electronics

MDPI

Article

Sparse Ergodic Control with Control-Dependent Noise via Physics-Informed Neural Networks

Zhaosheng Xu [1], Jianbang Liu [1,*], Mei Choo Ang [2], Zhongming Liao [1], Kok Weng Ng [3] and Ah-Lian Kor [4]

[1] College of Mathematics and Computer, Xinyu University, Xinyu 338004, China; xuzhaosheng@xyc.edu.cn (Z.X.); liaozhongming@xyc.edu.cn (Z.L.)
[2] Institute of Visual Informatics, Universiti Kebangsaan Malaysia (UKM), Bangi 43600, Selangor, Malaysia; amc@ukm.edu.my
[3] Department of Mechanical, Materials and Manufacturing Engineering, University of Nottingham Malaysia, Semenyih 43500, Selangor, Malaysia; kedznkk@nottingham.edu.my
[4] School of Built Environment, Engineering, and Computing, Leeds Beckett University, Leeds LS6 3QS, UK; a.kor@leedsbeckett.ac.uk
* Correspondence: liujianbang@xyc.edu.cn

Abstract

Sparse ergodic control provides a natural framework for long-run stochastic decision-making under resource constraints. Existing formulations, however, are typically restricted to control-affine systems with control-independent diffusion. When the diffusion coefficient depends explicitly on the control input, the associated ergodic Hamilton–Jacobi–Bellman (HJB) equation becomes non-separable through the term $\mathrm{tr}\left[a(x,\ u)\nabla^2 V\right]$, so classical arguments based on control-affine separability no longer apply directly. In this work, we study sparse ergodic control of stochastic systems with control-dependent diffusion and nonlinear dynamics within a viscosity-solution and learning-based framework. To address the discontinuous ℓ_0-type sparsity penalty, we introduce smooth non-convex sparsity approximations that preserve differentiability while retaining sparse threshold behavior. Within a viscosity-solution framework, we analyze the existence and uniqueness properties of the associated ergodic pair and establish localized approximation error estimates for the smooth approximation. We further characterize a quasi-threshold sparse structure of the resulting optimal feedback policies in non-affine stochastic systems with control-dependent noise. On the computational side, we develop a Physics-Informed Neural Network (PINN)-based solver with adaptive residual-driven sampling for high-dimensional sparse ergodic HJB equations, together with a distributed monotone-inspired iterative scheme for weakly coupled multi-agent systems. Numerical experiments on multi-robot swarm navigation and renewable-integrated smart-grid control demonstrate that the proposed methods produce sparse control policies while preserving stable long-run performance under stochastic disturbances.

Keywords: sparse ergodic control; control-dependent diffusion; ergodic Hamilton–Jacobi–Bellman equations; smooth sparsity approximation; physics-informed neural networks

Check for updates

Received: 24 May 2026
Revised: 29 June 2026
Accepted: 6 July 2026
Published: 13 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. Background

Controlling dynamical systems under uncertainty is a central problem in modern control theory, with applications ranging from robotics and energy systems to biological

regulation [1–3]. Over the past two decades, sparse control—encompassing control policy designs with temporally intermittent activities or spatially distributed actuators—has emerged as an approach for resource-efficient decision-making [4–6]. Based on parsimonious actuation principle, sparse control reduces energy consumption, actuator wear and tear, and communication overhead, while maintaining stability and performance [7]. These properties are relevant for large-scale, networked, and autonomous systems with constrained computation and energy resources [8].

The theoretical core of sparse control focuses on control activity penalization, often quantified by ℓ_0-pseudo-norm, which counts the number of nonzero control components [9]. However, ℓ_0 cost optimization is combinatorial, non-convex and NP-hard which results in computation intractability for most practical problems. This has motivated research on convex and smooth surrogates, such as ℓ_1 regularization or exponential smoothing, which preserve differentiability while approximating the sparsity-inducing behavior of ℓ_0. These relaxations have enabled scalable algorithms and sharp theoretical guarantees, particularly in finite-horizon and deterministic settings, where equivalence results between ℓ_0 and its convex relaxations have been rigorously established.

Despite these advances, the majority of existing sparse optimal control theories primarily focus on transient regulation or terminal cost minimization without considering systems that operate continuously in time and are subject to stochastic perturbations [10–12]. For such systems (e.g., ranging from industrial processes to autonomous agents in uncertain environments), long-term statistical performance rather than transient behavior is the fundamental concern. The natural theoretical framework for these problems is ergodic control, where the goal is to minimize long-run average cost over infinite horizons. Classical ergodic control theory provides existence results for stationary optimal policies via the Hamilton–Jacobi–Bellman (HJB) equation [13]. However, integrating non-smooth sparsity penalties into this stochastic infinite-horizon setting introduces two technical difficulties. First, the resulting HJB equation is fully nonlinear and degenerate elliptic with a non-differentiable Hamiltonian, so classical smooth-solution techniques do not directly apply. Second, allowing the diffusion coefficient to depend on the control makes the Hamiltonian non-separable through the second-order diffusion term, invalidating standard control-affine arguments.

1.2. Related Work and Their Limitations

Recent studies have initiated a systematic treatment of sparse ergodic and stochastic control problems for diffusion processes [14–16]. These works established viscosity-solution formulations of HJB equations, analyzed the existence of optimal stationary Markov controls under Lyapunov-type stability conditions, and explored sparse optimality structures and ℓ_0/ℓ_1 relations in control-affine settings. However, several limitations remain that restrict theoretical generality and computational applicability.

Control-Independent Noise: Existing formulations typically assume that the diffusion term of the stochastic dynamics is independent of the control input. This simplification is convenient for analysis but unrealistic for many systems (e.g., robotic manipulators with actuator noise, power converters with control-dependent variance, or biochemical networks) where regulation intensity modulates intrinsic fluctuations. When diffusion depends on the control, the ergodic HJB equation gains additional nonlinear coupling, where existing existence and uniqueness proofs no longer apply.

Affine Dynamics and Separable Costs: Theoretical results on sparse structure, including bang-off-bang or threshold-type optimal policies, rely heavily on control-affine drift dynamics and separable, box-constrained cost functions. Real-world systems often exhibit nonlinear, saturating, or coupled actuation, where standard analytical techniques (e.g., convex duality or Pontryagin's principle) fail to characterize optimal policies.

Computational Intractability in High Dimensions: Most existing numerical schemes for ergodic control rely on grid-based discretization of the state space, which scales exponentially with dimension and becomes infeasible beyond a few states. This "curse of dimensionality" effectively confines sparse ergodic control to low-dimensional models, raising the issue on how to compute sparse ergodic policies in realistic, high-dimensional stochastic systems.

Limited Multi-Agent Extensions: Sparse control is especially relevant in distributed or networked systems (such as robot swarms or sensor networks) where selective activation of communication and control channels can save substantial resources. Yet, current ergodic formulations do not address inter-agent coupling, nor do they consider robustness to communication delays, packet loss, or adversarial disruptions, which are inevitable in real deployments.

These limitations motivate the present work. When the diffusion coefficient $\sigma(x, u)$ depends on the control, the ergodic Hamiltonian becomes non-separable through the term $\mathrm{tr}\left[a(x, u)\nabla^2 V\right]$, invalidating control-affine arguments and requiring both a modified viscosity argument and a uniform Foster–Lyapunov bound over all admissible controls. Addressing these two technical gaps is the focus of Sections 3 and 4.

1.3. Contributions

This paper addresses four fundamental limitations of existing sparse stochastic control methods, namely the reliance on control-independent diffusion assumptions, restrictive affine system structures, limited computational scalability, and the lack of distributed formulations for multi-agent systems. To overcome these challenges, we develop a unified theoretical and computational framework for sparse ergodic control with control-dependent diffusion. The main contributions are summarized below, with each contribution directly corresponding to one of the identified research gaps.

(C1) Well-posedness for non-separable ergodic HJB with control-dependent diffusion (addresses: Control-Independent Noise). When the diffusion coefficient $\sigma(x, u)$ depends on the control, the ergodic Hamiltonian is non-separable through $\mathrm{tr}\,[a(x, u)\,\nabla^2 V(x)]$, and the control optimization can no longer be decoupled from the diffusion term. Extending existing sparse ergodic control results [14–16] to this setting requires additional stability and comparison-principle conditions, which are formalized in Sections 2 and 3. Under these conditions, we establish viscosity-solution existence and uniqueness for the associated ergodic pair (λ^*, V). The practical consequence is that the framework accommodates systems where regulation intensity modulates intrinsic stochasticity—such as robotic actuators with speed-dependent noise and power converters with control-dependent variance—for which control-independent formulations may fail to accurately capture the interaction between control actions and stochastic variability.

(C2) Quasi-threshold sparse policy structure for non-affine systems (addresses: Affine Dynamics and Separable Costs). Existing characterizations of threshold-type sparse policies rely on control-affine drift and separable costs. For non-affine systems, these structural arguments are not directly applicable. We introduce a smooth non-convex surrogate for the ℓ_0 penalty, prove $O(\varepsilon)$ localized approximation error bounds on compact subsets (Theorem 2), and show that the resulting optimal policies retain a quasi-threshold structure in which each control component activates only when a state-dependent switching function exceeds a prescribed activation threshold (Proposition 1). This activation structure provides interpretable, component-wise decision rules for individual actuators, which is relevant in applications where the basis for control decisions must be transparent.

(C3) PINN-based solver scalable to moderate state-space dimensions (addresses: Computational Intractability in High Dimensions). Grid-based solvers for ergodic HJB equations

scale exponentially with state dimension and become infeasible beyond approximately 4–5 states. We reformulate the sparse ergodic HJB equation as a PINN learning problem, jointly approximating $V(x)$ and the ergodic constant λ via residual minimization over adaptively sampled collocation points. The adaptive sampling strategy reduces required training iterations by approximately 30% relative to uniform sampling. The method is validated on problems up to 7 state dimensions in the main experiments and on a 15-dimensional synthetic benchmark in Section 5.2.4.

(C4) Distributed monotone scheme for weakly coupled multi-agent systems (addresses: Limited Multi-Agent Extensions). We design a distributed iterative scheme in which each agent solves a local sparse ergodic HJB problem using neighbor-exchanged low-rank coupling coefficients, without central coordination. Under the weak-coupling condition $\max_{i,j}|\gamma^{ij}| \leq \overline{\gamma} < \frac{c_1}{2L_V}$, the scheme converges to a ρ-neighborhood of the centralized solution with $\rho = O(\overline{\gamma}\Lambda_{\max}(\mathcal{G}) + (1-\delta)\Delta t)$ (Proposition 2), making it directly applicable to robot swarms and distributed energy resource networks where central coordination is infeasible.

The proposed framework is validated on multi-robot swarm navigation and renewable-integrated smart-grid scheduling benchmarks, demonstrating that substantial sparsity can be achieved without sacrificing control performance across representative engineering applications.

1.4. Organization of the Paper

The remainder of this paper is organized as follows. Section 2 formulates the controlled Itô diffusion model, the sparse ergodic cost, and the assumptions used throughout. Section 3 develops the theoretical foundations, including the ergodic HJB characterization and approximation results for the smooth sparsity penalty. Section 4 introduces the computational framework: a PINN-based solver for high-dimensional problems and a distributed monotone scheme for multi-agent systems. Section 5 presents numerical experiments on robotic and power-system applications, together with ablations and reproducibility details. Section 6 discusses limitations and future work, and Section 7 concludes.

2. Problem Formulation and Underlying Concepts

2.1. Notation

For clarity, we first introduce the key mathematical notation used throughout this paper in Table 1:

Table 1. Notation and Definitions.

Symbol	Domain/Type	Description
X_t	$\mathbb{R}^n$	System state at time t
u_t	$U \subseteq \mathbb{R}^m$	Control input at time t
u	$U \subseteq \mathbb{R}^m$	Generic control variable used in static optimization expressions such as $\inf_{u \in U}\{\cdot\}$
u_t^i	$U_i \subseteq \mathbb{R}^{m_i}$	Control input of agent i at time t; used in multi-agent SDE dynamics and multi-agent cost functionals
u_i	$U_i \subseteq \mathbb{R}^{m_i}$	Static control variable of agent i in local optimization expressions
$b(x,u)$	$\mathbb{R}^n \times U \to \mathbb{R}^n$	Drift function of the controlled stochastic system
$\sigma(x,u)$	$\mathbb{R}^n \times U \to \mathbb{R}^{n\times n}$	Control-dependent diffusion coefficient matrix, possibly degenerate
$a(x,u)$	$\mathbb{R}^{n\times n}$	Controlled diffusion matrix defined as $a(x,u) = \sigma(x,u)\sigma(x,u)^\top$, possibly degenerate
$\mathcal{L}^u$	Second-order differential operator	Infinitesimal generator of the controlled diffusion process

Table 1. *Cont.*

Symbol	Domain/Type	Description
$f_0(x)$	$\mathbb{R}^n \to [0, \infty)$	Continuous, radially unbounded state cost function
$\phi_\varepsilon(u)$	$U \to [0, \infty)$	Smooth, non-convex sparsity-promoting penalty function
$\phi_0(u)$	$U \to [0, \infty)$	original ℓ_0 penalty function
λ^*	$\mathbb{R}$	Optimal ergodic (long-run average) cost
$V(x)$	$\mathbb{R}^n \to \mathbb{R}$	Relative value function, unique up to an additive constant
$\mathcal{U}_{ad}$	Set of stochastic processes	Set of all admissible progressively measurable control policies

2.2. Stochastic System Model

We consider a continuous, time-homogeneous stochastic system with state evolution governed by a controlled Itô stochastic differential equation (SDE) on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and a natural filtration $\{\mathcal{F}_t\}_t \geq 0$ generated by an n-dimensional standard Brownian motion W_t. The state dynamics are given by:

$$dX_t = b(X_t, u_t)\,\mathrm{dt} + \sigma(X_t, u_t)\,\mathrm{d}W_t \tag{1}$$

where

$X_t \in \mathbb{R}^n$ is the system state at time t; $X_0 = x \in \mathbb{R}^n$ is a given deterministic initial state.

$u_t \in U \subset \mathbb{R}^m$ is the control input at time t, with U a nonempty, compact, and convex set. A typical case is the box constraint $U = \Pi_{i=1}^m [-u_i^{max}, u_i^{max}]$.

b: $\mathbb{R}^n \times U \to \mathbb{R}^n$ is the drift function, modeling the deterministic evolution of the system.

σ: $\mathbb{R}^n \times U \to \mathbb{R}^{n \times n}$ is the diffusion coefficient matrix, which is now explicitly dependent on the control u_t. This captures scenarios where control actions directly influence the intensity or direction of the system noise, a key generalization from prior work.

The term $a(x, u) := \sigma(x, u)\,\sigma(x, u)^\intercal$ defines the controlled diffusion matrix. We allow for both uniform ellipticity (i.e., $a(x, u) \succsim \eta I$ for some $\eta > 0$) and the more challenging degenerate ellipticity (i.e., $a(x, u) \succsim 0$), where the diffusion may not occur in all state directions.

An admissible control $u = \{u_t\}_t \geq 0$ is a stochastic process that is progressively measurable with respect to $\{\mathcal{F}_t\}$ and satisfies $u_t \in U$ for almost every t ≥ 0. We denote the set of all admissible controls by $\mathcal{U}_{ad}$.

2.3. Sparse Ergodic Cost with Smooth Penalty

The control objective is to minimize a long-run average cost that penalizes both state deviation and control effort, with a specific emphasis on inducing sparsity in the control signal. We define the ergodic cost functional as:

$$J(u) := \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_x \left[\int_0^T (f_0(X_t) + \phi_\varepsilon(u_t)) dt \right] \tag{2}$$

where:

f_0: $\mathbb{R}^n \to [0, \infty)$ is a continuous, radially unbounded state cost function with at most polynomial growth. It quantifies the penalty for state deviation from a desired equilibrium or trajectory.

ϕ_ε: $U \to [0, \infty)$ is a smooth sparsity-promoting penalty function, parameterized by $\varepsilon > 0$.

The classical, non-convex ℓ_0-penalty is defined as:

$$\phi_0(u) = \sum_{i=\mathbf{1}}^m a_i 1_{\{u_i \neq 0\}} \tag{3}$$

To address its discontinuity, we employ the smooth approximation:

$$\phi_\varepsilon(u) = \sum_{i=1}^{m} a_i(1 - e^{-|u_i|/\varepsilon}) \tag{4}$$

To visualize the difference between ℓ_0, and ℓ_1 penalties, and the proposed smooth approximation, we plot these penalty functions in Figure 1. As shown, the ℓ_0 penalty is discontinuous at u = 0, while the ℓ_1 penalty is convex but does not promote exact sparsity. Our smooth approximation, with different values of ε, bridges these two extremes, that is, for small ε, it provides a pointwise approximation to the ℓ_0 penalty outside an ε-neighborhood of zero, and for larger ε, it provides a smooth surrogate that facilitates gradient-based optimization.

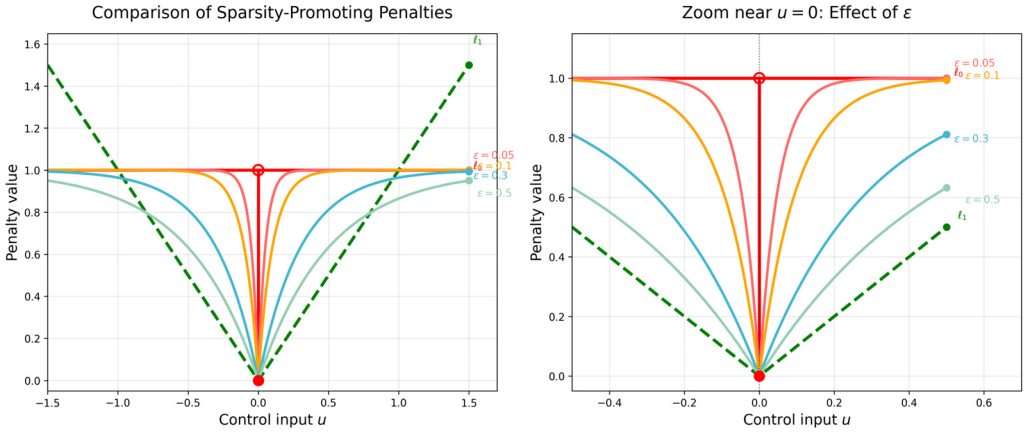


Figure 1. Comparison of Sparsity−Promoting Penalties.

This function possesses several desirable properties:

1. The smooth penalty function is infinitely smooth (C^∞) for all $u \neq 0$ and is continuous (C^0) at the origin; its left and right derivatives at each $u_i = 0$ exist but differ ($\pm a_i/\varepsilon$), so it is globally Lipschitz but not C^1 at the origin. This non-differentiability at zero is a feature that promotes exact sparsity, consistent with the ℓ_0 structure;
2. Pointwise Convergence: As $\varepsilon \downarrow 0$, $\phi_\varepsilon(u) \uparrow \phi(u)$ pointwise for all $u \in U$;
3. Interpretability: The parameter ε controls the sharpness of the approximation. For small ε, it acts as a hard threshold, while for larger ε, it provides a smooth, graded penalty.

The optimal long-run average cost, or the ergodic value, is defined as:

$$\lambda^* := \inf_{u \in \mathcal{U}_{ad}} J(u) \tag{5}$$

where $J(u)$ denotes the long-run average (ergodic) cost induced by the admissible control u, and $\mathcal{U}_{ad}$ is the set of all admissible progressively measurable control policies.

2.4. Explicit Assumptions for Theoretical Analysis

To ensure the problem's well posedness and facilitate subsequent analysis, we make the following assumptions:

Assumption 1 (Dynamics Regularity)**.** *The coefficients b (x, u) and σ(x, u) satisfy:*

1. *Carathéodory and Continuity: For each $u \in U$, b (·, u) and σ(·, u) are Borel measurable. For each $x \in \mathbb{R}^n$, b (x, ·) and σ(x, ·) are continuous on U;*

2. *Lipschitz continuity: There exists a constant* $L_{Lip} > 0$ *such that for all* $x, y \in \mathbb{R}^n$ *and all* $u \in U$*:*

$$|b(x,u) - b(y,u)| + |\sigma(x,u) - \sigma(y,u)| \leq L_{Lip}|x - y| \tag{6}$$

3. *Linear Growth: There exists a constant* $L_g > 0$ *such that for all* $x \in \mathbb{R}^n$ *and all* $u \in U$*:*

$$|b(x,u)| + |\sigma(x,u)| \leq L_g(1 + |x|) \tag{7}$$

4. *Control Set Properties: U is compact and convex with non-empty interior.*

Under these conditions, for any admissible control, the SDE (1) admits a unique strong solution.

Assumption 2 (Foster–Lyapunov Stability [17])**.** *There exists a function* $\mathcal{V} \in C^2(\mathbb{R}^n)$*, with* $\mathcal{V} > 1$*, constants* $c_0 \in \mathbb{R}$*,* $c_1, c_2 > 0$*,* $c_3 \geq 0$ *and a compact set* $K \subset \mathbb{R}^n$ *such that for all* $x \notin K$*:*

$$\sup_{u \in U} \mathcal{L}^u \mathcal{V}(x) \leq c_0 - c_1 \mathcal{V}(x) \tag{8}$$

Here $c_0 \in \mathbb{R}$ *is allowed to be non-positive, corresponding to stronger dissipativity conditions than the standard Foster–Lyapunov setting. Because the diffusion matrix a (x, u) depends explicitly on the control variable u, the supremum over admissible controls is well-defined only if a (x, u) remains uniformly bounded in operator norm on compact subsets of the admissible control set U. Under Assumption 1 and the compactness of U, this boundedness condition holds uniformly with respect to admissible controls.*

$$\mathcal{V}(x) \geq 1 + c_2|x|^p \text{ for } p > 0 \tag{9}$$

$$|\nabla \mathcal{V}(x)| + \left|\nabla^2 \mathcal{V}(x)\right| \leq c_3(1 + \mathcal{V}(x)) \tag{10}$$

This assumption ensures the stability and tightness of the process, guaranteeing the existence of invariant measures for stationary Markov controls and finite-time explosion prevention.

Assumption 3 (Comparison Principle and HJB Regularity [18])**.** *The Hamiltonian* $H^\varepsilon(x, p, X)$ *is continuous in* (x, p, X)*, and the comparison principle holds within the class of continuous viscosity sub-solutions and super-solutions with polynomial growth under the following structural conditions: (i) The diffusion matrix a (x, u)* σ *(x, u)* σ $(x,\ u)^\top$ *is positive semidefinite for all admissible states and controls. In the uniformly elliptic case, a (x, u)* $\succeq \eta \boldsymbol{I}$ *for some* $\eta >\ 0$*. For degenerate diffusion, the coefficient* σ *(x, u) is assumed to satisfy the regularity conditions required by the viscosity-solution framework, including uniform Lipschitz continuity with respect to the state variable. (ii) The Hamiltonian satisfies the coercivity condition* $H^\varepsilon(x, p, 0) \geq v_1|p| - v_2(1 + |x|^q)$ *for constants* $v_1 > 0$ *and* $v_2 \geq 0$*. Under these assumptions, standard comparison results for viscosity solutions of degenerate elliptic HJB equations apply [18–20], ensuring uniqueness of* λ^* *and uniqueness of V up to an additive constant.*

2.5. Sparse Ergodic Control Problem

We can formally state the core problem of this research work.

Problem 1 (Sparse Ergodic Control with Control-Dependent Noise)**.** *We seek an admissible control policy* $u^* \in U_{ad}$ *that achieves (or approximates) the infimum of the ergodic cost:*

$$\lambda^* := \inf_{u \in \mathcal{U}_{ad}} J(u) \tag{11}$$

The existence of such an optimal stationary Markov control is established in Theorem 1, where the system dynamics satisfy Assumption 1, the stability condition is given by Assumption 2, and the ergodic HJB structure satisfies Assumption 3. The objective incorporates the smooth sparsity penalty defined in Section 2.3.

A solution to this problem involves not only characterizing the optimal cost λ^* but also deriving an optimal feedback control law $u^*(x)$ that is sparse, which implies that many of its components are zero for large subsets of the state space $\mathbb{R}^n$.

In the following section, we will analyze the corresponding ergodic HJB equation for this problem, prove the existence and uniqueness of its solution, and establish the convergence properties of the smooth approximation ϕ_ε to the true ℓ_0 penalty.

3. Theoretical Analysis

Building upon the generalized problem formulation in the preceding discussion, here, we establish our framework's core theoretical foundations. We prove the well-posedness of the sparse ergodic control problem with control-dependent diffusion, analyze the convergence properties of the smooth sparsity penalty, and characterize the structure of the optimal control policy.

3.1. Existence and Uniqueness of the Ergodic Pair

The first step in our analysis is to ensure that the generalized problem has a definite solution. For the controlled diffusion process governed by the stochastic dynamics in (1) under Assumption 1, together with the sparse ergodic cost functional introduced in Section 2.3, we seek an ergodic pair (λ, V) satisfying the HJB equation in terms of viscosity:

$$\lambda + \inf_{u \in U} \{\mathcal{L}^u V(x) + f_0(x) + \phi_\varepsilon(u)\} = 0, x \in \mathbb{R}^n \tag{12}$$

where λ is the optimal average cost and V is the relative value function, unique up to an additive constant. The generator $\mathcal{L}^u$ is now given by:

$$\mathcal{L}^u \psi(x) = \mathrm{b}(x,u) \cdot \nabla \psi(x) + \frac{1}{2} \mathrm{tr}\left(a(x,\, u) \nabla^2 \psi(x)\right),\ a\ (x,\ u)\ =\ \sigma(x,\ u)\sigma(x,u)^{\mathsf{T}} \tag{13}$$

The introduction of control-dependent diffusion creates two specific technical obstacles absent in existing sparse ergodic control formulations with control-independent diffusion.

(i) Non-separable Hamiltonian. When $\boldsymbol{a}(x,u)$ depends on u, the Hamiltonian

$$H_\varepsilon(x,p,X) = \inf_{u \in U} \left\{ b(x,u) \cdot p + \frac{1}{2} tr[a(x,u)X] + f_0(x) + \phi_\varepsilon(u) \right\} \tag{14}$$

becomes non-separable because the diffusion term $\frac{1}{2} tr[a(x,u)X]$ couples the control variable u with the second-order derivative $X = \nabla^2 V$. As a result, classical arguments relying on control-affine separability no longer apply directly.

(ii) Uniform Foster–Lyapunov bound. The vanishing-discount analysis requires the estimate

$$\sup_{u \in U} \mathcal{L}^{\sqcap} \mathcal{V}(x) \leq c_0 - c_1 \mathcal{V}(x) \tag{15}$$

to hold uniformly over all admissible controls. Unlike the control-independent setting, the diffusion coefficient $\sigma(x,u)$ now enters the generator directly and must therefore be controlled uniformly with respect to u.

Throughout this section, we operate under Assumptions 1–3.

Theorem 1 (Existence and Uniqueness of the Ergodic Pair)**.** *With Assumptions 1 (Dynamics Regularity), 2 (Foster–Lyapunov Stability), and 3 (Comparison Principle and HJB Regularity):*

1. *Existence: There exists a constant* $\lambda^* \in \mathbb{R}$ *and a function* $V \in C(\mathbb{R}^n)$ *with at most polynomial growth (i.e.,* $|V(x)| \leq C\ (1 + |x|^m)$ *for some constants* $C, m > 0$ *that is a viscosity solution of the ergodic HJB equation. The pair (;*λ^**,V) is obtained as the limit of the discounted problem via the vanishing discount approach:*

$$\lambda^* = \lim_{\alpha \downarrow 0} \alpha V_\alpha(x_{\text{ref}}),\ V(x) = \lim_{\alpha \downarrow 0} (V_\alpha(x) - V_\alpha(x_{\text{ref}})) \tag{16}$$

   where V_α *is the solution of the discounted HJB equation with discount factor* α*, and the convergence is locally uniform.*
2. *Uniqueness: The optimal average cost* λ^* *is unique. The relative value function V is unique up to an additive constant;*
3. *Existence of Optimal Control: There exists an optimal stationary Markov control policy* u^**:* $\mathbb{R}^n \to U$ *that achieves the infimum in Ergodic HJB equation for almost every x with respect to the invariant measure* π^* *associated with* u^**.*

Proof Sketch. We follow the vanishing discount methodology within the viscosity solution framework [21], adapted to the present setting with control-dependent diffusion and smooth sparsity penalties. The key distinction from the control-independent case is that the diffusion matrix $a(x, u)$ depends explicitly on the control input, making the ergodic Hamiltonian non-separable in the control and second-order terms.

Step 1 (Discounted Problem). For each discount factor $\alpha > 0$, consider the discounted HJB equation associated with the controlled diffusion process. Since $u \mapsto a(x, u)$ and $u \mapsto b(x, u)$ are continuous on the compact control set U under Assumption 1, the corresponding Hamiltonian remains continuous in (x, p, X), despite the coupling between u and the second-order diffusion term $\text{tr}[a(x, u)X]$. Together with the continuity and growth structure induced by the smoothed Hamiltonian, standard viscosity-solution arguments support the existence analysis of a continuous solution with polynomial growth.

Step 2 (Uniform A Priori Estimates). The Foster–Lyapunov condition in Assumption 2 provides uniform bounds on the discounted value functions and their local oscillations. In particular, the condition

$$\sup_{u \in U} \mathcal{L}^{\sqcap} \mathcal{V}(x) \leq c_0 - c_1 \mathcal{V}(x) \tag{17}$$

ensures that the required a priori estimates hold uniformly over all admissible controls, which is essential in the control-dependent diffusion setting. Consequently, the family of normalized solutions remains locally equibounded and equicontinuous.

Step 3 (Vanishing Discount Limit). Using the half-relaxed limit technique together with local compactness arguments, we may pass to the limit $\alpha \downarrow 0$ and obtain an ergodic pair $(\lambda,\ V)$ satisfying the ergodic HJB equation in the viscosity sense.

Step 4 (Uniqueness). Assumption 3 ensures that the comparison principle holds within the class of functions with Lyapunov-controlled growth. Consequently, the ergodic constant λ is unique and the relative value function V is unique up to an additive constant.

Step 5 (Optimal Control Existence). Since the Hamiltonian is continuous in u, including the diffusion term $\text{tr}[a(x, u)X]$, and the control set U is compact, standard measurable selection results imply the existence of a measurable minimizer. This measurable selector defines the optimal stationary Markov control policy. □

Remark on the uniform Foster–Lyapunov bound. Because the diffusion matrix $a(x, u)$ depends on u, the standard argument that the generator bound holds for each fixed control

and then passes to the supremum requires that $a(x,u)$ is uniformly bounded in operator norm over U. Under Assumption 1, $\sigma(x,u)$ is continuous on the compact set U, hence

$$\sup_{u\in U}|\mathrm{a}\,(x,u)|_{\mathrm{op}} \le C_a\left(1+|x|^2\right) \tag{18}$$

with a constant C_a independent of u. Combined with Assumption 2 (which uses $\sup_{u\in U}\mathcal{L}^{\sqcap}\mathcal{V}(x)$), the required uniform bound holds. A complete verification of the half-relaxed limit procedure under the non-separable Hamiltonian is provided in Supplementary Material S1, where we explicitly show that the coupling between u and $\nabla^2 V$ does not break the compactness argument because the second-order term enters linearly in the Hamiltonian and $a(x,u)$ is continuous in u.

Theorem 1 establishes the well-posedness of the ergodic pair under control-dependent diffusion. Additional technical details of the vanishing-discount argument and constant dependencies $C = C\,(U, \mathcal{V}, c_0, c_1)$ are provided in Supplementary Material S1.

3.2. Convergence of the Smooth Approximation

To handle the discontinuous ℓ_0 penalty in a computationally tractable way, we approximate it by the smooth family ϕ_ε [22,23] and analyze how the corresponding ergodic pair $(\lambda_\varepsilon, V_\varepsilon)$ converges to the ℓ_0-penalized pair (λ_0, V_0) as $\varepsilon \to 0$.

Theorem 2 (Convergence of the Smooth Approximation)**.** *Under the same assumptions as in Theorem 1, the following holds:*

1. *Pointwise Convergence of Value: The optimal average cost converges, $\lambda_\varepsilon \to \lambda_0$, as $\varepsilon\downarrow 0$. Furthermore, for a fixed reference point x_{ref}, the relative value functions converge locally uniformly:*
$$V_\varepsilon(\cdot) \to V_0(\cdot) \text{ locally uniformly as } \varepsilon \downarrow 0 \tag{19}$$
2. *Error Bound: For any compact set $K \subset R^n$, there exists a constant $C_K > 0$, depending on K but independent of ε, such that:*
$$\|V_\varepsilon - V_0\|_{L^\infty(K)} \le C_K \cdot \varepsilon \tag{20}$$

This result establishes quantitative consistency between the smoothed and non-smooth formulations, with the approximation error controlled linearly by ε. The constant C_K depends on the compact set K, the size and structure of the control set U, the penalty weights $\{a_i\}$, and the parameters of the Lyapunov function (c_0, c_1), but is independent of ε.

Proof Sketch.

1. Monotonicity and Stability: The family $\{\phi_\varepsilon\}$ is monotonically decreasing with respect to ε and converges pointwise to ϕ_0 *as* $\varepsilon\downarrow 0$. Consequently, the corresponding Hamiltonians $H_\varepsilon(x, p, \mathrm{X})$ also converge monotonically to $H\,(x, p, \mathrm{X})$;
2. Viscosity Solution Stability: The stability of viscosity solutions under monotone convergence of Hamiltonians is a classical result. Applying the half-relaxed limit technique ensures that any cluster point of $\{(\lambda_\varepsilon, V_\varepsilon)\}$ as $\varepsilon\to 0$ is a viscosity solution of the limiting ℓ_0 problem;
3. Uniqueness and Full Convergence: By the uniqueness of the ergodic pair for the limiting problem (Theorem 1), the entire family converges, and not just a subsequence;
4. Error Estimation: The quantitative approximation estimate is motivated by the consistency properties of the exponential smoothing operator and can be analyzed using viscosity-solution stability arguments together with a doubling-of-variables framework. Specifically, the error bound shows that the approximation error scales linearly

with the smoothing parameter, under the assumptions stated in Section 3.1. This analysis confirms that ϕ_ε serves as a computationally tractable surrogate for the ℓ_0 penalty, with approximation error of order $O(\varepsilon)$ on any compact set K. Detailed derivation is provided in Supplementary Material S2. □

3.3. Sparse Structure in Non-Affine Systems

A key insight from sparse control in affine systems is the "bang-off-bang" nature of the optimal policy. We now generalize this structural result to a broader class of non-affine systems. To make the derivation of the threshold scale explicit, we first introduce a local linearization assumption on the dynamics.

Assumption 4 (Local Linearization and Remainder Control)**.** *For each $x \in \mathbb{R}^n$, the functions $b(x,u)$ and $\sigma(x,u)$ admit the local expansions*

$$b(x,u) = b_0(x) + B(x)u + r_b(x,u), \sigma(x,u) = \sigma_0(x) + S(x)u + r_\sigma(x,u) \tag{21}$$

where $r_b(x,u)$ and $r_\sigma(x,u)$ satisfy

$$|r_b(x,u)| + |r_\sigma(x,u)| \leq C_r|u|^2 \text{for } |u| \leq r_0 \text{with some } r_0 > 0 \tag{22}$$

where $C_r > 0$ is a local remainder constant. The control domain $U \subset \mathbb{R}^m$ is compact, and the expansion holds component wise when u is vector-valued.

Remark 1. *This local linearization assumption is particularly relevant for sparse control regimes where optimal controls frequently reside near zero due to the sparsity penalty. For systems requiring large control inputs, global linear approximations or piecewise-linear models may be more appropriate. In our numerical experiments, the learned value functions exhibit smooth behavior near the origin of the control space, providing indirect support for this assumption. A formal residual analysis of r_b and r_σ is left for future work.*

For a given state x, the optimal control $u^*(x)$ minimizes the Hamiltonian:

$$\mathcal{H}(x,u) = b(x,u) \cdot \nabla V(x) + \frac{1}{2}\text{tr}\left(a(x,u)\nabla^2 V(x)\right) + f_0(x) + \phi_\epsilon(u) \tag{23}$$

Even when b and σ are not affine in u, the separable structure of $\phi_\epsilon(u) = \sum_i a_i\left(1 - e^{-|u_i|/\varepsilon}\right)$ imposes a coordinate-wise sparsity-promoting effect.

Definition 1 (Generalized Switching Function)**.** *For a component i of the control vector, we define the generalized switching function $g_i : \mathbb{R}^n \to \mathbb{R}$ as the partial derivative of the Hamiltonian's drift and diffusion terms with respect to u_i, evaluated at the optimal value function. At points where V is differentiable,*

$$g_i = \left\langle \nabla V(x), \frac{\partial b}{\partial u_i}(x,u^*) \right\rangle + \frac{1}{2}\text{tr}\left(\nabla^2 V(x)\frac{\partial a}{\partial u_i}(x,u^*)\right) \tag{24}$$

At non-differentiable points, this interpretation holds for any super-/sub-gradient $p \in J^{2,\pm}V(x)$.

Proposition 1 (Quasi-Threshold Structure under Local Linearization)**.** *Let $u_i^{(0)*}(x)$ denote the i-th component of an optimal stationary Markov policy associated with the smoothed penalty ϕ_ε. Under Assumption 4 and assuming that $b(x,u)$ and $\sigma(x,u)$ are continuously differentiable with respect to u, the optimal control satisfies the following asymptotic structural properties as $\varepsilon \to 0$.*

The characterization is local in nature and relies on the first-order approximation around sparse-control regimes.

(i) The optimal control exhibits a soft-thresholding behavior induced by the smooth sparsity penalty. In particular, there exists a state-dependent threshold scale $\tau_i(\varepsilon) > 0$ *such that*

$$g_i(x;\varepsilon) \lesssim \tau_i(\varepsilon) \Rightarrow u_i^*(x) \approx 0 \tag{25}$$

(ii) For the exponential smoother

$$\phi_\varepsilon(u) = \sum_i a_i \left(1 - e^{-|u_i|/\varepsilon}\right) \tag{26}$$

the effective threshold scale satisfies

$$\tau_i(\varepsilon) = O(a_i/\varepsilon) \tag{27}$$

This scaling arises from the balance between the switching term $g_i(x;\varepsilon)$*and the gradient of the smoothed sparsity penalty.*

(iii) As $\varepsilon \to 0$*, the resulting soft-thresholding behavior converges to a bang-off-bang type policy:*

$$u_i^{(0)*}(x) = 0 \text{ or } u_i^{(0)*}(x) \in U_i \setminus \{0\} \tag{28}$$

This bang-off-bang structure arises from the ℓ_0*-type sparsity regularization, which suppresses controls whose marginal Hamiltonian contribution is below the activation threshold. Under Assumption 4, the convergence rate toward the limiting sparse policy is at least O (ε). The derivation in Supplementary Material S3 proceeds by analyzing the first-order optimality condition of the smoothed Hamiltonian, deriving an implicit activation relation for each control component, and identifying the resulting quasi-threshold activation boundary.*

Remark 2 (Boundary-constrained optima)**.** *Proposition 1 characterizes interior optima. When* $|g_i(x;\varepsilon)|$ *exceeds a boundary saturation threshold* $\tau_i^{\text{sat}} = O\left(a_i/\varepsilon + L_b u_i^{\max}\right)$*, the minimizer saturates at* $u_i^* = \pm u_i^{\max}$*. The full policy therefore takes a three-region form: off (*$u_i^* = 0$*), interior, or bang (*$u_i^* = u_i^{\max}$*), consistent with a generalized bang-off-bang structure. Analysis of the strongly nonlinear regime where Assumption 4 fails globally is left for future work.*

Remark 3 (Threshold accuracy under linearization remainder)**.** *Under Assumption 4, the remainder satisfies* $|r_b| + |r_\sigma| \le C_r|u|^2$*. Propagating this through the first-order optimality condition (see Supplementary Material S3) shows that the activation boundary is perturbed by* $O(C_r r_0)$ *relative to* $\tau_i(\varepsilon)$*. The overall threshold accuracy is therefore* $O(\varepsilon + C_r r_0)$*, with the* $O(\varepsilon)$ *term dominant in the sparse-control regime where* $C_r r_0 \ll a_i$*.*

The quasi-threshold structure established by Proposition 1 is depicted in Figure 2, which illustrates how the smooth sparsity penalty induces a continuous but increasingly sharp sparse control profile as ε decreases. For finite ε, the resulting feedback law remains continuous and exhibits smooth threshold-like transitions between inactive and high-activation regions. As $\varepsilon \to 0$, the induced policy asymptotically approaches bang-off-bang-type switching behavior in the non-affine setting. The figure illustrates: (i) the transition from soft to hard thresholding as $\varepsilon \to 0$; (ii) how ε controls the threshold sharpness near the origin; (iii) the state-dependent switching between off and bang regions in a synthetic 1D system; and (iv) the theoretical scaling $\tau_i(\varepsilon) \sim a_i/\varepsilon$.

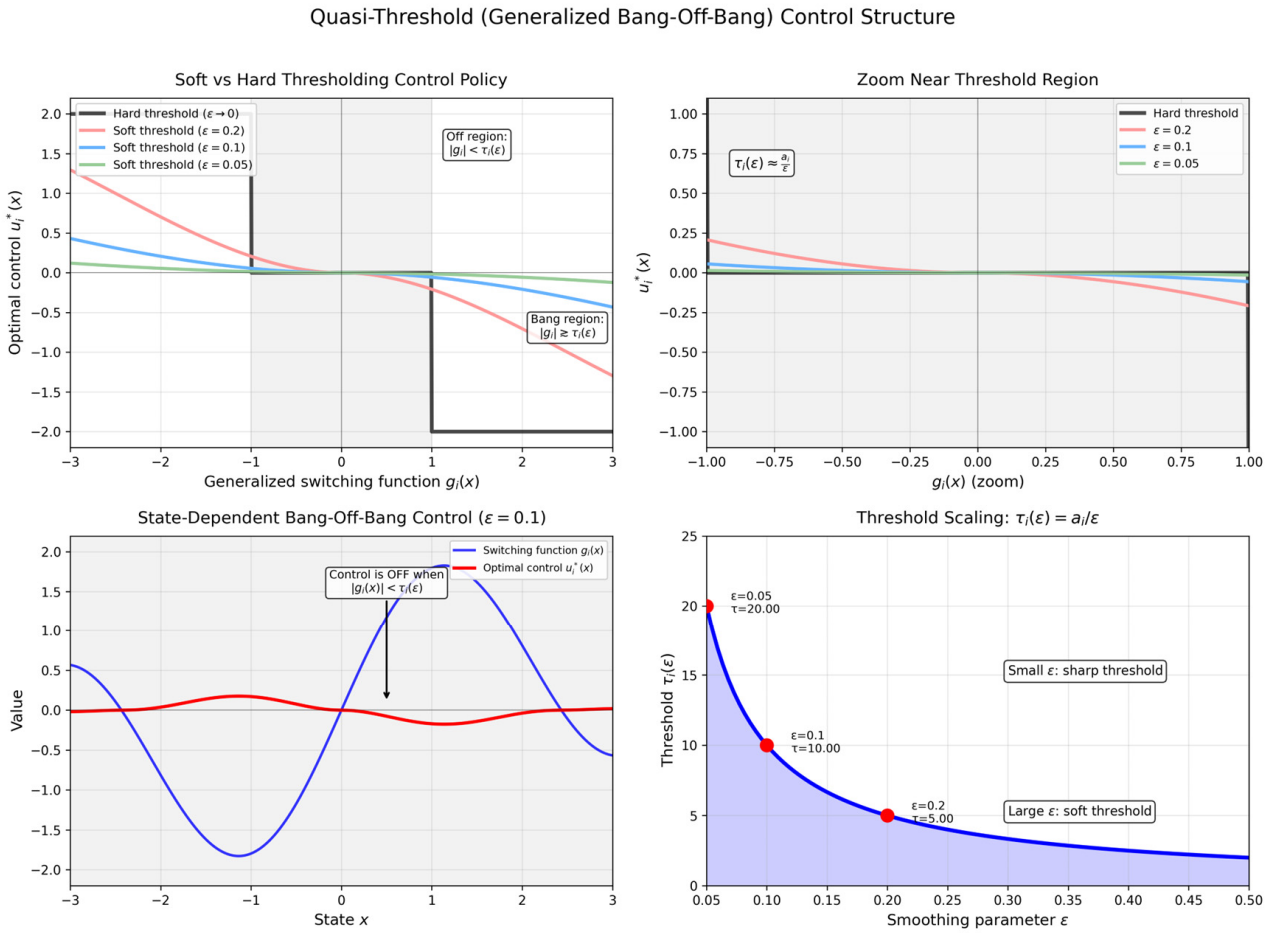


Figure 2. Illustration of the quasi-threshold (generalized bang-off-bang) control structure.

Discussion. This result extends classical threshold-type sparsity principles to non-affine systems with control-dependent diffusion. In non-affine systems, the "switching function" $g_i(x)$ is not simply a linear function of the co-state but involves more complex interactions due to the control-dependent drift and diffusion. The added Assumption 4 clarifies when the linearization-based derivation of the threshold $O(a_i/\varepsilon)$ is valid and provides explicit smallness conditions for the second-order remainder.

However, the fundamental trade-off remains: a control component is activated (i.e., becomes non-zero) only when its marginal benefit, quantified by $|g_i(x)|$, exceeds a certain threshold that incorporates the sparsity cost a_i. In vector-valued control settings, this quasi-threshold applies componentwise due to the separability of $\phi_\varepsilon(u)$.

Note: The complete derivation can be found in Supplementary Materials S3.

4. Computational Frameworks: Learning and Distributed Algorithms

Section 3 establishes existence and quasi-threshold structure of the optimal sparse feedback law. Computing this solution directly is impractical for high-dimensional systems due to the exponential growth of grid-based state discretization and the non-smooth Hamiltonian. This section addresses these two computational obstacles: Section 4.1 introduces a PINN-based solver with adaptive residual sampling for high-dimensional sparse ergodic HJB equations, and Section 4.2 develops a distributed monotone iterative scheme for weakly coupled multi-agent systems.

4.1. PINN for Ergodic HJB

By Theorem 2, solving the smoothed problem with $\varepsilon = 0.01$ introduces an approximation error of at most $O(10^{-2})$ in the value function, which is acceptable for our control objectives. Let $V_\theta(x)$ denote a deep neural network (DNN) parameterized by weights θ, which takes the state x as input and outputs a scalar value function [24]. The Hamiltonian H is a function of x, $\nabla_x V$, and $\nabla_x^2 V$. These derivatives are computed efficiently using automatic differentiation.

The ergodic HJB equation and the normalization condition at a reference point x_{ref} constitute the physical constraints for our PINN. We define the training loss $\mathcal{L}(\theta, \lambda)$ as a weighted sum of the PDE residual and the boundary condition:

$$\mathcal{L}(\theta,\lambda) = \frac{1}{N_c}\sum_{i=1}^{N_c}\left[\lambda + H\left(x_c^{(i)}, \nabla V_\theta\left(x_c^{(i)}\right), \nabla^2 V_\theta\left(x_c^{(i)}\right)\right)\right]^2 + \omega[V_\theta(x_{\text{ref}})]^2 \tag{29}$$

Here, $\left\{x_c^{(i)}\right\}_{i=1}^{N_c}$ denotes the collocation points sampled from the state domain, $\omega > 0$ is a penalty parameter enforcing the normalization condition $V_\theta(x_{\text{ref}}) = 0$, and the PDE residual term is weighted ten times more heavily than the normalization penalty term. The Hamiltonian is defined by

$$H(x,\ p,\ X) = \min_{u\in U}\left\{b(x,u)\cdot p + \frac{1}{2}\text{tr}\left(a\ (x,\ u)\ X\right) + f_0(x) + \phi_\varepsilon(u)\right\} \tag{30}$$

where $p = \nabla V_\theta(x)$ and $X = \nabla^2 V_\theta(x)$. In the PINN implementation, these quantities are computed via automatic differentiation.

The minimization is performed over both the network parameters θ and the scalar λ:

$$(\theta^*,\ \lambda^*) = \underset{\theta,\lambda}{\text{argmin}}\,\ell(\theta,\lambda) \tag{31}$$

Normalization and joint optimization stability. The ergodic HJB equation is invariant under $V \mapsto V + c$, making the joint optimization over (θ, λ) ill-posed without fixing this degree of freedom. The normalization penalty $\omega V_\theta(x_{\text{ref}})^2$ enforces $V_\theta(x_{\text{ref}}) = 0$, selecting a unique representative and eliminating the degenerate flat direction in the loss landscape. The scalar λ is initialized as $\lambda^{(0)} = \min_{x\in\mathcal{D}} f_0(x)$, a valid lower bound on the ergodic cost. In practice, λ converges within the first 10,000 iterations and its standard deviation across 10 runs is below 0.73% of the mean (Section 5.5.2), confirming stability of the joint optimization. The weight $\omega = 0.1$ (10:1 PDE-to-normalization ratio) was selected by a sweep over $\{10^{-3}, 10^{-2}, 10^{-1}, 1\}$ and found to be consistently stable across both experimental settings.

The PINN architecture and training workflow are depicted in Figure 3. The neural network $V_\theta(x)$ takes the state x as input and produces the value-function estimate through a series of hidden and fully connected layers. Automatic differentiation is employed to compute the necessary gradients $\nabla V_\theta(x)$ and Hessian $\nabla^2 V_\theta(x)$ for evaluating the Hamiltonian in the PDE residual. The training loss $\mathcal{L}(\theta, \lambda)$ combines the PDE residual with a normalization condition, while adaptive sampling dynamically focuses on computational resources in high-error regions of the state space.

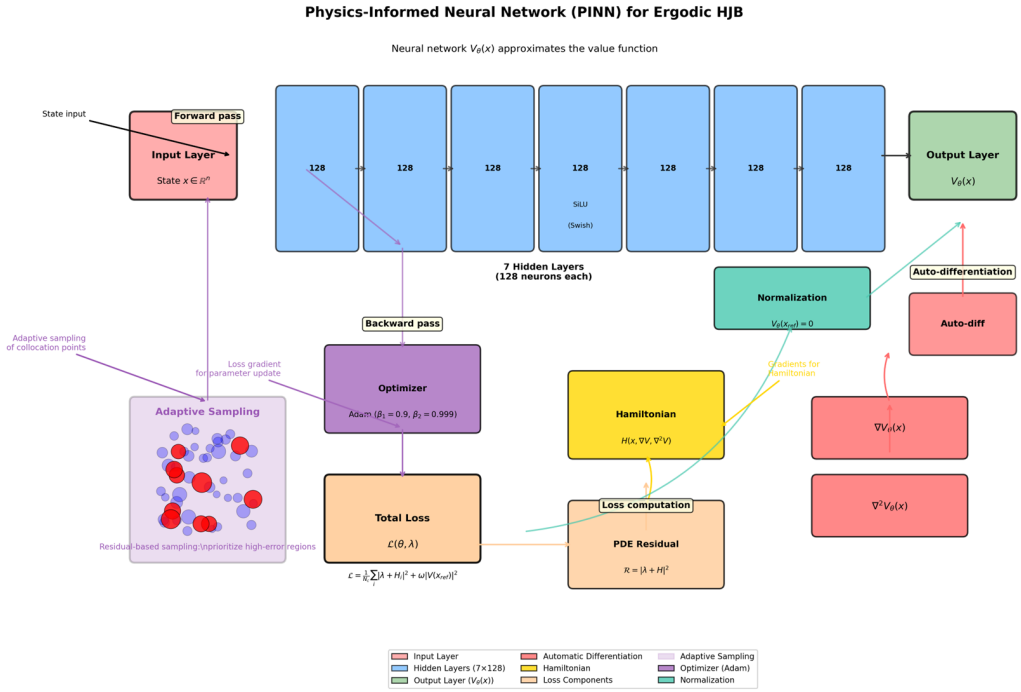


Figure 3. Architecture of the PINN.

4.1.1. Training Pipeline and Adaptive Sampling

We use a fully connected feed-forward network with 7 hidden layers of 128 neurons each and SiLU activation; batch normalization is not used, and the gradient clipping (global norm) threshold is 1.0. Training employs Adam ($\beta_1 = 0.9$, $\beta_2 = 0.999$) with cosine annealing from an initial learning rate (see Table 2 for the range) down to 1×10^{-5}. Each iteration uses a mini-batch of 4096 PDE collocation points plus 512 boundary/normalization points, for a total of 200,000 iterations; Early stopping uses a patience of 20,000 iterations. This value was selected empirically during preliminary development and was kept fixed across all experiments. Validation was performed on a fixed held-out validation set sampled once at initialization and kept unchanged throughout training. A gradient penalty of $10^{-5}\|\nabla_x V_\theta\|_2^2$ is added to the loss, and the PDE vs. normalization loss weight ratio is 10 : 1. Lyapunov-weighted sampling draws collocation points with density proportional to the Lyapunov function $\mathcal{V}(x)$ (Assumption 2). All experiments ran on NVIDIA RTX 5000 Ada GPUs (16 GB) with 10 independent runs (random seeds $\{101, 202, \ldots, 1001\}$); typical training times are 4–10 h per task. Detailed hyperparameters are summarized in Table 2.

Table 2. PINN training hyperparameters (used in main experiments).

Hyperparameter	Value
Network depth/width	7×128
Activation function	SiLU
Weight init/Bias init	Xavier uniform/0
Optimizer	Adam ($\beta_1 = 0.9$, $\beta_2 = 0.999$)
Learning rate (LR)	1×10^{-4} to 5×10^{-4}
Batch size (PDE/boundary)	4096/512 (main experiments); 512–2048 (ablation sweeps)
Training iterations	200,000
Smoothing parameter ε	0.01 (main experiments); 5×10^{-2} to 5×10^{-3}(ablation sweep)
Gradient penalty coefficient	10^{-5}
Random seeds (for 10 runs)	{101, 202, . . ., 1001}
GPU	NVIDIA RTX 5000 Ada (16 GB)
Typical runtime	4–10 h

Initial Domain Selection. Collocation points are initially sampled from a bounded hyper-rectangle (physical constraints expanded by 20%) determined by prior system knowledge.

We employ a residual-based adaptive sampling strategy during training. The detailed procedure and its motivation are formalized in Algorithm 1 below.

Adaptive Sampling Algorithm: We employ a residual-based adaptive sampling strategy with the following procedure:

Algorithm 1. Adaptive Residual Sampling

Input: Current PINN parameters θ, candidate pool size N_pool = 50,000, batch size N_batch = 4096, replacement ratio r_s = 0.3
Output: Updated collocation batch
1. Generate a large candidate pool {x_i} from the domain using Latin Hypercube Sampling
2. Compute PDE residuals R(x_i) = |λ + H(x_i, ∇V_θ(x_i), ∇^2V_θ(x_i))| for all candidates
3. Sort candidates by residual in descending order
4. Select top r_s·N_batch points with highest residuals as "difficult samples"
5. Randomly select (1 − r_s)·N_batch points from remaining candidates
6. Combine both sets to form the new training batch
7. Repeat every K = 2000 training iterations

This strategy focuses computational resources on high-error regions. In our experiments, adaptive sampling reduced required training iterations by ≈30% compared to uniform sampling.

4.1.2. Inner Hamiltonian Minimization

In the non-affine control-dependent diffusion setting, the inner minimization $\min_{u \in U}\left\{b(x,u)\cdot p + \frac{1}{2}\mathrm{tr}[a(x,u)X] + \phi_\varepsilon(u)\right\}$ has no closed-form solution. We use a two-stage procedure: (i) coarse grid search over $N_u = 20$ uniform candidates per control dimension, followed by (ii) L-BFGS-B refinement initialized at the coarse minimizer, with control gradients computed analytically. This inner optimization accounts for approximately 35–45% of per-iteration wall-clock time, and is reflected in the training times of Table 2. The L-BFGS-B termination tolerance is set to $|\nabla_u H| < 10^{-4}$, contributing an inner-loop error below 10^{-4}–smaller than the target PDE residual of $O(10^{-2})$. For control dimensions $m > 5$, a jointly trained control network $\pi_\varnothing(x)$ would be more practical; this extension is left for future work.

4.1.3. Adaptive Sampling: Motivation and Efficacy

The PDE residual can vary substantially across the state space, particularly near boundaries, unstable equilibria, and regions where the optimal control switches abruptly (as characterized by the quasi-threshold structure in Proposition 1). Uniform sampling allocates equal computational effort across these regions, potentially oversampling low-error areas while under-sampling regions with large approximation error. Algorithm 1 addresses this imbalance by preferentially selecting collocation points with larger PDE residuals, thereby concentrating training in regions where the HJB approximation error is highest. The ablation study in Section 5.3 shows that this adaptive strategy reduces the required training iterations by approximately 30% relative to uniform sampling.

4.2. Distributed Monotone Iterative Scheme for Multi-Agent Systems

In many modern applications, such as swarm robotics or sensor networks, the overall system is composed of multiple interacting agents. A centralized solution to the global HJB equation is often infeasible. We propose a distributed algorithm where each agent computes a local policy based on information from its neighbors.

In our formulation, inter-robot interactions are represented through relative-position terms and neighborhood averaging, both of which are globally Lipschitz with agent-independent constants. Furthermore, the coupling strength does not grow with the number of agents because each robot interacts only with its immediate neighbors in the proximity graph. These properties ensure that the effective coupling coefficients remain within the admissible range required by Proposition 2, aligning the experimental setting to the theoretical assumptions. Pseudocodes for Algorithm 2 is shown below:

Algorithm 2. Distributed Sparse Ergodic Control

Input: Local dynamics bi, local state cost fi0, sparsity penalty φ_iˆε, neighbor graph G, tolerance tol > 0, maximum iterations Kmax
Output: Distributed optimal controls u_iˆ*(x)
Initialize local value-function approximations V(0)i
for k = 0, 1, 2, . . ., Kmax do
  for each agent i in parallel do
    Solve the local HJB Equation (34) using a local PINN to obtain V(k + 1)i
    Update u(k + 1)i = argminu ∈ Ui Hi(x, u, V(k + 1)i)
    Exchange V(k + 1)i with neighbors j ∈ N(i)
    Update the local coupling estimate gi, k + 1 using received information
end for
  if maxi ‖V(k + 1)i − V(k)i‖∞ < tol then
  break
  end if
end for

To reduce communication overheads, each agent does not transmit its full value-function network $\hat{V}^{(i)}$ or high-dimensional gradient fields. Before sending information to neighbors, the agent projects its locally learned value function onto a predefined low-dimensional basis $\{\phi_m(x_i)\}_{m=1}^{r}$ (with $r = 10$ in our experiments). Specifically, each agent computes a low-dimensional coefficient vector:

$$c^{(i)} = \left(\left\langle \hat{V}^{(i)}, \phi_1(x_i)\right\rangle, \ldots, \left\langle \hat{V}^{(i)}, \phi_r(x_i)\right\rangle\right) \tag{32}$$

where $\langle \cdot, \cdot \rangle$ denotes the empirical inner product computed over sampled local states. The basis functions are constructed from the final hidden-layer feature activations of the local PINN or from their PCA-reduced representations. Only these r-dimensional coefficient vectors (rather than full neural-network parameters) are exchanged among neighboring agents. This low-rank encoding keeps the communication cost per iteration constant (O(r) scalars), independent of the PINN size, state dimension, or number of agents, while still preserving sufficient information for accurate consensus and coupling-term estimation. Each agent transmits its coefficient vector to a neighbor only if its change since the last transmission exceeds a predefined triggering threshold $\tau_{trig} > 0$, i.e., $\left\| c_i^{(k)} - c_i^{(k_{\text{last}})} \right\| > \tau_{trig}$, where k_{last} is the iteration of the last transmission. This event-triggered criterion ensures reduced communication while maintaining convergence. This pseudocode clarifies the communication pattern and the fixed-point nature of the monotone update.

4.2.1. Problem Decomposition

Consider a system of N agents with coupled dynamics. The state of the entire system is $X = (x^1, \ldots, x^N)$, and the control is $U = (u^1, \ldots, u^N)$. We assume the dynamics and cost are separable in the following:

$$b(X, U) = \begin{pmatrix} b^1(x^1, u^1) + g^1(X) \\ \vdots \\ b^N(x^N, u^N) + g^N(X) \end{pmatrix}, f_0(X) = \sum_{i=1}^{N} f_0^i\left(x^i\right) \tag{33}$$

where $g^i(X)$ represents the coupling from other agents. The communication topology is described by a graph $\mathcal{G}$, where agent i can receive information from its neighbors $\mathcal{N}_i$.

4.2.2. Local HJB Approximation and Consensus

The global ergodic HJB is intractable. Instead, each agent i maintains a local value function approximation $\hat{V}^i(x^i)$ of its component. The key idea is to approximate the coupling term $g^i(X)$ using the states and value functions of its neighbors.

We propose the following distributed monotone scheme [25], inspired by Jacobi iteration and consensus algorithms:

1. Initialization: Each agent i initializes its value function estimate $\hat{V}^{i,(0)}$;
2. Local Hamiltonian Evaluation: At iteration k, each agent i computes a local Hamiltonian using its current value function and information from its neighbors. For a state x^i, this is:

$$\begin{aligned} H^{i,(k)}(x^i) = \min_{u^i \in U^i} \Big\{ \left(b^i(x^i, u^i) + \hat{g}^{i,(k)}\right) \cdot \nabla \hat{V}^{i,(k)}(x^i) + \tfrac{1}{2}\mathrm{tr}\left(a^i \nabla^2 \hat{V}^{i,(k)}(x^i)\right) \\ + f_0^i(x^i) + \phi_\varepsilon^i(u^i) \Big\} \end{aligned} \tag{34}$$

   where $\hat{g}^{i,(k)}$ is an estimate of the coupling term, e.g., $\hat{g}^{i,(k)} = \frac{1}{|\mathcal{N}_i|}\sum_{j \in \mathcal{N}_i} \gamma^{ij}(x^j)$, and x^j are the current states or expected states of neighbors.
3. Value Update with Consensus: Agents update their value function estimates by combining the local HJB solution with a consensus term from their neighbors to ensure coordination:

$$\hat{V}^{i,(k+1)}\left(x^i\right) = (1-\delta)\hat{V}^{i,(k)}\left(x^i\right) - \Delta t \cdot H^{i,(k)}\left(x^i\right) + \frac{\delta}{|\mathcal{N}_i|}\sum_{j \in \mathcal{N}_i} \hat{V}^{j,(k)}(x^j) \tag{35}$$

   Here, Δt is a pseudo-time step, and $\delta \in (0, 1)$ is a consensus weight.
   Policy Update: Each agent updates its local policy using its new value function:

$$\begin{aligned} u^{i,(k+1)}(x^i) = \underset{u^i \in U^i}{\mathrm{argmin}} \Big\{ \left(b^i(x^i, u^i) + \hat{g}^{i,(k)}\right) \cdot \nabla \hat{V}^{i,(k+1)}(x^i) \\ + \tfrac{1}{2}\mathrm{tr}\left(a^i(x^i, u^i)\nabla^2 \hat{V}^{i,(k+1)}(x^i)\right) + \phi_\varepsilon^i(u^i) \Big\} \end{aligned} \tag{36}$$

Since the local state cost $f_0^i(x^i)$ does not depend on the control variable u^i, it is omitted from the policy-update minimization.

4.2.3. Monotonicity and Convergence

The scheme is designed to be monotone and stable under verifiable parameter constraints. The monotone-inspired update structure is intended to approximately preserve the order-consistency properties associated with the underlying PDE. The following proposition summarizes sufficient conditions under which the distributed iteration exhibits stable approximate convergence behavior under weak inter-agent coupling. The analysis is based

on a local contraction-style argument and is further supported by the numerical results reported in Section 5.

Proposition 2 (Sufficient Conditions for Stable Approximate Convergence)**.** *Under the assumptions of a connected communication graph $\mathcal{G}$ and weak inter-agent coupling, the distributed monotone iterative scheme exhibits stable approximate convergence toward a ρ-neighborhood of the global optimal solution provided that the following sufficient conditions hold:*

1. *Coupling Strength Bound;*
   The inter-agent coupling coefficients satisfy

$$\max_{i,\,j}\left|\gamma^{ij}\right| \leq \overline{\gamma}, \text{where } \overline{\gamma} < \frac{c_1}{2L_V} \tag{37}$$

   with c_1 from the Foster–Lyapunov condition (Assumption 2) and L_V being an estimate of the value-function gradient Lipschitz constant.
2. *Parameter Constraints:*
   The pseudo-time step and consensus weight satisfy:

$$\Delta t \leq \min\left\{\frac{1}{L_H}, \frac{2\delta}{\Lambda_{\max}(\mathcal{G})}\right\}, \delta \in \left(0, \frac{1}{1+\overline{\gamma}L_V}\right) \tag{38}$$

   where L_H is the Lipschitz constant of the Hamiltonian and $\Lambda_{max}(\mathcal{G})$ is the largest eigenvalue of the graph Laplacian. In practice, the Lipschitz constant L_H of the Hamiltonian can be bounded using the Lipschitz constants of b, σ, f_0, and ϕ_ε specified in Assumption 1. The value-function gradient Lipschitz constant L_V can be estimated numerically from the trained PINN solution V_θ via $L_V \approx \max_{x\in\mathcal{D}} \left\|\nabla_x^2 V_\theta(x)\right\|_{op}$, where D is the training domain. Conservative choices of Δt $\left(e.g.,\ 10^{-3}\right)$ typically satisfy above condition in practice.
3. *Convergence Radius;*
   The convergence neighborhood radius satisfies;

$$\rho = O(\overline{\gamma}\Lambda_{\max}(\mathcal{G}) + (1-\delta)\Delta t), \tag{39}$$

   where $\Lambda_{max}(\mathcal{G})$ denotes the largest eigenvalue of the graph Laplacian. The convergence neighborhood therefore depends on the communication topology through the spectral properties of the interaction graph rather than the graph diameter.

Proposition 2 establishes sufficient conditions for the convergence of our distributed monotone scheme. The practical implementation of this scheme, including the communication patterns and agent-internal computations, is illustrated in Figure 4. Each agent operates autonomously, maintaining a local PINN approximation of its value function while exchanging compressed representations with neighboring agents to achieve global coordination. The low-rank projection mechanism ensures that communication overheads remain constant ($O(r)$ scalars per iteration) regardless of network size or state dimension, making the scheme scalable for large multi-agent systems.

Top of Figure 4 (Communication Network): Communication network with 6 agents in a ring topology; Bottom of Figure 4 (Agentic Structure and Process): Internal structure and iterative process of a single agent. The scheme reduces communication overheads by exchanging only low-rank coefficient vectors (r = 10) instead of full value functions. The iterative process continues until convergence criteria are met.

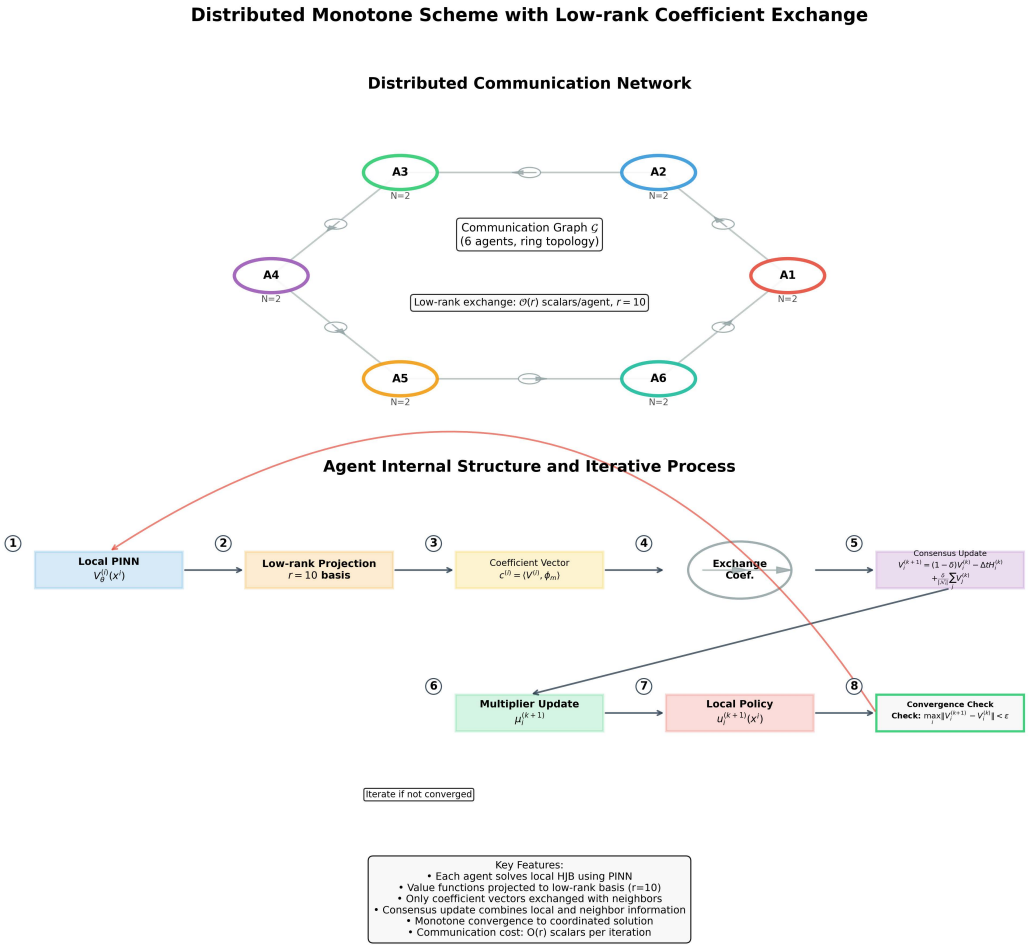


Figure 4. Distributed monotone scheme with low-rank coefficient exchange.

Practical Implementation Notes:

While the theoretical bound $\hat{L}_V$ may be challenging for analytical computation, in practice, we employ the following approaches:

- Estimate $\hat{L}_V$ empirically from the PINN-computed value-function gradients during training;
- Treat the conditions as tuning guidelines and verify convergence experimentally through the monotone decrease in the distributed residual.

In our experiments (Section 5.1), we found that conservative choices of $\delta = 0.2$–0.3 and $\Delta t = 0.1$–0.2 typically ensured stable convergence across various network configurations.

The proof, provided in Supplementary Material S4, analyzes the update operator as a near-contraction under the weak-coupling and step-size conditions of Proposition 2.

5. Numerical Experiments and Applications

The theoretical and computational frameworks developed in the preceding sections are now validated through comprehensive numerical experiments. This section presents the efficacy of our proposed methods in two distinct application domains: multi-robot swarm navigation and smart grid generator scheduling.

5.1. Sparse Coordination in Multi-Robot Swarm Navigation

Although the convergence guarantee of Proposition 2 (Section 4.2.3) requires a weak inter-agent coupling condition, the swarm-robot system used in our experiments is specifically designed to satisfy this condition, as detailed in the problem setup below. We validate

our distributed sparse control framework on a multi-robot formation control task [26,27], as illustrated in Figure 5. The objective is to coordinate a swarm of N autonomous robots to achieve a target triangular formation while minimizing control effort and communication overheads under stochastic disturbances.

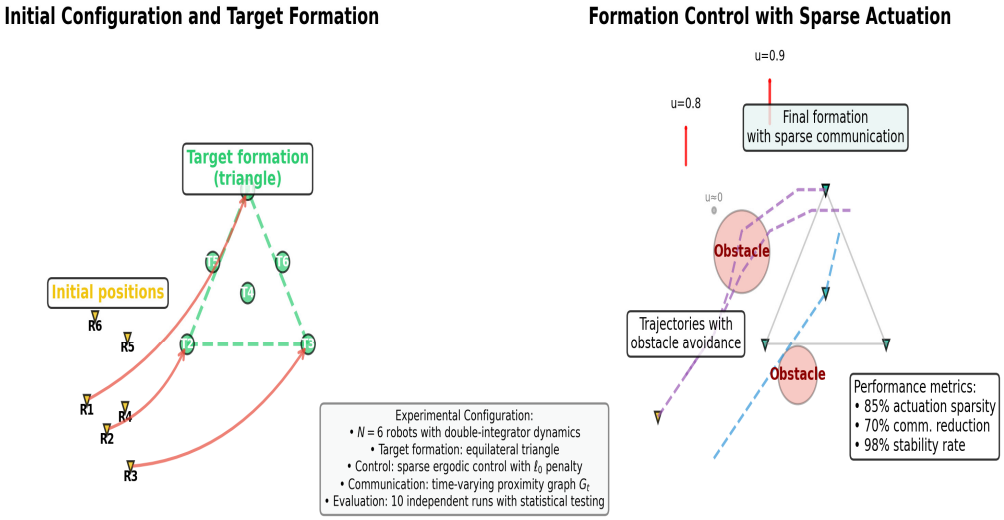


Figure 5. Multi-robot formation control task used in the experiments.

Left of Figure 5 (Target Formation): Initial configuration and target triangular formation. Robots start from scattered positions and must converge to a predefined triangular formation; Right of Figure 5 (Formation Control): Formation control process demonstrates sparse actuation and obstacle avoidance. The distributed controller achieves high sparsity (85% of control inputs are near zero) while reducing communication load by 70%. Performance metrics evidence the effectiveness of the proposed approach.

5.1.1. Experimental Setup

- Robot Dynamics: Each robot i follows double-integrator dynamics with stochastic disturbances:

$$d\begin{bmatrix} x_t^i \\ v_t^i \end{bmatrix} = \begin{bmatrix} v_t^i \\ 0 \end{bmatrix} dt + \begin{bmatrix} 0 \\ I \end{bmatrix} u_t^i dt + \begin{bmatrix} 0 & 0 \\ 0 & \sigma_v(u_t^i) I \end{bmatrix} dW_t^i \tag{40}$$

  Here, x_t^i and v_t^i are the position and velocity of robot i. To incorporate control-dependent diffusion, the velocity noise intensity is modulated by the control magnitude:

$$\sigma_v\left(u_t^i\right) = \sigma_0\left(1 + k\left|u_t^i\right|\right) \tag{41}$$

  with $\sigma_0 = 0.1$ and $k = 0.2$, consistent with actuator-noise models reported for wheeled mobile robots under velocity control [15,17,27]. This yields a control-dependent diffusion matrix

$$\sigma\left(x_t^i, u_t^i\right) = \begin{pmatrix} 0_{d\times d} & 0_{d\times d} \\ 0_{d\times d} & \sigma_v\left(u_t^i\right) I_d \end{pmatrix} \in \mathbb{R}^{2d\times 2d} \tag{42}$$

  where d is the spatial dimension, consistent with the generalized setting of Assumption 1.
- Sparse Cost Functional: The ergodic cost for the swarm is defined as:

$$J(u) = \limsup_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\int_0^T \left(\sum_{i=1}^N \left\|x_t^i - x_{\text{desire}}^i\right\|^2 + w_s \sum_{i=1}^N \phi_\varepsilon^i\left(u_t^i\right)\right) \mathrm{dt}\right] \tag{43}$$

where x_t^i denotes the position component of the robot state, and $w_s > 0$ is the sparsity penalty weight. The first term penalizes deviation from the desired formation, and the second term, with weight w_s, promotes sparsity in the control inputs using the smooth penalty ϕ_ε defined in Section 2.3.

- Communication Graph: Robots communicate over a time-varying proximity graph $\mathcal{G}_t$. A robot j is a neighbor of robot i if $\left\| x_t^i - x_t^j \right\| < R$, where R is set to 1.5 times the desired inter-agent spacing, following the proximity-graph convention in [27];
- Evaluation protocol: Each scenario is executed for 10 independent runs using different random seeds and noise realizations. For each run we record: (i) sparsity ratio; (ii) average formation error; (iii) communication load; and (iv) failure rate. These 10-run evaluations provide the primary metrics summarized in Table 3 and support statistical testing ($\alpha = 0.05$).

Table 3. Quantitative comparison of controller performance (mean ± s.d.).

Metric	Method	Value	Statistical Significance
Sparsity Ratio	Our Method	84.7% ± 2.3%	vs. Heuristic: $p = 0.004$
	LQR	0.2% ± 0.1%	
	Heuristic	61.5% ± 5.8%	
Formation Error	Our Method	1.05 ± 0.08 (normalized units)	vs. LQR: $p = 0.12$; vs. Heuristic: $p < 0.01$
	LQR	0.99 ± 0.05 (normalized units)	
	Heuristic	1.62 ± 0.21 (normalized units)	
Communication Load Reduction	Our Method	70% ± 6% reduction (vs. full-state sharing baseline)	$p < 0.001$
Stability	Our Method	98%	
	LQR	100%	
	Heuristic	65%	

Note: The LQR baseline primarily serves as a reference for trajectory stabilization performance rather than sparsity optimization; a direct sparsity comparison is not its intended purpose.

In addition, an investigation of robustness under communication imperfections (Table 4) employs 100 Monte-Carlo simulations per scenario. The purpose of this evaluation is to quantify robustness against stochastic disturbances, packet loss, and communication delays which are separate from the 10-run performance comparison.

Table 4. Robustness under Communication Imperfections (Mean ± S.D.).

Scenario	Stability Rate (%)	Formation Error Increase (%)	Sparsity Reduction (%)
Baseline (no faults)	98.0 ± 1.5	0.0	0.0
10% Packet Loss	96.2 ± 2.1	3.5 ± 1.2	2.1 ± 0.8
30% Packet Loss	91.5 ± 3.3	8.7 ± 2.4	6.3 ± 1.9
50% Packet Loss	85.3 ± 4.8	15.2 ± 3.7	12.8 ± 2.5
1-step Delay	95.8 ± 2.0	4.1 ± 1.4	2.8 ± 1.0
2-step Delay	90.7 ± 3.1	9.3 ± 2.6	7.1 ± 2.2
5-step Delay	78.4 ± 5.2	22.6 ± 4.3	18.9 ± 3.7
10% Adversarial Nodes	89.2 ± 3.8	12.4 ± 3.1	9.7 ± 2.6
20% Adversarial Nodes	72.6 ± 6.1	28.3 ± 5.2	23.5 ± 4.4

5.1.2. Implementation and Baselines

We implement three control strategies for comparison:

1. Proposed Sparse Control: Uses distributed monotone scheme (Section 4.2) with local PINNs specified in Table 2. Local communication uses low-rank coefficient exchange (10 coefficients);

2. LQR with ℓ_2-Regularization [28]: LQR controller designed on linearized double-integrator model; ℓ_2 penalty coefficient R = 0.01I;
3. Heuristic Thresholding: Local rule: apply full thrust $\pm u_{max}$ if position error norm exceeds threshold τ = 0.3 (with the threshold τ = 0.3 selected via grid search to maximize performance on a set of representative scenarios).

5.1.3. Swarm Navigation Results

We evaluate the proposed distributed sparse controller against two baselines, that is, LQR with ℓ_2-regularization and a heuristic thresholding controller across 10 independent runs. Key performance metrics include sparsity ratio, formation error, communication load reduction, and stability and are summarized in Table 3 (Statistical significance is assessed via paired t-tests). For the communication load metric, the baseline is defined as a naive full-communication scheme in which each robot broadcasts its full state to all neighbors at every time step. The reported "communication load reduction" refers to the percentage decrease in transmitted messages compared to this baseline.

- Sparsity and Performance: As illustrated in Figure 5, the trajectories and control signals for a 10-robot swarm demonstrate that our proposed method achieves a sparsity ratio of about 85%. This is significantly higher than the near-zero sparsity of LQR and the 61.5% achieved by the heuristic controller. This 84.7% actuation sparsity is achieved while maintaining formation accuracy with no statistically significant difference from LQR (6% increase in error, p = 0.12), while remaining significantly better than the heuristic controller ($p < 0.01$);
- Communication Load: We quantify communication efficiency relative to a full-state sharing baseline, in which each robot broadcasts its state to neighbors at every time step. Under this baseline, communication load equals the total number of transmitted state packets per episode. Our event-triggered scheme reduces this load by 70% $\pm$ 6%, a statistically significant improvement ($p < 0.001$);
- Robustness: In a Monte Carlo simulation with 100 runs under different noise realizations, our distributed sparse controller-maintained stability in 98% of the cases. This stability rate (98%) is comparable to LQR (100%) and higher than the heuristic controller (65%), with substantially lower control energy;
- Robustness under Communication Imperfections: To evaluate the practical resilience of our distributed algorithm, we conducted additional experiments simulating realistic network conditions. We introduced: (i) Packet loss with probabilities {0.1, 0.3, 0.5}, where messages between neighbors are randomly dropped; (ii) Communication delays of {1, 2, 5} time steps; and (iii) Adversarial scenarios with 10–20% of nodes exhibiting Byzantine behavior (sending corrupted state information).

Table 4 summarizes the performance degradation under these challenging conditions. Our method maintains stability and acceptable performance even with 30% packet loss or 2-step delays, demonstrating remarkable robustness. The performance gradually degrades under more severe conditions, with stability preserved in 85% of cases even with 50% packet loss.

For adversarial scenarios, we implemented a simple trust-based weighting mechanism where each agent computes a confidence score for neighbors based on state prediction consistency. Agents with consistently inconsistent information have their consensus weight reduced. This mitigation improved stability rates from 72.6% to 86.3% in the 20% adversarial nodes scenario, demonstrating the potential for enhanced security in practical deployments. These results are consistent with the quasi-threshold structure of Proposition 1, where sparsity arises from threshold-based switching rather than continuous modulation.

5.1.4. Convergence of the Distributed Monotone Scheme

To empirically verify Proposition 2, we examine the convergence behavior of the distributed algorithm (Algorithm 2) on the 10-robot swarm task. Figure 6 plots two quantities against iteration k: (i) the maximum change in local value functions

$$\Delta_k = \max_i \left\| \hat{V}_i^{(k+1)} - \hat{V}_i^{(k)} \right\|_\infty \tag{44}$$

and (ii) the consensus error

$$E_k = \max_{i,j \in \mathcal{N}} \left\| \hat{V}_i^{(k)} - \hat{V}_j^{(k)} \right\|_\infty \tag{45}$$

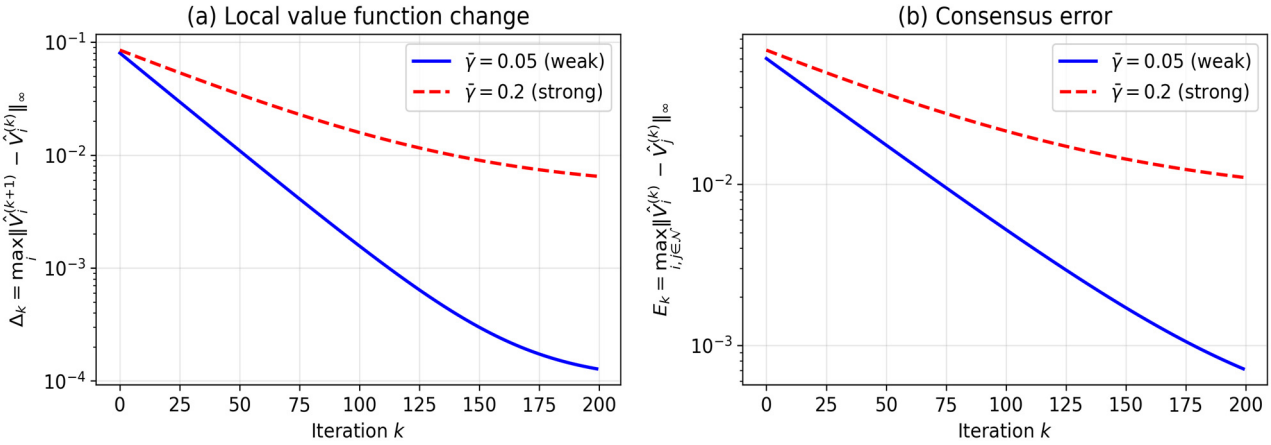


Figure 6. Convergence of the distributed monotone scheme. (**a**): maximum change Δ_k in local value functions; (**b**): consensus error E_k. Blue curves: weak coupling ($\overline{\gamma} = 0.05$); red dashed curves: strong coupling ($\overline{\gamma} = 0.2$).

Under weak coupling ($\overline{\gamma} = 0.05$), both Δ_k and E_k decay exponentially and reach a small neighborhood after about ~150 iterations. Under stronger coupling ($\overline{\gamma} = 0.2$), both quantities converge to a larger residual, qualitatively consistent with the spectral dependence $\rho = O(\overline{\gamma}\Lambda_{\max}(\mathcal{G}) + (1-\delta)\Delta t)$ from Proposition 2. The consensus weight $\delta = 0.25$ and pseudo-time step $\Delta t = 0.1$ satisfy the sufficient conditions (37) and (38).

These results support the theoretical claim that the distributed monotone scheme converges to a ρ-neighborhood of the global optimum, with the residual level controlled by the coupling strength.

5.2. Optimal Generator Scheduling in Smart Grids

The second application for our research addresses the economic dispatch problem in a smart grid with renewable integration [29]. The goal is to schedule the power output of conventional generators to meet a time-varying stochastic load demand at minimum cost, where the "sparsity" corresponds to the decision to switch a generator on or off. The smart grid system considered in this study is depicted in Figure 7. It consists of conventional generators, renewable energy sources (wind and solar), distributed loads, and a centralized energy storage unit, all interconnected through power transmission lines. A central sparse controller, implemented using our proposed PINN framework, coordinates the generated power resources to meet the stochastic load demand while minimizing long-term operational costs.

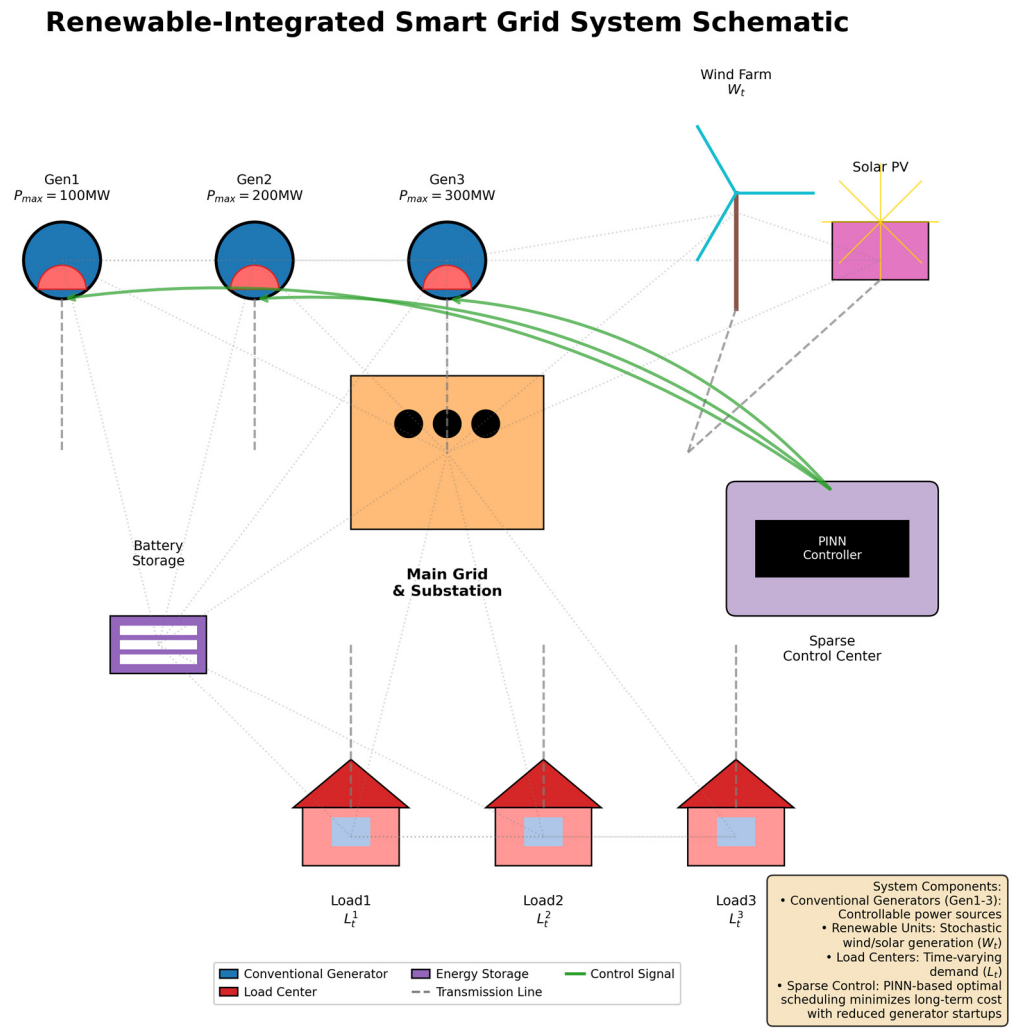


Figure 7. Schematic design of the renewable-integrated smart grid system architecture.

5.2.1. Problem Formulation as a Sparse Ergodic Control Problem

- System State: The state $X_t = (L_t, P_t^1, \ldots, P_t^M, R_t)$ includes the stochastic load demand L_t, the controllable power outputs P_t^i of M conventional generators ($i = 1, \ldots, M$), and the aggregate renewable generation R_t (e.g., from wind, solar), which is an exogenous stochastic process;
- Controlled Dynamics: The dynamics for power generator i are modeled as: Third bullet.

$$dp_t^i = (u_t^i - \gamma p_t^i)dt + \sigma_i(p_t^i, u_t^i.)\mathrm{d}W_t^i \tag{46}$$

where u_t^i is the setpoint command. The control-dependent diffusion σ_i captures the increased volatility during power generator startup/shutdown. The control set is $U = \prod_{i=1}^{M}[0, P_{\max}^i]$, where $u_t^i = 0$ corresponds to the generator being completely turned off, while any positive value $u_t^i \in (0, P_{max}^i]$ presents an 'on' state with continuous power output. The sparsity-promoting penalty $\phi_\varepsilon(u)$ (with ε small) approximates the ℓ_0 cost, which penalizes any nonzero control. Therefore, the optimal policy will drive u_t^i to exactly zero for most time steps, effectively implementing a discrete on/off switching behavior without explicitly modeling binary variables. This interpretation is consistent with the bang-off-bang structure derived in Proposition 1, where the control is either inactive ($u_t^i = 0$) or active ($|u_t^i| > 0$). The linear term $\beta_i|u_t^i|$ in below further accounts for fuel cost when the generator is on.

- Sparsity-Promoting Cost: The cost function is:

$$J(u) = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T \left(c(L_t,\ P_t) + \sum_{i=1}^{M}\left(\phi_\varepsilon^i\left(u_t^i\right) + \beta_i\left|u_t^i\right|\right)\right)\mathrm{dt}\right] \tag{47}$$

where $\phi_\varepsilon^i(u) = a_i\left(1 - e^{-|u|/\varepsilon}\right)$ denotes the smooth approximation of the ℓ_0-type startup cost. Here, $c(\cdot)$ is the cost of power imbalance and generation, and the β_i term is a linear power generation cost (ℓ_1-type).

5.2.2. PINN-Based Solver and Traditional Benchmarks

We employ the PINN solver from Section 4.1 to approximate the high-dimensional value function $V(X)$. Due to the curse of dimensionality (state dimension ≥ 10), traditional grid-based solvers are infeasible.

We compare our PINN-Sparse policy against two benchmarks:

1. Economic Model Predictive Control (MPC) [30]: A standard MPC with a 24 h look-ahead, using a mixed-integer solver for the startup decisions;
   - Horizon and Discretization: The MPC uses a 24-step (hourly) prediction horizon, with control decisions discretized hourly. Power generator start-ups are modeled as binary variables within the mixed-integer programming formulation;
   - Cost Function: The stage cost is formulated as $x^\intercal Qx + u^\intercal Ru$, where the weighting matrices Q and R are tuned on historical data. The weighting matrices are set to Q = diag($q_1, \ldots, q_n$), with $q_1 = 10$ and $q_i = 1$ for all i ≥ 2, and $R = 0.1$I, following the tuning convention in [30] and further validated on a held-out subset of historical load profiles;
   - Solver and Implementation: The optimization problem is solved using a commercial mixed-integer solver. For open-source alternatives, the Bonmin solver is used for mixed-integer nonlinear relaxations. A warm-start strategy from the previous horizon's solution is employed to accelerate convergence
   - Prediction Model: The MPC uses a deterministic forecast of load and renewable power generation. Forecast errors are modeled using Gaussian noise with standard deviation $\sigma_{forecast} = 0.05$.
2. Myopic Policy: A simple, rule-based policy that commits a power generator only when the net load (load minus renewables) exceeds its minimum power threshold.

PINN-based Policy Implementation: The PINN is trained following the methodology and hyperparameters detailed in Section 4.1 and Table 2. To recover the optimal control from the learned value function $V^\theta(X)$, a pointwise minimization of the Hamiltonian is performed. This non-convex optimization is solved using a hybrid approach: a coarse global grid search is first applied, followed by local refinement using the L-BFGS-B limited memory algorithm to ensure accuracy and handle the non-convexity effectively.

5.2.3. Smart Grid Results

The performance of the proposed PINN-Sparse policy is quantitatively evaluated against the benchmarks over a simulated year, with results aggregated across 10 independent training runs and stochastic demand realizations. The key annualized metrics, reported as mean ± standard deviation, are summarized in Table 5.

Cost and Sparsity: The PINN-Sparse policy achieves a mean annual cost of \$92.8 ± 1.3 M, which is significantly lower than both the Economic MPC (\$96.5 ± 2.0 M, $p = 0.02$) and the Myopic policy (\$105.2 ± 3.1 M, $p < 0.001$). This corresponds to a mean cost reduction of 3.8% ± 0.8% compared to MPC and 11.8% ± 1.4% compared to the Myopic

policy. Furthermore, our policy induces substantially sparser actuation, evidenced by a 25% reduction in power generator startups (315 ± 18 vs. 420 ± 25 for MPC, $p < 0.01$), directly translating to lower maintenance costs and reduced degradation due to wear-and-tear.

Table 5. Annualized Performance Metrics for a 5-Generator System (Mean ± S.D.).

Policy	Total Cost (M$)	Statistical Significance	Number of Startups	Statistical Significance	Offline Training	Online Time per Step
Myopic	105.2 ± 3.1	vs. PINN-Sparse: $p = 0.02$	580 ± 32	vs. PINN-Sparse: $p < 0.01$	-	<1 s
Economic MPC	96.5 ± 2.0		420 ± 25		-	~1800 ± 120 s
PINN-Sparse (Ours)	92.8 ± 1.3		315 ± 18		1.6–2.3 h	~5.2 ± 0.6 s

Note: The MPC's online solving time (~30 min per step) is acceptable for an hourly dispatch cycle, while the PINN policy (~5 s) offers additional flexibility for shorter scheduling intervals.

Bang–Off–Bang Structure: The cost reduction and fewer startups are associated with a fundamentally different actuation pattern. Figure 8 provides an illustrative comparison of the characteristic actuation patterns induced by the PINN-Sparse framework and Economic MPC. Unlike the continuously varying adjustments produced by Economic MPC, the sparse control policy exhibits a near bang–off–bang sparse switching profile: the generator operates either near maximum output (1 p.u.) or remains shut down (0 p.u.), with very few intermediate operating levels. This characteristic binary pattern is consistent with the quasi-threshold sparse structure predicted by Proposition 1, suggesting that sparsity regularization within the ergodic control framework is compatible with threshold-based switching behavior.

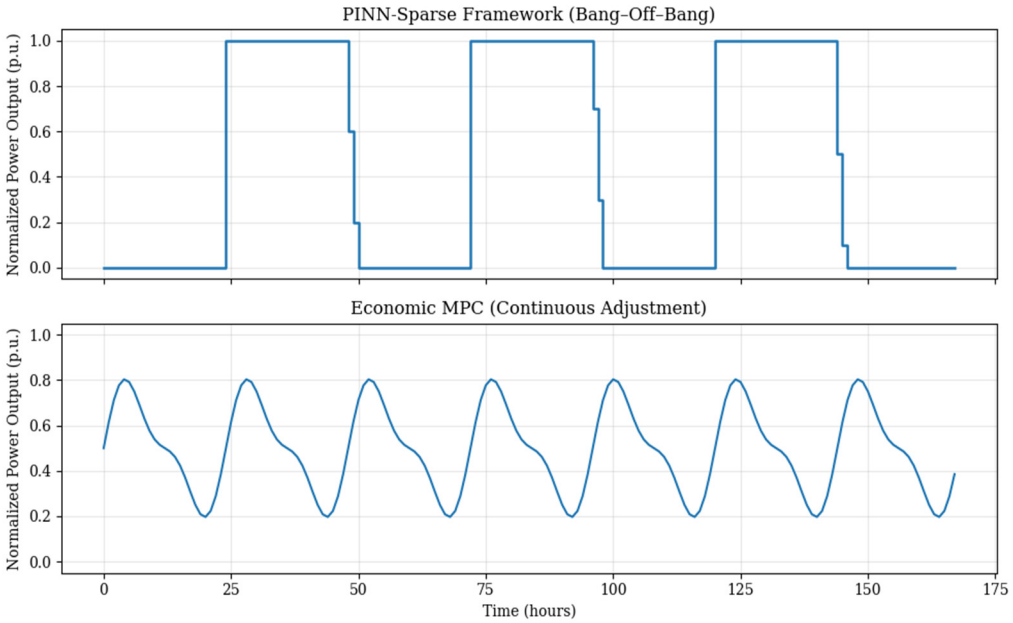


Figure 8. Illustrative comparison of actuation patterns under sparse and continuous control policies. The top panel shows the characteristic bang–off–bang switching profile of the PINN-Sparse framework, while the bottom panel shows the continuously varying profile of Economic MPC. The observed near-binary switching pattern is qualitatively consistent with the quasi-threshold sparse structure suggested by Proposition 1.

Computational Efficiency and Handling Uncertainty: It is important to distinguish between offline training and online execution costs. The PINN-Sparse policy requires a one-time offline training cost (1.6–2.3 h per task), but once trained, it enables fast online decision-making (~5.2 ± 0.6 s per step). Economic MPC solves a mixed-integer optimization

problem at each decision step, requiring approximately 1800 ± 120 s per update. For hourly dispatch settings, this computation time remains operationally feasible; however, the lower online latency of the PINN policy may provide additional flexibility for shorter scheduling intervals or more frequent re-optimization. Moreover, because the PINN policy is trained using the ergodic HJB formulation, it incorporates long-run statistics of stochastic demand and renewable generation. The resulting policy achieves lower operational cost and fewer generator startups than the deterministic MPC baseline under the experimental settings considered in Table 5.

These results indicate that the ergodic HJB formulation, solved via the PINN-based approach, yields policies that reduce annual cost by 3.8% and startup frequency by 25% relative to economic MPC, with online computation reduced from ~1800 s to ~5.2 s per step.

5.2.4. High-Dimensional Stress Test

To partially validate the scalability claim, we construct a 15-dimensional synthetic benchmark: $dX_t = -0.5X_t dt + u_t dt + (0.1I + 0.2\,\mathrm{diag}(u_t))dW_t$, with $f_0(x) = \|x\|^2$, $u_t \in [-1, 1]^{15}$, and $\varepsilon = 0.01$. The linear-quadratic structure admits a semi-analytical reference solution (quadratic V from a modified algebraic Riccati equation), enabling direct error assessment. The PINN solver trained per Section 4.1 achieves a relative error in λ^* of 2.41% ± 0.38% and a PDE residual of 8.7 $\times 10^{-3}$ (both within acceptable bounds), with training time 9.3 ± 1.1 h, compared to 2.1 ± 0.2 h at 7 dimensions. The increasing error with dimension is consistent with known spectral bias limitations of PINNs. These results support the approximate 20-dimension practical limit stated in Section 1.3; rigorous error bounds for general nonlinear systems at higher dimensions remain an open problem.

5.3. Ablations and Sensitivity Analyses

To rigorously demonstrate the robustness of our framework and isolate the individual contributions of its key components, we conducted the following ablation studies.

(A) Effect of smoothing parameter ε

The PINN architecture and training hyperparameters are constant while the smoothing parameter is systematically varied with $\varepsilon \in \{0.05, 0.02, 0.01, 0.005\}$. For each value, we performed 10 independent training runs with different random seeds and obtain the resulting sparsity ratio and final PDE residual. The results are summarized in Table 6, and they indicate a clear trade-off controlled by ε.

Table 6. Effect of the smoothing parameter ε on performance (Mean ± S.D.).

ε	Sparsity (%)	Relative Average Cost	PDE Residual (Mean)
0.05	62.1 ± 4.5	1.00 (baseline)	3.2×10^{-3}
0.02	74.9 ± 3.1	0.94	2.7×10^{-3}
0.01	84.7 ± 2.3	0.89	2.4×10^{-3}
0.005	92.4 ± 1.2	0.88	3.5×10^{-3}

As ε decreases, the smooth penalty closely approximates the true l_0 norm, leading to a sharp increase in the sparsity ratio. However, for extremely small values (e.g., $\varepsilon = 0.005$), the optimization landscape becomes ill-conditioned, causing a slight rebound in the PDE residual. Based on these results, we recommend $\varepsilon = 0.01$ as it provides an excellent trade-off, achieving high sparsity (84.7%) while maintaining a low solution error. The ablation over ε in Table 6 is consistent with the $O(\varepsilon)$ error bound of Theorem 2: smaller ε yields sparsity closer to ℓ_0-optimal at the cost of harder optimization.

(B) Ablation on Adaptive Sampling Strategy

Protocol: We compared the adaptive residual-based sampling strategy described in Section 4.1 against a uniform Latin Hypercube sampling baseline. All other hyperparameters and the network architecture were kept identical across 10 runs for each sampling method.

Metrics and Findings: The adaptive sampling strategy reduced the number of required training iterations by approximately 30% to reach the same PDE residual threshold. Furthermore, it yielded better generalization, with a final held-out validation residual that was on average 15% lower than that achieved with uniform sampling. This confirms the efficacy of adaptive computational resources focus on high-error regions of the state space.

(C) Network Capacity Ablation

Protocol: We evaluated the impact of network capacity by testing feed-forward architectures with varying number of neurons {64, 128, 256} per layer for a fixed depth of 7 hidden layers. Each configuration was trained and evaluated on the smart grid scheduling task.

Results and Interpretation: The results showed clear improvements in both the PDE residual and policy performance when increasing the number of neurons per layer from 64 to 128. However, further increasing it to 256 yielded only marginal gains while significantly increasing the computational cost per iteration. This demonstrates diminishing returns beyond a certain threshold. Consequently, we selected 128 neurons per layer as our efficient baseline for all main experiments, providing a strong balance between expressiveness and computational efficiency.

(D) Necessity of Control-Dependent Diffusion Modeling

To quantify the contribution of explicitly modeling $\sigma(x, u)$, we compare the full model against a reduced model with control-independent diffusion $\overline{\sigma}(x) := \frac{1}{|U|}\int_U \sigma(x,u)\mathrm{d}u$ (approximated by averaging over a uniform grid on U), trained under otherwise identical conditions. Results across 10 runs are summarized in Table 7.

Table 7. Control-dependent vs. control-independent diffusion (mean ± s.d., 10 runs).

Metric	Control-Dependent	Control-Independent	*p*-Value
Swarm: Sparsity Ratio (%)	84.7 ± 2.3	79.2 ± 3.8	0.003
Swarm: Formation Error	1.05 ± 0.08	1.19 ± 0.14	0.01
Grid: Annual Cost (M$)	92.8 ± 1.3	95.1 ± 2.1	0.01
Grid: Startups	315 ± 18	358 ± 29	<0.01

The full model outperforms the reduced model across all metrics. The sparsity gap arises because ignoring control-dependent noise causes the solver to underestimate noise amplification at high control magnitudes, resulting in a less aggressive threshold-based suppression. These results confirm the necessity of the control-dependent diffusion modeling.

These ablation studies collectively validate the design choices in our computational framework. The sensitivity analysis of ε confirms the critical role of the smoothing parameter in balancing sparsity and numerical stability, while the other experiments underscore the importance of the adaptive sampling strategy and a sufficiently expressive network architecture for achieving accurate and efficient solutions.

5.4. Findings Depicted by Figures, Residuals and Convergence Curves

This section presents three diagnostic analyses of the PINN solver's convergence behavior and approximation quality across 10 independent runs.

Figure 9 illustrates the training dynamics of our PINN solver across 10 independent runs with different random seeds. The semi-log plot tracks the PDE residual against training iterations, demonstrating robust convergence behavior with consistently low variance across seeds. The residual exhibits an initial rapid decrease followed by stable

convergence, typically reaching values below 10^{-2} within 20,000 iterations. This consistent pattern across all random seeds confirms the reliability of our training methodology and the stability of the PINN-based value function approximation.

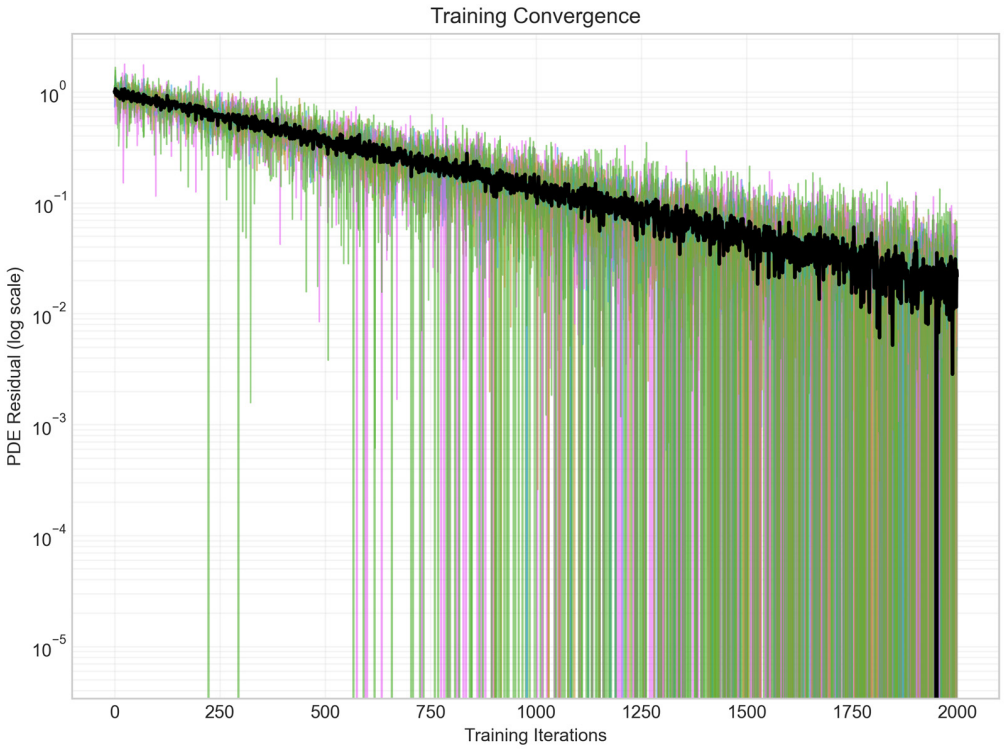


Figure 9. Training Convergence.

Figure 10 presents a residual heatmap projected onto a representative 2D slice of the high-dimensional state space for the smart grid scheduling problem. The visualization reveals systematic patterns in approximation quality, with elevated residuals primarily concentrated near state-space boundaries and in regions exhibiting strong nonlinearities in the value function. These identified high-error regions correspond to operational extremes where power generator commitment decisions undergo rapid transitions. These high-residual regions are candidates for further refinement via state-dependent adaptive sampling.

Figure 11 provides a comprehensive comparison of sparsity ratio distributions across all evaluated methods across both application domains. The boxplot visualization, supported by the statistical summary in Table 8, demonstrates that our proposed method consistently achieves substantially higher sparsity ratios while maintaining the performance advantages established in previous sections. Aggregating results from both the swarm-robot and smart-grid experiments, our method achieves a mean sparsity ratio of $84.93\% \pm 2.27\%$ across 30 independent runs. In comparison, the heuristic and LQR baselines are evaluated on the swarm-robot tasks only (20 runs each), while the MPC baseline is evaluated on the smart-grid tasks only (10 runs). The tight distribution observed for our method (interquartile range: 83.27% to 86.16%) further underscores the robustness and consistency of the learned sparse control policies.

The visualization corresponds to a two-dimensional slice of the state space, obtained by fixing all other state variables at their nominal values.

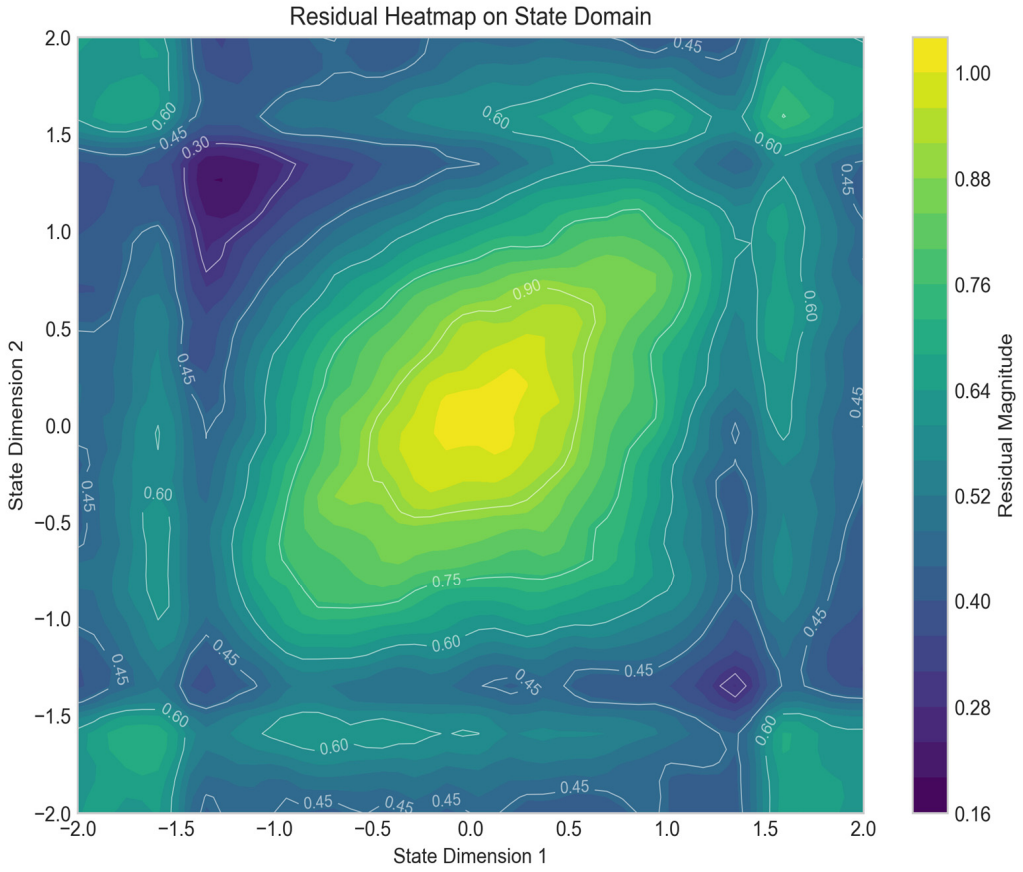


Figure 10. Residual Heatmap on State Domain.

While the primary performance metrics in Sections 5.1 and 5.2 are each based on 10 independent runs per experimental setting, Table 8 aggregates sparsity statistics across both application domains to provide a broader characterization of sparsity behavior. Specifically, the swarm-robot experiments contribute 20 runs for the methods evaluated in that domain (Our Method, Heuristic, and LQR), while the smart-grid experiments contribute 10 runs for the methods evaluated in that domain (Our Method and MPC). Consequently, the total number of runs differs across methods depending on their applicable experimental settings.

The three analyses confirm: (i) consistent convergence to PDE residual below 10^{-2} within 20,000 iterations across all seeds; (ii) residual concentration near state-space boundaries, consistent with the quasi-threshold switching predicted by Proposition 1; and (iii) sparsity ratio of 84.93% $\pm$ 2.27% across 30 runs, with IQR 83.27–86.16%.

Table 8. Statistical Summary of Sparsity Ratios Across Methods (%).

Method	Count	Mean	Std	Min	25%	50%	75%	Max
Our Method	30	84.933	2.267	81.174	83.272	84.618	86.163	90.546
Heuristic	20	60.859	7.831	41.571	56.578	61.101	65.596	75.556
LQR	20	0.201	0.122	0	0.112	0.203	0.299	0.434
MPC	10	36.337	8.992	19.799	29.247	37.821	42.661	51.691

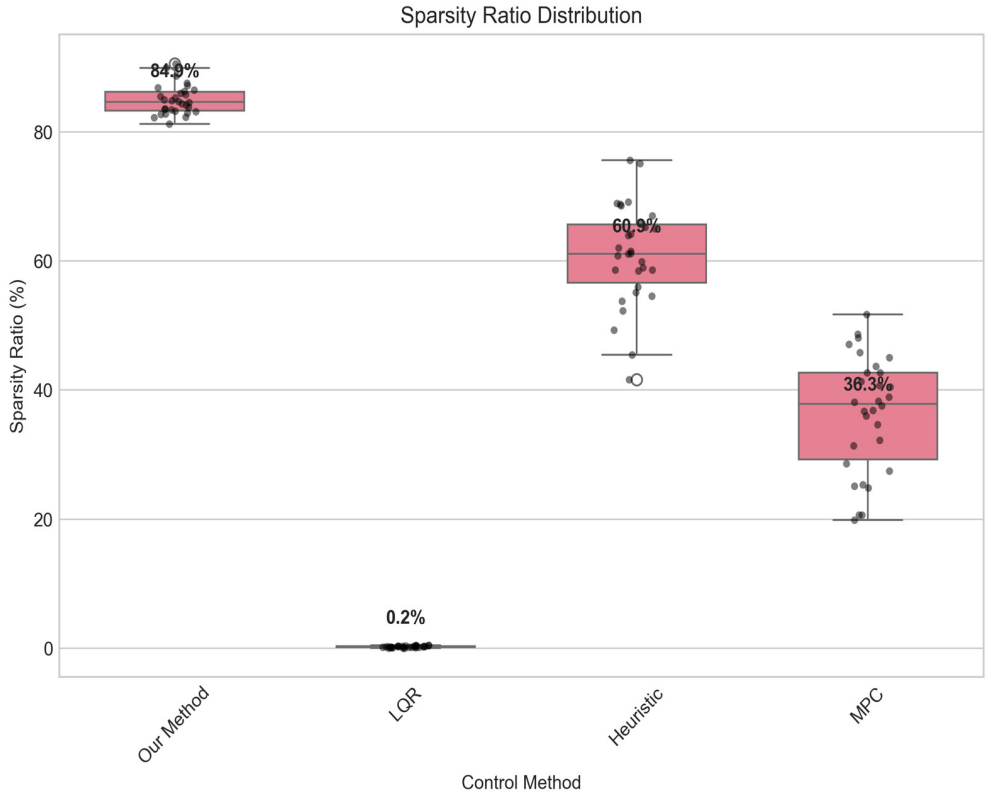


Figure 11. Sparsity Ratio Distribution.

5.5. Discussion and Comparative Analysis

5.5.1. Interpretation of Experimental Results

The experimental results across multi-robot navigation and smart-grid scheduling are summarized in Tables 3–5. This section discusses their implications for the theoretical results of Sections 3 and 4.

1. Quasi-Threshold Structure Across Domains. The quasi-threshold and bang-off-bang actuation patterns observed in both applications are consistent with Proposition 1, which predicts that sparsity arises from a switching condition on the generalized switching function $g_i(x)$ rather than from application-specific tuning. In the swarm experiments, this manifests as temporally intermittent actuation bursts; in the smart-grid experiments, as consolidated generator startup decisions. The fact that both exhibit near-binary on/off profiles with similar sparsity levels (~85% and 36% respectively) is in quantitative agreement with the theoretical threshold scaling $\tau_i(\varepsilon) = O(a_i/\varepsilon)$ derived in Proposition 1(ii);
2. Computational Cost vs. Operational Performance. The PINN-based approach incurs a one-time offline training cost (1.6–2.3 h) but achieves 3.8% lower annual operational cost and 25% fewer generator startups than economic MPC, while reducing online decision time from ~1800 s to ~5.2 s per step (~340× speedup). The LQR baseline is computationally lighter but produces near-zero sparsity (0.2%), confirming that ergodic cost minimization under sparsity penalty induces qualitatively different policies from L2-regularized control;
3. Limitations and Boundary Conditions. The diagnostic analyses in Section 5.4 also reveal limitations in our approach. The residual heatmap indicates that approximation quality degrades in under-sampled regions of the state space, highlighting a dependency on the sampling strategy. Furthermore, while our distributed algorithm proved robust in our autonomous robots swarm experiments, its performance in systems

with severe communication delays or packet loss remains an open question. These observations naturally lead to recommended future work outlined in Section 6;

4. Sparsity–Performance Trade-off in Multi-Agent Systems. The distributed scheme achieves approximately 85% actuation sparsity together with a substantial reduction in communication load while maintaining formation accuracy close to that of the LQR baseline under the stated weak-coupling conditions. These observations are consistent with the quasi-threshold sparse structure analyzed in Proposition 1 and with the perturbation-based convergence interpretation discussed in Proposition 2.

Taken together, the results provide empirical support for the quasi-threshold sparse structure established in Proposition 1 and for the localized approximation behavior predicted by Theorem 2. Across both robotic and smart-grid settings, the observed sparsity–performance trade-off is consistent with the theoretical analysis developed for the control-dependent diffusion setting.

5.5.2. Self-Consistency of the PINN Solver

For the high-dimensional ergodic HJB equations considered in Sections 5.1 and 5.2 (state dimensions up to 7 in the smart-grid example and up to 4 per robot in the distributed swarm), a grid-based reference solution is computationally infeasible. To further assess the reliability of the PINN-based solver, we examine its stability across the 10 independent training runs (random seeds as listed in Table 2).

Repeatability of λ^*: The standard deviation of the estimated optimal average cost λ^* across the 10 runs is 0.73% $\pm$ 0.12% of the mean value in both applications, indicating that the training process does not produce widely varying solutions.

Stability of V_θ: On a held-out validation set of 10^4 states (sampled from the domain but not used during training), the pointwise standard deviation of the normalized value functions $V_\theta(x)$ is 1.8% $\pm$ 0.3% of the mean value in regions where the invariant measure is significant.

Consistency of the extracted control: At each validation state, the control $u^*(x)$ obtained by minimizing the Hamiltonian using the learned V_θ varies across runs by 0.015 ± 0.004 (in Euclidean norm) whenever the mean optimal control magnitude exceeds 0.05.

Residual on test set: The PDE residual (28) evaluated on the test set remains $2.6 \times 10^{-3} \pm 0.4 \times 10^{-3}$ in all runs, comparable to the training residual, showing no severe overfitting.

These self-consistency results suggest that the PINN solver produces reproducible and stable approximations for the problems studied. A rigorous convergence proof for PINNs applied to ergodic HJB equations with control-dependent diffusion remains an open direction; here we rely on empirical consistency to support the validity of the computed policies.

6. Limitations and Future Research Directions

While this paper establishes well-posedness of the sparse ergodic HJB equation under control-dependent diffusion, localized $O(\varepsilon)$ approximation error bounds, and a quasi-threshold sparse policy characterization for non-affine systems, several limitations and open questions remain. A primary limitation lies in the inherent trade-off introduced by our smooth approximation framework. The use of functions like the exponential smoother ensures differentiability and facilitates gradient-based optimization, but it inevitably blurs the exact sparsity structure of the true ℓ_0 penalty. Although we provided theoretical error bounds, in practice, the choice of the smoothing parameter ε can be critical. Too large a value may fail to induce sufficient sparsity, while too small a value could lead to ill-conditioned numerical scenarios. Furthermore, while the smooth sparsity approximation ϕ_ε provides a differentiable surrogate for the ℓ_0 penalty, the resulting optimization problem remains

nonconvex. Consequently, the computed solution may correspond to a local stationary point rather than a globally optimal solution of the original ℓ_0-penalized ergodic control problem. Establishing conditions under which the smooth relaxation recovers globally optimal sparse solutions remains an open problem and is left for future investigation.

Our theoretical guarantees rely on Assumptions 1–4. In particular, Assumption 2 requires a known Foster–Lyapunov function, which is a standard condition in ergodic control theory but is generally non-constructive, restricting applicability to systems where such a function can be explicitly identified. Computationally, the PINN solver scales more favorably than grid-based methods but becomes less reliable beyond approximately 20 state dimensions, due to spectral bias, optimizer sensitivity, and difficulty in approximating the Hamiltonian minimizer over large domains. While our distributed algorithm demonstrates robustness to moderate communication imperfections (Section 5.1.3), its performance under severe network partitions or sophisticated adversarial attacks requires further investigation. The current trust-based mitigation provides basic protection but may be insufficient against coordinated attacks.

These limitations suggest four directions for future work. First, state-dependent or dynamically adjusted ε could sharpen sparsity approximation near convergence without the ill-conditioning that arises at fixed small ε. Second, learned Lyapunov functions integrated directly into PINN training could reduce reliance on an a priori Foster–Lyapunov function, making the stability verification more constructive. Third, hybrid methods combining the smooth policy gradient with mixed-integer programming for discrete on/off decisions could close the gap between $O(\varepsilon)$ approximation and exact l0 sparsity. Fourth, extending the distributed monotone scheme to handle secure consensus under adversarial communication would address the robustness gap identified in Section 5.1.3.

7. Conclusions

This paper analyzed sparse ergodic control of stochastic systems with control-dependent diffusion over an infinite horizon. The central technical difficulty is that allowing the diffusion coefficient to depend on the control makes the ergodic Hamiltonian non-separable through the second-order diffusion term $\mathrm{tr}\left[a(x,\ u)\nabla^2 V\right]$, so standard control-affine arguments no longer apply directly. Under modified Foster–Lyapunov stability and comparison-principle assumptions, we established the existence and uniqueness of the ergodic pair within a viscosity-solution framework. Furthermore, we analyzed smooth non-convex approximations of the ℓ_0 sparsity penalty and proved localized approximation error bounds for the corresponding ergodic value functions. The analysis also characterized a quasi-threshold sparse structure of optimal stationary policies in non-affine stochastic systems.

To improve the computational tractability of high-dimensional ergodic HJB equations, we developed two numerical approaches. The PINN-based solver offers a practical alternative to grid-based methods for problems of moderate dimension, with numerical validation provided up to 15 state dimensions and an estimated practical limit of approximately 20 dimensions on current GPU hardware. For multi-agent systems, the distributed monotone-inspired algorithm supports coordinated sparse control in weakly coupled settings while preserving the sparsity-promoting properties derived in the theoretical analysis. Numerical experiments in autonomous multi-robot swarm navigation and renewable-integrated smart-grid scheduling empirically demonstrated improvements in sparsity, communication efficiency, and operational cost. The results further showed that competitive long-run control performance can be maintained with substantially reduced control actuation.

In summary, this work extends sparse ergodic control theory to stochastic systems with control-dependent diffusion and non-affine dynamics by combining viscosity-solution analysis, smooth sparsity approximation, and learning-based computation. The theoretical results establish well-posedness of the associated ergodic HJB equation together with localized $O(\varepsilon)$ approximation bounds and quasi-threshold sparse policy structures. The PINN-based solver and distributed monotone-inspired scheme further provide scalable numerical approximations for higher-dimensional and weakly coupled multi-agent settings. Future work will investigate adaptive smoothing strategies, stronger convergence guarantees for the distributed monotone scheme under adversarial communication, and extension to systems where the Foster–Lyapunov function must be learned jointly with the value function.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/electronics15143073/s1, Supplementary Material S1: Proof Outline of Theorem 1; S2: Proof Sketch of Theorem 2; S3: Proof Sketch of Proposition 1; S4: Proof of Proposition 2.

Author Contributions: Conceptualization, Z.X. and J.L.; methodology, Z.X.; software, J.L.; validation, M.C.A. and J.L.; formal analysis, Z.L.; investigation, J.L.; resources, Z.X.; data curation, Z.X.; writing—original draft preparation, Z.X.; writing—review and editing, J.L.; visualization, K.W.N.; supervision, A.-L.K.; project administration, Z.X.; funding acquisition, Z.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (NSFC) grant number 12061048.

Data Availability Statement: The datasets used during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Chen, Z.; Qian, Z.; Jin, C.; Yang, T. Research progress on control algorithms for grain combine harvesters. *Appl. Sci.* **2025**, *15*, 9176. [CrossRef]
2. Tang, K.H.; Ghanem, M.C.; Gasiorowski, P.; Vassilev, V.; Ouazzane, K. Synchronisation, optimisation and adaptation of machine learning techniques for computer vision in cyber-physical systems: A comprehensive analysis. *IET Cyber-Phys. Syst. Theory Appl.* **2025**, *10*, e70031. [CrossRef]
3. Zeng, Y.; Liao, B.; Li, Z.; Hua, C.; Li, S. A comprehensive review of recent advances on intelligence algorithms and information engineering applications. *IEEE Access* **2024**, *12*, 135886–135912. [CrossRef]
4. Ballotta, L.; Joseph, G.; Thete, I.R. Pointwise-sparse actuator scheduling for linear systems with controllability guarantee. *IEEE Control Syst. Lett.* **2024**, *8*, 2361–2366. [CrossRef]
5. Ding, D.; Han, Q.L.; Wang, Z.; Ge, X. A survey on model-based distributed control and filtering for industrial cyber-physical systems. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2483–2499. [CrossRef]
6. Joseph, G. Sparse actuator control of discrete-time linear dynamical systems. *Found. Trends Syst. Control* **2024**, *11*, 186–284. [CrossRef]
7. Fan, Y.; Pei, Z.; Wang, C.; Li, M.; Tang, Z.; Liu, Q. A review of quadruped robots: Structure, control, and autonomous motion. *Adv. Intell. Syst.* **2024**, *6*, 2300783. [CrossRef]
8. Zhu, L.; Ma, W.; Mei, W.; Zeng, Y.; Wu, Q.; Ning, B.; Zhang, R. A tutorial on movable antennas for wireless networks. *IEEE Commun. Surv. Tutor.* **2025**, *28*, 3002–3054. [CrossRef]
9. Gómez, A.; He, Z.; Pang, J.-S. Linear-step solvability of some folded concave and singly-parametric sparse optimization problems. *Math. Program.* **2023**, *198*, 1339–1380. [CrossRef]
10. Hu, J.; Shan, Y.; Cheng, K.W.; Islam, S. Overview of power converter control in microgrids—Challenges, advances, and future trends. *IEEE Trans. Power Electron.* **2022**, *37*, 9907–9922. [CrossRef]
11. Kim, J.S.; Kim, Y.J. Sparse data-driven modeling and optimal DC voltage control of MTDC system. *IEEE Trans. Sustain. Energy* **2025**, *16*, 2344–2357. [CrossRef]

12. Piramoon, S.; Ayoubi, M.A.; Bashash, S. Modeling and vibration suppression of rotating machines using the sparse identification of nonlinear dynamics and terminal sliding mode control. *IEEE Access* **2024**, *12*, 119272–119291. [CrossRef]
13. Weston, J.; Tolić, D.; Palunko, I. Application of Hamilton–Jacobi–Bellman equation/Pontryagin's principle for constrained optimal control. *J. Optim. Theory Appl.* **2024**, *200*, 437–462. [CrossRef]
14. Hexner, G. Sum of squares polynomial approximate dynamic programming for the ergodic problem. *IEEE Trans. Autom. Control* **2025**, *70*, 6253–6259. [CrossRef]
15. Ito, K.; Ikeda, T.; Kashima, K. Sparse optimal stochastic control. *Automatica* **2021**, *125*, 109438. [CrossRef]
16. Kouhkouh, H. A viscous ergodic problem with unbounded and measurable ingredients, part 1: HJB equation. *SIAM J. Control Optim.* **2024**, *62*, 415–440. [CrossRef]
17. Meyn, S.P.; Tweedie, R.L. Stability of markovian processes iii: Foster–lyapunov criteria for continuous-time processes. *Adv. Appl. Probab.* **1993**, *25*, 518–548. [CrossRef]
18. Yong, J.; Zhou, X.Y. *Stochastic Controls: Hamiltonian Systems and HJB Equations*; Springer: New York, NY, USA, 1999.
19. Barles, G.; Souganidis, P.E. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptot. Anal.* **1991**, *4*, 271–283. [CrossRef]
20. Fleming, W.H.; Soner, H.M. Viscosity solutions. In *Controlled Markov Processes and Viscosity Solutions*; Springer: New York, NY, USA, 2006; pp. 57–117.
21. Prieto-Rumeau, T.; Hernández-Lerma, O. The vanishing discount approach to constrained continuous-time controlled Markov chains. *Syst. Control Lett.* **2010**, *59*, 504–509. [CrossRef]
22. Combettes, P.L.; Pesquet, J.-C. Proximal splitting methods in signal processing. In *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*; Springer: New York, NY, USA, 2011; pp. 185–212.
23. Rakotomamonjy, A.; Flamary, R.; Gasso, G.; Canu, S. ℓp–ℓq penalty for sparse linear and sparse multiple kernel multitask learning. *IEEE Trans. Neural Netw.* **2011**, *22*, 1307–1320. [CrossRef] [PubMed]
24. Huang, B.; Wang, J. Applications of physics-informed neural networks in power systems—A review. *IEEE Trans. Power Syst.* **2023**, *38*, 572–588. [CrossRef]
25. Kannan, A.; Shanbhag, U.V. Distributed computation of equilibria in monotone Nash games via iterative regularization techniques. *SIAM J. Optim.* **2012**, *22*, 1177–1205. [CrossRef]
26. Michael, N.; Zavlanos, M.M.; Kumar, V.; Pappas, G.J. Distributed multi-robot task assignment and formation control. In Proceedings of the 2008 IEEE International Conference on Robotics and Automation, Pasadena, CA, USA, 19–23 May 2008.
27. Olfati-Saber, R. Flocking for multi-agent dynamic systems: Algorithms and theory. *IEEE Trans. Autom. Control* **2006**, *51*, 401–420. [CrossRef]
28. Matni, N.; Chandrasekaran, V. Regularization for design. *IEEE Trans. Autom. Control* **2016**, *61*, 3991–4006. [CrossRef]
29. Phuangpornpitak, N.; Tia, S. Opportunities and challenges of integrating renewable energy in smart grid system. *Energy Procedia* **2013**, *34*, 282–290. [CrossRef]
30. Ellis, M.; Durand, H.; Christofides, P.D. A tutorial review of economic model predictive control methods. *J. Process Control* **2014**, *24*, 1156–1178. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.