi}$ scales down, corroborating the theoretically guaranteed $\mathcal{O}(\bar{\pi}^{\frac{1}{2}})$ convergence rate of the suboptimality gap established in Theorem 5.

VI. Conclusions

In this paper, we have developed an analytical method for synthesizing near-optimal mixed strategies in zero-sum linear-quadratic differential games (ZSLQDGs). By constructing a surrogate pure-strategy stochastic differential game via second-order moment matching, we have effectively bypassed the intractability of optimizing over infinite-dimensional probability measure spaces. The resulting optimal strategy laws are governed by a Generalized Riccati Differential Equation

(GRDE), which unveils a fundamental dynamic energy allocation law for variance injection in adversarial control.

We have demonstrated that this surrogate construction yields a high-order $\mathcal{O}(\bar{\pi}^2)$ weak approximation of the original system dynamics. Furthermore, we rigorously certified that both the global value approximation error and the strategy suboptimality gaps are bounded by $\mathcal{O}(\bar{\pi}^{1/2})$, providing explicit performance guarantees for our proposed dual-routing execution architecture. Our results offer robust analytical benchmarks for ZSLQDGs and clarify the physical mechanisms of randomization in continuous-time conflict. Extending this moment-matching framework to N-player general-sum differential games is a promising direction for future research, where the equilibrium analysis must generalize from a unique saddle point to more complex Nash equilibria.

REFERENCES

[1] T. Xu, W. Xi, and J. He, "Differential game with mixed strategies: A weak approximation approach," in *2023 62nd IEEE Conference on Decision and Control (CDC)*, Dec. 2023, pp. 5216–5221.

[2] R. Isaacs, *Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization*. Courier Corporation, 1999.

[3] T. Basar, "On the uniqueness of the nash solution in linear-quadratic differential games," *International Journal of Game Theory*, vol. 5, pp. 65–90, 1976.

[4] D. Jacobson, "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games," *IEEE Transactions on Automatic control*, vol. 18, no. 2, pp. 124–131, 2003.

[5] A. J. Weeren, J. M. Schumacher, and J. C. Engwerda, "Asymptotic analysis of linear feedback nash equilibria in nonzero-sum linear-quadratic differential games," *Journal of Optimization Theory and Applications*, vol. 101, pp. 693–722, 1999.

[6] Y. Ho, A. Bryson, and S. Baron, "Differential games and optimal pursuit-evasion strategies," *IEEE Transactions on Automatic Control*, vol. 10, no. 4, pp. 385–389, Oct. 1965.

[7] P. Menon and A. Calise, "Guidance laws for spacecraft pursuit-evasion and rendezvous," in *Guidance, Navigation and Control Conference*. Minneapolis,MN,U.S.A.: American Institute of Aeronautics and Astronautics, Aug. 1988.

[8] V. Turetsky and J. Shinar, "Missile guidance laws based on pursuit–evasion game formulations," *Automatica*, vol. 39, no. 4, pp. 607–618, Apr. 2003.

[9] D. Li and J. B. Cruz, "Defending an asset: A linear quadratic game approach," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 1026–1044, Apr. 2011.

[10] A. Jagat and A. J. Sinclair, "Nonlinear control for spacecraft pursuit-evasion game using the state-dependent riccati equation method," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 6, pp. 3032–3042, Dec. 2017.

[11] S. Aggarwal, T. Başar, and D. Maity, "Linear quadratic zero-sum differential games with intermittent and costly sensing," *IEEE Control Systems Letters*, vol. 8, pp. 1601–1606, 2024.

[12] C. Harris, P. Reny, and A. Robson, "The existence of subgame-perfect equilibrium in continuous games with almost perfect information: A case for public randomization," *Econometrica: Journal of the Econometric Society*, pp. 507–544, 1995.

[13] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems*, vol. 27, 2014.

[14] S. Perkins, P. Mertikopoulos, and D. S. Leslie, "Mixed-strategy learning with continuous action sets," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 379–384, Jan. 2017.

[15] L. Peters, D. Fridovich-Keil, L. Ferranti, C. Stachniss, J. Alonso-Mora, and F. Laine, "Learning mixed strategies in trajectory games," in *Robotics: Science and Systems XVIII*. Robotics: Science and Systems Foundation, Jun. 2022.

[16] J. Hespanha, M. Prandini, and S. Sastry, "Probabilistic pursuit-evasion games: A one-step nash approach," in *Proceedings of the 39th IEEE Conference on Decision and Control*, vol. 3, Dec. 2000, pp. 2272–2277 vol.3.

[17] R. Vidal, O. Shakernia, H. Kim, D. Shim, and S. Sastry, "Probabilistic pursuit-evasion games: Theory, implementation, and experimental evaluation," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 5, pp. 662–669, Oct. 2002.

[18] V. Isler, S. Kannan, and S. Khanna, "Randomized pursuit-evasion in a polygonal environment," *IEEE Transactions on Robotics*, vol. 21, no. 5, pp. 875–884, Oct. 2005.

[19] ——, "Randomized pursuit-evasion with local visibility," *SIAM Journal on Discrete Mathematics*, vol. 20, no. 1, pp. 26–41, Jan. 2006.

[20] R. J. Aumann, "Spaces of measurable transformations," *Bulletin of the American Mathematical Society*, vol. 66, no. 4, pp. 301–304, 1960.

[21] ——, "Borel structures for function spaces," *Illinois Journal of Mathematics*, vol. 5, no. 4, pp. 614–630, 1961.

[22] ——, "28. mixed and behavior strategies in infinite extensive games," in *Advances in Game Theory. (AM-52)*. Princeton University Press, Dec. 1964, pp. 627–650.

[23] J. Engwerda, *LQ Dynamic Optimization and Differential Games*, 1st ed. Chicester, West Sussex, England Hoboken, NJ: Wiley, Jun. 2005.

[24] T. Xu, W. Xi, and J. He, "Near-optimal mixed strategy for zero-sum differential games," *arXiv preprint arXiv:2308.01144*, 2026.

[25] P. Cardaliaguet, "Differential games with asymmetric information," *SIAM Journal on Control and Optimization*, vol. 46, no. 3, pp. 816–838, Jan. 2007.

[26] R. Buckdahn, J. Li, and M. Quincampoix, "Value function of differential games without isaacs conditions. an approach with nonanticipative mixed strategies," *International Journal of Game Theory*, vol. 42, no. 4, pp. 989–1020, Nov. 2013.

[27] P. Cardaliaguet, C. Jimenez, and M. Quincampoix, "Pure and random strategies in differential game with incomplete informations," *Journal of Dynamics and Games*, vol. 1, no. 3, pp. 363–375, 2014.

[28] R. Buckdahn, J. Li, and M. Quincampoix, "Value in mixed strategies for zero-sum stochastic differential games without isaacs condition," *The Annals of Probability*, vol. 42, no. 4, pp. 1724–1768, 2014.

[29] W. H. Fleming and D. Hernandez-Hernandez, "Mixed strategies for deterministic differential games," *Communications on Stochastic Analysis*, vol. 11, no. 2, Jun. 2017.

[30] L. D. Berkovitz and N. G. Medhin, "Relaxed controls," in *Nonlinear Optimal Control Theory*. Chapman and Hall/CRC, 2012.

[31] M. Sîrbu, "Stochastic perron's method and elementary strategies for zero-sum differential games," *SIAM Journal on Control and Optimization*, vol. 52, no. 3, pp. 1693–1711, Jan. 2014.

[32] K. Kunisch, P. Trautmann, and B. Vexler, "Optimal control of the undamped linear wave equation with measure valued controls," *SIAM Journal on Control and Optimization*, vol. 54, no. 3, pp. 1212–1244, Jan. 2016.

[33] A. M. G. Cox, S. Källblad, M. Larsson, and S. Svaluto-Ferro, "Controlled measure-valued martingales: A viscosity solution approach," *The Annals of Applied Probability*, vol. 34, no. 2, pp. 1987–2035, Apr. 2024.

[34] R. Elliott, "The existence of value in stochastic differential games," *SIAM Journal on Control and Optimization*, vol. 14, no. 1, pp. 85–94, Jan. 1976.

[35] G. Wang and Z. Yu, "A pontryagin's maximum principle for non-zero sum differential games of bsdes with applications," *IEEE Transactions on Automatic Control*, vol. 55, no. 7, pp. 1742–1747, Jul. 2010.

[36] W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*. Springer Science & Business Media, Feb. 2006.

[37] P. Kumar, "Optimal mixed strategies in a dynamic game," *IEEE Transactions on Automatic Control*, vol. 25, no. 4, pp. 743–749, Aug. 1980.

[38] Z. Dou, X. Yan, D. Wang, and X. Deng, "Finding mixed strategy nash equilibrium for continuous games through deep learning," *arXiv preprint arXiv:1910.12075*, 2019.

[39] C. Martin and T. Sandholm, "Finding mixed-strategy equilibria of continuous-action games without gradients using randomized policy networks," in *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, ser. IJCAI '23, Aug. 2023, pp. 2844–2852.

[40] A. Levant, "Chattering analysis," *IEEE Transactions on Automatic Control*, vol. 55, no. 6, pp. 1380–1389, Jun. 2010.

[41] B. D. Anderson and J. B. Moore, *Optimal control: linear quadratic methods*. Courier Corporation, 2007.

[42] J. H. Braslavsky, R. H. Middleton, and J. S. Freudenberg, "Feedback stabilization over signal-to-noise ratio constrained channels," *IEEE Transactions on automatic control*, vol. 52, no. 8, pp. 1391–1403, 2007.

[43] W. M. Wonham, "Optimal stationary control of a linear system with state-dependent noise," *SIAM Journal on Control*, vol. 5, no. 3, pp. 486–500, 1967.
[44] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. SIAM, 1998.
[45] G. N. Mil'shtein, "Weak approximation of solutions of systems of stochastic differential equations," *Theory of Probability & Its Applications*, vol. 30, no. 4, pp. 750–766, 1986.

APPENDIX A
PROOF OF THEOREM 1

Proof. The proof proceeds by backward induction. At the terminal time $t_N = T$, the game value uniquely exists as $V_m^\pi(T,x) = x^\top Q_T x$.

Assume the unique game value $V_m^\pi(t_k,\cdot)$ exists at step k. For the local stage game over $[t_{k-1},t_k)$ given $x(t_{k-1}) = x$, let $\mathcal{P}_x^{(1)}$ and $\mathcal{P}_x^{(2)}$ denote the convex sets of probability measures satisfying the energy constraints in Definition 1 with capacities γ_1 and γ_2, respectively. Under Assumption 3, the local expected cost-to-go evaluates to:

$$J_{k-1}(x,\mu,\nu) \triangleq \int_{U\times V} \Bigg\{ \int_{t_{k-1}}^{t_k} h(t,x(t),u,v)dt + V_m^\pi(t_k,x(t_k)) \Bigg\} \mu(du)\nu(dv),$$

To equate the upper and lower values (i.e., $\inf_\mu \sup_\nu J_{k-1} = \sup_\nu \inf_\mu J_{k-1}$), we verify the requisite conditions of Sion's Minimax Theorem: 1) *Weak Compactness:* The uniform second-moment bounds in Definition 1 guarantee that the measure families $\mathcal{P}_x^{(1)}$ and $\mathcal{P}_x^{(2)}$ are tight. By Prokhorov's Theorem, this tightness ensures relative weak compactness. Furthermore, since the second-moment functional is lower semi-continuous in the weak topology, $\mathcal{P}_x^{(i)}$ are closed, and thus strictly weakly compact. 2) *Bilinearity and Continuity:* As an expectation, the functional J_{k-1} is strictly bilinear (hence concave-convex) and continuous with respect to (μ,ν) in the weak topology.

With all conditions satisfied, Sion's Minimax Theorem guarantees the existence of a unique local saddle point, confirming $V_{m,+}^\pi(t_{k-1},x) = V_{m,-}^\pi(t_{k-1},x) \triangleq V_m^\pi(t_{k-1},x)$. By backward induction down to t_0, the global upper and lower value functions coincide, i.e., $V_{m,+}^\pi(t_0,x_0) = V_{m,-}^\pi(t_0,x_0) = V_m^\pi(t_0,x_0)$, completing the proof. □

APPENDIX B
PROOF OF THEOREM 2

Proof. Our certification of the SDG weak approximation proceeds via a two-stage procedure derived from the classical weak approximation theory of SDEs [45]: i) investigate the local one-step approximation errors between the original game dynamics and the designed SDG, and ii) if the local one-step approximation errors for the first three moments are bounded by $\mathcal{O}(\bar\pi^3)$, then the global approximation error over the entire time horizon possesses an order of 2, i.e., $\mathcal{O}(\bar\pi^2)$.

Unlike classical SDE weak approximation, an SDG approximation requires both the state trajectory and the accumulated cost function to be jointly approximated. To facilitate this, we first transform the original game $\mathbf{G}_{lq}$ and the surrogate SDG $\tilde{\mathbf{G}}_{lq}$ into auxiliary forms where the running costs are absorbed into an augmented state variable y. Let $z(t) = [u(t)^\top, v(t)^\top]^\top$, $R \triangleq \mathrm{diag}(R_1,-R_2)$ and $B \triangleq [B_1,B_2]$. The auxiliary games are constructed as:

$$(\mathbf{G}_{lq}^{aux}):\begin{cases} \dot x(t) = Ax(t) + Bz(t), \\ \dot y(t) = \mathbb{E}\left[x^\top(t)Qx(t) + z^\top(t)Rz(t)\right], \\ x(t_0) = x_0, \quad y(t_0) = 0, \\ J = \mathbb{E}\left[y(T) + x^\top(T)Q_T x(T)\right]. \end{cases}$$

$$(\tilde{\mathbf{G}}_{lq}^{aux}):\begin{cases} d\tilde x(t) = (A\tilde x(t) + B\tilde\Gamma(t))dt + \delta^{\frac12}_\pi(t)B\tilde\Lambda(t)dW_t, \\ d\tilde y(t) = \mathbb{E}\Big[\tilde x^\top(t)Q\tilde x(t) + \tilde\Gamma^\top(t)R\tilde\Gamma(t) + \mathrm{tr}(\tilde\Sigma(t)R)\Big]dt, \\ \tilde x(t_0) = x_0, \quad \tilde y(t_0) = 0, \\ \tilde J = \mathbb{E}\left[\tilde y(T) + \tilde x^\top(T)Q_T\tilde x(T)\right]. \end{cases}$$

We now proceed to the local one-step error analysis. Let $\delta = t_1 - t_0$ be the length of the first commitment interval, satisfying $\delta \le \bar\pi$. Let $\Delta = [\Delta_x^\top, \Delta_y]^\top$ be the exact one-step increment for $\mathbf{G}_{lq}^{aux}$, and $\tilde\Delta = [\tilde\Delta_x^\top, \tilde\Delta_y]^\top$ be the increment for the surrogate $\tilde{\mathbf{G}}_{lq}^{aux}$.

Lemma 3 (One-step local approximation)**.** *Under the strategy projection map ϕ defined in Lemma 5, the differences between the local moments of Δ and $\tilde\Delta$ conditional on the initial state are strictly bounded by:*

(i) $\|\mathbb{E}[\Delta_x] - \mathbb{E}[\tilde\Delta_x]\| = 0$,
(ii) $\|\mathbb{D}[\Delta_x] - \mathbb{D}[\tilde\Delta_x]\| = \mathcal{O}(\bar\pi^4)$,
(iii) $\|\mathbb{E}[\Delta_x^{\otimes 3}] - \mathbb{E}[\tilde\Delta_x^{\otimes 3}]\| = \mathcal{O}(\bar\pi^3)$,
(iv) $|\Delta_y - \tilde\Delta_y| = \mathcal{O}(\bar\pi^5)$.

Proof of Lemma 3. For the exact dynamics of $\mathbf{G}_{lq}^{aux}$ under the Zero-Order Hold (ZOH) pattern, the control $z(t)$ is constant over $[t_0,t_1)$. The exact state transition is $x(t) = \Phi(t)x_0 + \Psi(t)z(t_0)$, where $\Phi(t) = e^{A(t-t_0)}$ and $\Psi(t) = \int_{t_0}^t e^{A(t-s)}Bds$. By definition of the projection map, $\tilde\Gamma(t_0) = \mathbb{E}[z(t_0)]$. Furthermore, owing to the conditional independence of the players' sampling mechanisms (Assumption 3), the cross-covariance is zero, yielding $\tilde\Sigma(t_0) = \mathbb{D}[z(t_0)] = \mathrm{diag}(\tilde\Sigma_1(t_0),\tilde\Sigma_2(t_0))$.

Proof of (i): Taking the expectation of the exact state yields $\mathbb{E}[x(t_1)] = \Phi(t_1)x_0 + \Psi(t_1)\tilde\Gamma(t_0)$. Due to the linearity of the surrogate drift and $\mathbb{E}[dW_t] = 0$, the surrogate state expectation strictly evaluates to $\mathbb{E}[\tilde x(t_1)] = \Phi(t_1)x_0 + \Psi(t_1)\tilde\Gamma(t_0)$. Thus, the first moment difference is identically 0.

Proof of (ii): The exact covariance of the ZOH dynamics is governed by the matrix $\Psi(t)$:

$$\mathbb{D}[x(t_1)] = \Psi(t_1)\tilde\Sigma(t_0)\Psi(t_1)^\top.$$

Using the Taylor expansion of the state transition integral $\Psi(t_1) = \int_{t_0}^{t_1} e^{A(t_1-s)}Bds = \delta B + \frac{\delta^2}{2}AB + \mathcal{O}(\delta^3)$, we obtain:

$$\mathbb{D}[x(t_1)] = \delta^2 B\tilde\Sigma B^\top + \frac{\delta^3}{2}\left(AB\tilde\Sigma B^\top + B\tilde\Sigma B^\top A^\top\right) + \mathcal{O}(\delta^4). \tag{21}$$

Concurrently, applying the Itô isometry to the diffusion term of $\tilde{\mathbf{G}}_{lq}^{aux}$, the surrogate covariance is:

$$\mathbb{D}[\tilde x(t_1)] = \delta\int_{t_0}^{t_1} e^{A(t_1-s)}B\tilde\Sigma(t_0)B^\top e^{A^\top(t_1-s)}ds.$$

Expanding the integrand yields:

$$\mathbb{D}[\tilde{x}(t_1)] = \delta^2 B\tilde{\Sigma}B^\top + \frac{\delta^3}{2}\left(AB\tilde{\Sigma}B^\top + B\tilde{\Sigma}B^\top A^\top\right) + \mathcal{O}(\delta^4). \tag{22}$$

Remarkably, owing to the global linearity of the dynamics and the specific $\delta^{\frac{1}{2}}$ scaling design of the diffusion term, the second and third-order Taylor terms in (21) and (22) structurally cancel each other out. Thus, $\|\mathbb{D}[x(t_1)] - \mathbb{D}[\tilde{x}(t_1)]\| = \mathcal{O}(\delta^4) = \mathcal{O}(\bar{\pi}^4)$.

Proof of (iii): In the exact ZOH formulation, the random control input $z(t_0)$ is sampled once and held constant over the interval $[t_0, t_1)$, meaning the exact state increment Δ_x is entirely devoid of intra-interval stochastic driving processes (i.e., no Brownian motion). Thus, the exact increment rigorously evaluates to $\Delta_x = \delta(Ax_0 + Bz(t_0)) + \mathcal{O}(\delta^2)$. Taking the third moment strictly yields $\mathbb{E}[\Delta_x^{\otimes 3}] = \delta^3\mathbb{E}[(Ax_0 + Bz(t_0))^{\otimes 3}] + \mathcal{O}(\delta^4) = \mathcal{O}(\delta^3)$.

For the surrogate SDG, the increment comprises a drift component $D \sim \mathcal{O}(\delta)$ and a pure diffusion component κ. As established in (ii), the variance of the diffusion component scales as $\mathbb{E}[\kappa^{\otimes 2}] = \mathcal{O}(\delta^2)$. Since κ is an Itô integral of a deterministic matrix, it is strictly Gaussian, implying its odd moments identically vanish (i.e., $\mathbb{E}[\kappa] = 0$ and $\mathbb{E}[\kappa^{\otimes 3}] = 0$). Expanding the third moment $\mathbb{E}[(D + \kappa)^{\otimes 3}]$, the only non-zero terms are the pure drift $D^{\otimes 3} \sim \mathcal{O}(\delta^3)$ and the cross-terms involving one drift and two diffusion components. These cross-terms evaluate to permutations of $D \otimes \mathbb{E}[\kappa^{\otimes 2}]$, which specifically scale as $\mathcal{O}(\delta) \cdot \mathcal{O}(\delta^2) = \mathcal{O}(\delta^3)$. Thus, both exact and surrogate third moments inherently reside in $\mathcal{O}(\delta^3)$, bounding their difference rigidly at $\mathcal{O}(\bar{\pi}^3)$.

Proof of (iv): The expected control cost terms inside the integral match identically because $\mathbb{E}[z^\top Rz] = \tilde{\Gamma}^\top R\tilde{\Gamma} + \mathrm{tr}(\tilde{\Sigma}R)$. The accumulated state cost difference over $[t_0, t_1)$ solely depends on the state covariance difference established in (ii):

$$\begin{aligned}\Delta_y - \tilde{\Delta}_y &= \int_{t_0}^{t_1} \mathrm{tr}\left\{Q\left(\mathbb{D}[x(t)] - \mathbb{D}[\tilde{x}(t)]\right)\right\} dt \\ &= \mathrm{tr}\left\{Q\int_{t_0}^{t_1} \mathcal{O}((t - t_0)^4) dt\right\} = \mathcal{O}(\delta^5) = \mathcal{O}(\bar{\pi}^5).\end{aligned}$$

The local one-step proof is thus completed. □

According to Milshtein's theorem [45, Theorem 2], Lemma 3 guarantees that the local weak approximation errors for the first three moments are strictly bounded by $\mathcal{O}(\bar{\pi}^3)$. Accumulating this $\mathcal{O}(\bar{\pi}^3)$ error per step over $N = (T - t_0)/\bar{\pi}$ intervals yields a global approximation error of $\mathcal{O}(\bar{\pi}^2)$.

Consequently, the continuous state trajectory and the accumulated cost of the designed surrogate SDG $\tilde{\mathbf{G}}_{lq}^{aux}$ under pure strategies constitute a strict order 2 weak approximation of the original game $\mathbf{G}_{lq}^{aux}$ under mixed strategies. Reverting the auxiliary variables to the original cost functional yields $|J(t_0, x_0, \alpha, \beta) - \tilde{J}(t_0, x_0, \phi_1(\alpha), \phi_2(\beta))| = \mathcal{O}(\bar{\pi}^2)$, which completes the proof. □

Appendix C
Proof of Theorem 3

Proof. For brevity, we drop the time argument t and denote $\tilde{V}_p$ as $\tilde{V}$. The HJI equation for the continuous-time surrogate SDG is formulated as:

$$-\partial_t \tilde{V} = \inf_{\tilde{u}\in\tilde{\mathcal{C}}_1} \sup_{\tilde{v}\in\tilde{\mathcal{C}}_2} \left\{\tilde{h} + (\partial_x \tilde{V})^\top \tilde{f} + \frac{1}{2}\,\mathrm{tr}\left(\tilde{\Sigma}_{diff}\partial_x^2 \tilde{V}\right)\right\}, \tag{23}$$

where the equivalent drift is $\tilde{f} = Ax + B_1\tilde{\Gamma}_1 + B_2\tilde{\Gamma}_2$ and the equivalent diffusion covariance is explicitly evaluated as $\tilde{\Sigma}_{diff} = \delta_\pi(t)\left(B_1\tilde{\Sigma}_1 B_1^\top + B_2\tilde{\Sigma}_2 B_2^\top\right)$. Substituting the ansatz $\tilde{V}(t, x) = x^\top P(t)x$ into the HJI equation (23) with drift $\tilde{f}$ and diffusion $\tilde{\Sigma}_{diff}$, we observe that the min-max optimization decouples into two independent subproblems:

$$-\dot{P} = A^\top P + PA + Q + \inf_{\tilde{u}} \tilde{H}_1 + \sup_{\tilde{v}} \tilde{H}_2, \tag{24}$$

where the local decoupled Hamiltonians are defined as:

$$\tilde{H}_1 = 2x^\top PB_1\tilde{\Gamma}_1 + \tilde{\Gamma}_1^\top R_1\tilde{\Gamma}_1 + \mathrm{tr}\left((\delta_\pi B_1^\top PB_1 + R_1)\tilde{\Sigma}_1\right),$$

$$\tilde{H}_2 = 2x^\top PB_2\tilde{\Gamma}_2 - \tilde{\Gamma}_2^\top R_2\tilde{\Gamma}_2 + \mathrm{tr}\left((\delta_\pi B_2^\top PB_2 - R_2)\tilde{\Sigma}_2\right).$$

Step 1: Maximizer (Player 2). We solve $\sup_{\tilde{\Gamma}_2, \tilde{\Sigma}_2 \succeq \mathbf{0}} \tilde{H}_2$ subject to $\|\tilde{\Gamma}_2\|^2 + \mathrm{tr}(\tilde{\Sigma}_2) \le \gamma_2^2\|x\|^2$. Notice that the objective function is concave (strictly concave in $\tilde{\Gamma}_2$ and linear in $\tilde{\Sigma}_2$) and the constraints are convex. For any $x \neq 0$, Slater's condition is strictly satisfied (e.g., by selecting an interior point $\tilde{\Gamma}_2 = \mathbf{0}$ and $\tilde{\Sigma}_2 = \epsilon I \succ \mathbf{0}$), which mathematically guarantees strong duality. Thus, the Karush-Kuhn-Tucker (KKT) conditions are both necessary and sufficient for global optimality. Formulating the Lagrangian $\mathcal{L}_2 = 2x^\top PB_2\tilde{\Gamma}_2 - \tilde{\Gamma}_2^\top(R_2 + \lambda I)\tilde{\Gamma}_2 + \mathrm{tr}((Q_{2,eq} - \lambda I)\tilde{\Sigma}_2) + \lambda\gamma_2^2\|x\|^2$, we derive:

(i) *Dual Monotonicity and Optimal Mean:* For the supremum of $\mathcal{L}_2$ to be bounded, dual feasibility strictly requires $Q_{2,eq} - \lambda I \preceq \mathbf{0}$, hence $\lambda \ge \max(0, \lambda_{\max}(Q_{2,eq})) \triangleq \lambda_2(t)$. For any strictly feasible $\lambda > \lambda_2(t)$, the matrix $Q_{2,eq} - \lambda I$ is strictly negative definite, enforcing $\tilde{\Sigma}_2 = \mathbf{0}$, and the optimal unconstrained mean is $\tilde{\Gamma}_2^*(\lambda) = (R_2 + \lambda I)^{-1}B_2^\top Px$. Substituting these yields the dual function $g(\lambda) = x^\top PB_2(R_2 + \lambda I)^{-1}B_2^\top Px + \lambda\gamma_2^2\|x\|^2$. Its derivative evaluates to: $g'(\lambda) = \gamma_2^2\|x\|^2 - \|\tilde{\Gamma}_2^*(\lambda)\|^2$. Under Assumption 4, the matrix inequality $\gamma_2^2 I - PB_2(R_2 + \lambda_2 I)^{-2}B_2^\top P \succeq \mathbf{0}$ inherently holds, mathematically guaranteeing $g'(\lambda_2(t)) \ge 0$. Furthermore, since the matrix inverse $(R_2 + \lambda I)^{-2}$ strictly decreases as λ increases, the term $\|\tilde{\Gamma}_2^*(\lambda)\|^2$ strictly decreases, yielding $g'(\lambda) > 0$ for all $\lambda > \lambda_2(t)$. Because $g(\lambda)$ is strictly increasing on its feasible domain $[\lambda_2(t), \infty)$, the infimum of the dual problem $\min_\lambda g(\lambda)$ is uniquely attained at the left boundary $\lambda^* = \lambda_2(t)$. The optimal mean is therefore uniquely determined as $\tilde{\Gamma}_2^*(t) = (R_2 + \lambda_2(t)I)^{-1}B_2^\top Px \triangleq K_2(t)x$.

(ii) *Complementary Slackness and Optimal Variance:* The structure of the optimal variance strictly bifurcates based on $\lambda_{\max}(Q_{2,eq})$:

- If $\lambda_{\max}(Q_{2,eq}) \le 0$, the multiplier is $\lambda^* = 0$ and $Q_{2,eq} \preceq \mathbf{0}$. To maximize the non-positive trace term $\mathrm{tr}(Q_{2,eq}\tilde{\Sigma}_2)$, Player 2 optimally avoids any variance penalty by assigning $\tilde{\Sigma}_2^* = \mathbf{0}$.
- If $\lambda_{\max}(Q_{2,eq}) > 0$, the optimal multiplier is strictly positive ($\lambda^* = \lambda_{\max}(Q_{2,eq}) > 0$). The matrix $M \triangleq Q_{2,eq} - \lambda^* I \preceq \mathbf{0}$. To maximize $\mathrm{tr}(M\tilde{\Sigma}_2)$, the variance must align with the null space of M, resulting in a rank-1

form: $\tilde{\Sigma}_2^* = c \cdot v_2 v_2^\top$, where v_2 is the principal eigenvector of $Q_{2,eq}$. Crucially, the KKT complementary slackness requires $\lambda^*\big(\gamma_2^2\|x\|^2 - \|\tilde{\Gamma}_2^*\|^2 - \mathrm{tr}(\tilde{\Sigma}_2^*)\big) = 0$. Since $\lambda^* > 0$, the energy constraint must be strictly active, uniquely dictating that the variance must absorb all residual energy: $c = \gamma_2^2\|x\|^2 - \|K_2 x\|^2$.

Combining both cases structurally yields the optimal variance augmented with the indicator function $\mathbb{I}_{\{\lambda_{\max}(Q_{2,eq})>0\}}$ as presented in the theorem.

Step 2: Minimizer (Player 1). Given $Q_{1,eq} \succ \mathbf{0}$, the trace term $\mathrm{tr}(Q_{1,eq}\tilde{\Sigma}_1)$ is minimized at $\tilde{\Sigma}_1^* = \mathbf{0}$. Minimizing $\tilde{H}_1$ w.r.t $\tilde{\Gamma}_1$ yields $\tilde{\Gamma}_1^* = -R_1^{-1}B_1^\top P x$.

Step 3: Synthesis. Substituting the optimal controls back into the HJI equation and identifying coefficients quadratic in x yields the GRDE (9). The existence of $P(t)$ on $[t_0, T]$ is guaranteed by the capacity condition, completing the proof. □

Appendix D
Proof of Lemma 1

Proof. The bounds for ϵ_1, ϵ_2, and ϵ_3 are evaluated systematically through a four-step procedure: error dynamics analysis, state bound estimation, Grönwall's inequality application, and cost difference expansion. Let $e(t) = \tilde{x}^d(t) - \tilde{x}^*(t)$ be the state tracking error.

Part I: Bound for $\epsilon_1 = \mathcal{O}(\bar{\pi})$

Evaluating ϵ_1 corresponds to replacing $\tilde{\alpha}^*$ with $\mathrm{ZOH}_\pi[\tilde{\alpha}^*]$ against $\phi_2(\beta^*)$. From Theorem 3, Player 1's analytical optimal variance strictly vanishes ($\tilde{\Lambda}_1^* \equiv \mathbf{0}$). Thus, the ZOH substitution only affects the drift. Let $\rho_1 = \max_t \|K_1(t)\|$.

Step 1: Error dynamics. Applying Itô's formula to $\|e(t)\|^2$ for $t \in [t_{k-1}, t_k)$:

$$\begin{aligned} d\|e(t)\|^2 = 2e(t)^\top \big\{ & A\tilde{x}^d(t) - A\tilde{x}^*(t) - B_1[K_1(t_{k-1})\tilde{x}^d(t_{k-1}) \\ & - K_1(t)\tilde{x}^*(t)] + B_2[\tilde{\Gamma}_2^d(t_{k-1}) - \tilde{\Gamma}_2^*(t_{k-1})]\big\}dt \\ & + 2e(t)^\top \delta_\pi^{\frac{1}{2}}(t)[0, B_2\tilde{\Lambda}_2^d(t_{k-1}) - B_2\tilde{\Lambda}_2^*(t_{k-1})]dW_t \\ & + \delta_\pi(t)\left\|B_2[\tilde{\Lambda}_2^d(t_{k-1}) - \tilde{\Lambda}_2^*(t_{k-1})]\right\|_F^2 dt. \end{aligned}$$

Step 2: States upper bound. Given that the opponent's strategy $\phi_2(\beta^*)$ satisfies linear growth with parameter κ_2 (note that the linear growth property of the optimal strategies is a direct structural consequence of the state-dependent energy constraints imposed in Definition 1), applying the concavity of the 2-norm and Grönwall's inequality guarantees uniformly bounded second moments. Define $M^* = \|A\| + \rho_1\|B_1\| + \kappa_2\|B_2\|$. It rigorously follows that $\mathbb{E}\|\tilde{x}^*(t)\|^2 \le M_T^*$ and $\mathbb{E}\|\tilde{x}^d(t)\|^2 \le M_T^d$ for some positive constants M_T^* and M_T^d.

Step 3: Error upper bound. By expanding the drift difference into a cross-term $\mathbb{E}\|B_1K_1(t_{k-1})[\tilde{x}^d(t_{k-1}) - \tilde{x}^*(t)]\|$ and a Lipschitz term $\mathbb{E}\|B_1[K_1(t_{k-1}) - K_1(t)]\tilde{x}^*(t)\|$, we bound the difference strictly by $\mathcal{O}(\bar{\pi})$. Because $\phi_2(\beta^*)$ is globally Lipschitz (Assumption 5), there exist constants $a_1, a_2, b_1, b_2 > 0$ such that the expected drift difference is bounded by $a_1\mathbb{E}\|e(t)\| + b_1\bar{\pi}$, and the Frobenius norm difference satisfies $\mathbb{E}\|B_2[\tilde{\Lambda}_2^d - \tilde{\Lambda}_2^*]\|_F \le a_2\mathbb{E}\|e(t)\| + b_2\bar{\pi}$. Taking expectations on $d\|e(t)\|^2$:

$$\frac{d}{dt}\mathbb{E}\|e(t)\|^2 \le 2(\|A\| + a_1^2 + a_2^2\bar{\pi})\mathbb{E}\|e(t)\|^2 + 2b_1^2\bar{\pi}^2 + 2b_2^2\bar{\pi}^3.$$

Applying Grönwall's inequality yields the mean square state error $\mathbb{E}\|e(t)\|^2 = \mathcal{O}(\bar{\pi}^2)$.

Step 4: Cost difference. Substituting $\mathbb{E}\|e(t)\|^2 = \mathcal{O}(\bar{\pi}^2)$ into the quadratic cost functional, the state penalty deviation is bounded via the Cauchy-Schwarz inequality: $\mathbb{E}[\tilde{x}^{*\top}Q\tilde{x}^* - \tilde{x}^{d\top}Q\tilde{x}^d] \le 2\|Q\|\sqrt{\max(M_T^d, M_T^*)}\sqrt{\mathbb{E}\|e(t)\|^2} = \mathcal{O}(\bar{\pi})$. The linear-quadratic control cost deviations structurally follow $\mathcal{O}(\bar{\pi})$. Thus, $\epsilon_1 = \mathcal{O}(\bar{\pi})$.

Part II: Bound for $\epsilon_2 = \mathcal{O}(\bar{\pi}^{\frac{1}{2}})$

Evaluating ϵ_2 replaces $\tilde{\beta}^*$ with $\mathrm{ZOH}_\pi[\tilde{\beta}^*]$ against $\phi_1(\alpha^*)$. Because $\tilde{\beta}^*$ is the exact optimal response (Theorem 3), its variance involves the indicator function $\mathbb{I}_{\{\lambda_{\max}(Q_{2,eq})>0\}}$, which introduces finite non-Lipschitz jump discontinuities.

Step 1 & 2: The error dynamics and state bounding follow the symmetric structure as Part I, guaranteeing bounded second moments M_T^* and M_T^d.

Step 3: Error upper bound. Crucially, although the non-Lipschitz indicator function precludes standard Lipschitz bounding techniques, the structural state-dependent energy constraints universally truncate the maximum possible variance injection. By applying the parallelogram inequality and the submultiplicativity of the Frobenius norm, we have:

$$\begin{aligned} & \left\|B_2\tilde{\Lambda}_2^d(t_{k-1}) - B_2\tilde{\Lambda}_2^*(t_{k-1})\right\|_F^2 \\ & \le 2\|B_2\|^2\left(\|\tilde{\Lambda}_2^d(t_{k-1})\|_F^2 + \|\tilde{\Lambda}_2^*(t_{k-1})\|_F^2\right) \\ & = 2\|B_2\|^2\left(\mathrm{tr}(\tilde{\Sigma}_2^d(t_{k-1})) + \mathrm{tr}(\tilde{\Sigma}_2^*(t_{k-1}))\right). \end{aligned}$$

Because the traces of the optimal variances are explicitly bounded by the state-dependent energy capacity $\gamma_2^2\|\tilde{x}\|^2$, this dynamically evaluates to:

$$\begin{aligned} & \left\|B_2\tilde{\Lambda}_2^d(t_{k-1}) - B_2\tilde{\Lambda}_2^*(t_{k-1})\right\|_F^2 \\ & \le 2\|B_2\|^2\gamma_2^2\left(\|\tilde{x}^d(t_{k-1})\|^2 + \|\tilde{x}^*(t_{k-1})\|^2\right). \end{aligned}$$

Based on the uniform state bounds established in Step 2, there exists a constant $C_x > 0$ such that $\mathbb{E}\|\tilde{x}^d(t_{k-1})\|^2 + \mathbb{E}\|\tilde{x}^*(t_{k-1})\|^2 \le C_x$. Taking expectations on $d\|e(t)\|^2$ injects this bounded local variance jump scaled by $\delta_\pi(t) \sim \mathcal{O}(\bar{\pi})$:

$$\frac{d}{dt}\mathbb{E}\|e(t)\|^2 \le 2(\|A\| + a_1^2)\mathbb{E}\|e(t)\|^2 + 2b_1^2\bar{\pi}^2 + 2\|B_2\|^2\gamma_2^2 C_x\bar{\pi}.$$

Due to the $\mathcal{O}(\bar{\pi})$ constant term, applying Grönwall's inequality degrades the accumulated state error to $\mathbb{E}\|e(t)\|^2 = \mathcal{O}(\bar{\pi})$.

Step 4: Cost difference. Substituting $\mathbb{E}\|e(t)\|^2 = \mathcal{O}(\bar{\pi})$ into the cost functional, the state penalty expands via Cauchy-Schwarz as $\mathbb{E}[\tilde{x}^{*\top}Q\tilde{x}^* - \tilde{x}^{d\top}Q\tilde{x}^d] \le C_Q\sqrt{\mathbb{E}\|e(t)\|^2} = \mathcal{O}(\bar{\pi}^{\frac{1}{2}})$. Crucially, because the optimal variance $\tilde{\Sigma}_2^*(t, \tilde{x})$ operates as a state-feedback mechanism, its deviation must be rigorously decoupled into a state mismatch error and a time discretization error:

$$\begin{aligned} \tilde{\Sigma}_2^d(t) - \tilde{\Sigma}_2^*(t) = & \underbrace{\tilde{\Sigma}_2^*(t_{k-1}, \tilde{x}^d(t_{k-1})) - \tilde{\Sigma}_2^*(t_{k-1}, \tilde{x}^*(t_{k-1}))}_{\text{State mismatch error}} \\ & + \underbrace{\tilde{\Sigma}_2^*(t_{k-1}, \tilde{x}^*(t_{k-1})) - \tilde{\Sigma}_2^*(t, \tilde{x}^*(t))}_{\text{Time discretization error}}. \end{aligned}$$

For the state mismatch error, because $\tilde{\Sigma}_2^*(t,\tilde{x})$ is quadratic in $\tilde{x}$, Cauchy-Schwarz bounds the expected deviation by $\mathcal{O}(\sqrt{\mathbb{E}\|e(t_{k-1})\|^2}) = \mathcal{O}(\bar{\pi}^{\frac{1}{2}})$.

For the time discretization error, note that the indicator function $\mathbb{I}_{\{\lambda_{\max}(Q_{2,eq}(t))>0\}}$ relies exclusively on the deterministic Riccati solution $P(t)$, thus introducing only a finite number of deterministic jumps. Furthermore, the expected quadratic state term $\mathbb{E}[\tilde{x}^*(t)\tilde{x}^*(t)^\top]$ defines the state covariance matrix, which evolves according to a smooth deterministic ordinary differential equation. Consequently, the expectation $\mathbb{E}[\mathrm{tr}(\tilde{\Sigma}_2^*(t,\tilde{x}^*(t))R_2)]$ intrinsically forms a piecewise continuously differentiable function. As a function of bounded variation, its left Riemann sum approximation error mathematically guarantees an $\mathcal{O}(\bar{\pi})$ bound.

Summing these composite mechanisms, the $\mathcal{O}(\bar{\pi}^{\frac{1}{2}})$ tracking error strictly dominates the variance cost difference, yielding $\epsilon_2 = \mathcal{O}(\bar{\pi}^{\frac{1}{2}})$.

Part III: Bound for $\epsilon_3 = \mathcal{O}(\bar{\pi}^{\frac{1}{2}})$

Evaluating ϵ_3 involves ZOH substitutions for both continuous optimal strategies. The error dynamics inherit both the Lipschitz drift delays from Player 1 and the non-Lipschitz variance jumps from Player 2. Driven by the dominant diffusion discontinuity identified in Part II, Grönwall's inequality yields an identical mean square tracking bound $\mathbb{E}\|e(t)\|^2 = \mathcal{O}(\bar{\pi})$. By expanding the state and control cost functionals via the Cauchy-Schwarz inequality, the quadratic discrepancies are structurally governed by $\mathcal{O}(\sqrt{\mathbb{E}\|e(t)\|^2}) + \mathcal{O}(\bar{\pi}) = \mathcal{O}(\bar{\pi}^{\frac{1}{2}})$. Thus, $\epsilon_3 = \mathcal{O}(\bar{\pi}^{\frac{1}{2}})$. □

APPENDIX E
PROOF OF LEMMA 2

Proof. We rigorously establish the deviation bounds for ϵ_4 and ϵ_5 by analyzing the closed-loop continuous-time optimal tracking best responses against the opponent's strategy mismatch.

Part I: Bound for $\epsilon_4 = \mathcal{O}(\bar{\pi}^{\frac{1}{2}})$

Evaluating ϵ_4 quantifies $|\tilde{J}(\tilde{\alpha}^{br}, \mathrm{ZOH}_\pi[\tilde{\beta}^*]) - \tilde{J}(\tilde{\alpha}^{br}, \tilde{\beta}^*)|$. Because the opponent's strategy $\mathrm{ZOH}_\pi[\tilde{\beta}^*]$ preserves linear growth with respect to the state, the continuous optimal best response $\tilde{\alpha}^{br}$ amounts to solving an LQ tracking problem subject to piecewise-constant external signals. Standard linear-quadratic tracking theory ensures that $\tilde{\alpha}^{br}$ consists of a smooth linear state feedback term $\tilde{\Gamma}_1^{br}(x) = -K_1^{br}(t)x + d(t)$ and zero variance $\tilde{\Lambda}_1^{br} = \mathbf{0}$, meaning $\tilde{\alpha}^{br}$ is uniformly globally Lipschitz.

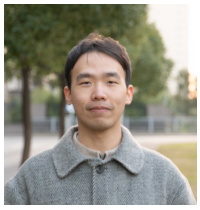
Let $e(t) = \tilde{x}^d(t) - \tilde{x}^*(t)$ denote the state error when evaluating the fixed tracking response $\tilde{\alpha}^{br}$ against $\mathrm{ZOH}_\pi[\tilde{\beta}^*]$ versus $\tilde{\beta}^*$. Because $\tilde{\beta}^*$ contains the non-Lipschitz variance indicator jumps (Theorem 3), the diffusion error over $[t_{k-1}, t_k)$ mimics the exact structural discrepancy detailed in Part II of Lemma 1. Namely, $\mathbb{E}\|B_2(\tilde{\Lambda}_2^d - \tilde{\Lambda}_2^*)\|_F^2 \le \mathcal{O}(1)$. Applying Itô's formula and Grönwall's inequality yields a mean square state error of $\mathbb{E}\|e(t)\|^2 = \mathcal{O}(\bar{\pi})$.

Bounding the quadratic state penalty difference exclusively through the Cauchy-Schwarz inequality provides $\mathbb{E}[\tilde{x}^{*\top} Q \tilde{x}^* - \tilde{x}^{d\top} Q \tilde{x}^d] \le C\sqrt{\mathbb{E}\|e(t)\|^2} = \mathcal{O}(\bar{\pi}^{\frac{1}{2}})$. Therefore, $\epsilon_4 = \mathcal{O}(\bar{\pi}^{\frac{1}{2}})$.

Part II: Bound for $\epsilon_5 = \mathcal{O}(\bar{\pi})$

Symmetrically, evaluating ϵ_5 requires to quantify $|\tilde{J}(\mathrm{ZOH}_\pi[\tilde{\alpha}^*], \tilde{\beta}^{br}) - \tilde{J}(\tilde{\alpha}^*, \tilde{\beta}^{br})|$. The optimal tracking response $\tilde{\beta}^{br}$ is naturally globally Lipschitz.

Crucially, the deviation is driven exclusively by the ZOH mechanism acting on Player 1's analytical strategy $\tilde{\alpha}^*$. Since Player 1's exact optimal variance is structurally zero ($\tilde{\Lambda}_1^* \equiv \mathbf{0}$), $\mathrm{ZOH}_\pi[\tilde{\alpha}^*]$ does not introduce any non-Lipschitz jump discrepancies in the diffusion matrix (i.e., $\mathbb{E}\|B_1(\tilde{\Lambda}_1^d(t_{k-1}) - \tilde{\Lambda}_1^*(t_{k-1}))\|_F^2 \equiv 0$). The state error $e(t)$ is propelled purely by the Lipschitz continuous drift delays, perfectly analogous to Part I of Lemma 1. Consequently, Grönwall's inequality preserves the higher-order tracking bound, yielding $\mathbb{E}\|e(t)\|^2 = \mathcal{O}(\bar{\pi}^2)$. Substituting this refined error back into the cost functional immediately bounds the state penalty difference via Cauchy-Schwarz by $\mathcal{O}(\sqrt{\mathbb{E}\|e(t)\|^2}) = \mathcal{O}(\bar{\pi})$. The control cost differences follow similarly, ensuring $\epsilon_5 = \mathcal{O}(\bar{\pi})$. □



Tao Xu (S'22) received a B.S. degree in the School of Mathematical Sciences from Shanghai Jiao Tong University (SJTU), Shanghai, China, in 2022. He is currently working toward the Ph.D. degree with the Department of Automation, SJTU. He is a member of Intelligent of Wireless Networking and Cooperative Control Group. His research interests include probabilistic prediction, distributionally robust optimization, dynamic games, and robotics.

Wang Xi (S'24) received the B.Eng. degree from the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University (SJTU), Shanghai, China. He is currently pursuing the Ph.D. degree with the School of Automation and Intelligent Sensing, SJTU. His research interests lie in the development of autonomous robotic systems across multiple abstraction layers, ranging from planning and control of mobile robots, to the design paradigms and architectures of robotic middleware, and further to real-time task scheduling and runtime mechanisms for individual robotic systems.

Jianping He (SM'19) is currently a full professor in the School of Automation and Intelligent Sensing at Shanghai Jiao Tong University, Shanghai, China. He received the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2013, and had been a research fellow in the Department of Electrical and Computer Engineering at University of Victoria, Canada, from Dec. 2013 to Mar. 2017. His research interests mainly include the distributed learning, control and optimization, security and privacy in network systems.