2026, № 1
Журнал обчислювальної та прикладної математики
Заснований в 1965 році

Journal of Numerical & Applied Mathematics
ISSN (Print): 2706-9680, ISSN (Online): 2706-9699
ЄДРПОУ видавця: 02070944
ROR: 02aaqv166

E-mail:
jnameditorialteam@gmail.com

Макарович А.В., Капустян О.А.

МЕТОД ДИНАМІЧНОГО ПРОГРАМУВАННЯ ДЛЯ ОДНІЄЇ ЛІНІЙНО-КВАДРАТИЧНОЇ ЗАДАЧІ ОПТИМАЛЬНОГО КЕРУВАННЯ В УМОВАХ НЕВИЗНАЧЕНОСТІ

Адреса редакції:
Кафедра обчислювальної математики,
факультет комп'ютерних наук та кібернетики,
Київський національнний університет імені Тараса Шевченка,
Україна, 01601, Київ, вул. Володимирська, 64/13.

Свідоцтво про державну реєстрацію КВ 4246 від 26.05.2000
Включено в Перелік наукових фахових видань у відповідності до Наказу МОН України № 409 від 17.03.2020

Журнал обчислювальної та прикладної математики **2026, № 1** Journal of Numerical & Applied Mathematics

DOI: 10.17721/2706-9699.2026.1.02

УДК 517.977, 517.95

MSC 49J20, 49N10

DYNAMIC PROGRAMMING APPROACH FOR ONE LINEAR QUADRATIC OPTIMAL CONTROL PROBLEM WITH UNCERTAINTY

A. V. Makarovych, O. A. Kapustian

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine,
E-mail: adalbert.makarovych@gmail.com, ORCID: 0009-0000-2352-9933,
E-mail: olenakapustian@knu.ua, ORCID: 0000-0002-2629-0750

МЕТОД ДИНАМІЧНОГО ПРОГРАМУВАННЯ ДЛЯ ОДНІЄЇ ЛІНІЙНО-КВАДРАТИЧНОЇ ЗАДАЧІ ОПТИМАЛЬНОГО КЕРУВАННЯ В УМОВАХ НЕВИЗНАЧЕНОСТІ

А. В. Макарович, О. А. Капустян

Київський національний університет імені Тараса Шевченка, Київ, Україна,
E-mail: adalbert.makarovych@gmail.com, ORCID: 0009-0000-2352-9933,
E-mail: olenakapustian@knu.ua, ORCID: 0000-0002-2629-0750

Abstract. This article focuses on solving the linear-quadratic (LQ) optimal control problem for a parabolic partial differential equation (PDE) operating under parametric uncertainty. To manage this uncertainty, we model the unknown parameter via a probability distribution and minimize the expected value of the cost functional.

Applying a Dynamic Programming approach to this system yields an exact optimal state-feedback control law, which is governed by a infinite-dimensional Integro-Differential Riccati Equation (IDRE). Because solving this equation directly is computationally prohibitive, we employ a Spectral Galerkin method combined with a polynomial chaos expansion to approximate the stochastic parameter space. This mathematical transformation reduces the complex, stochastic IDRE into a standard Matrix Riccati Differential Equation (MRDE).

By reducing the stochastic PDE control problem to an MRDE that can be solved in advance, our framework avoids the severe computational bottlenecks typically associated with uncertain environments. This provides a highly efficient baseline for robust controller design and future reinforcement learning implementations.

Keywords: optimal control, parabolic differential equations, dynamic programming, polynomial chaos expansion.

Corresponding author: A.V. Makarovych (adalbert.makarovych@gmail.com).

Анотація. Ця стаття присвячена розв'язанню лінійно-квадратичної (ЛК) задачі оптимального керування для параболічного диференціального рівняння (ПДР) в умовах параметричної невизначеності. Щоб впоратися з цією невизначеністю, ми моделюємо невідомий параметр за допомогою розподілу ймовірностей та мінімізуємо математичне сподівання функціонала якості.

Застосування апарату динамічного програмування до цієї системи дає точний закон оптимального керування зі зворотним зв'язком за станом, який визначається нескінченновимірним інтегро-диференціальним рівнянням Ріккаті (ІДРР). Оскільки безпосереднє розв'язання цього рівняння є обчислювально складним, ми застосовуємо спектральний метод Гальоркіна в поєднанні з розкладом поліноміального хаосу для апроксимації стохастичного простору параметрів. Це математичне перетворення зводить стохастичне ІДРР до стандартного матричного диференціального рівняння Ріккаті (МДРР).

Зводячи стохастичну задачу керування ПДР до МДРР, яке можна розв'язати заздалегідь, наш підхід дозволяє уникнути обчислювальних перешкод, що зазвичай виникають у середовищах з невизначеністю. Це створює високоефективну основу для проєктування робастних регуляторів і майбутніх реалізацій алгоритмів навчання з підкріпленням.
Ключові слова: оптимальне керування, параболічні диференціальні рівняння, динамічне програмування, розклад поліноміального хаосу.

1. Introduction

Optimal control problems for partial differential equations (PDEs) are widely used to model and manage complex dynamic systems evolving over time and space. Among these, the Linear Quadratic Regulator (LQR) problem represents a cornerstone of modern control theory, providing an optimal feedback strategy for linear PDEs while minimizing a quadratic cost functional [1]. However, classical LQR formulations inherently assume that the system dynamics are completely deterministic and perfectly known. In real-world physical and engineering systems, the governing parameters — such as the diffusion coefficient in a parabolic medium — are often subject to measurement errors or environmental variability.

To address this distributed parametric uncertainty, recent theoretical advances have moved beyond single deterministic models by representing uncertain parameters via probability distributions over a set of possible system dynamics [1]. This averaged optimal control framework seeks to minimize the expected value of the cost functional. The foundational properties of such averaged continuous-time systems, including the existence of solutions and the convergence of optimal controls, have been rigorously established in recent literature [1, 2].

The challenge of controlling systems with partially or completely unknown dynamics also forms the core of modern Reinforcement Learning (RL) [3]. The profound connection between RL and classical optimal control theory has long been recognized, with RL often viewed as a form of direct adaptive optimal control [4, 5]. Recent perspectives emphasize a unified view bridging continuous control and learning-based approaches [6]. For instance, novel model-based and online algorithms have been developed to simultaneously identify unknown system configurations and synthesize LQR controllers for both ordinary and partial differential equations [3, 7]. While these data-driven algorithms demonstrate remarkable empirical success, they heavily benefit from rigorous theoretical baselines that can pre-compute optimal averaged policies to guarantee robustness and stability during the learning process.

Despite its theoretical appeal, finding the optimal control strategy for an averaged PDE presents severe mathematical and computational challenges. Applying dynamic programming [8] to such a system results in a highly complex, infinite-dimensional equation known as an Integro-Differential Riccati Equation (IDRE). Solving this equation is computationally prohibitive. To overcome this obstacle, this paper proposes a framework to mathematically simplify the problem. We apply a Spectral Galerkin projection [9] combined with a Generalized Polynomial Chaos (gPC) expansion [10]. By representing the system's uncertainty using a specific set of orthogonal polynomials [11], we can take advantage of their structural properties to eliminate the complicated integral calculations. This transformation rigorously reduces the difficult IDRE into a standard, finite-dimensional Matrix Riccati Differential Equation (MRDE) [12]. Because this new matrix equation is strictly finite, it can be efficiently solved beforehand, providing an exact and reliable feedback control law.

This paper is organized as follows. Section 2 defines the mathematical formulation of the problem. Section 3 outlines the preliminary results, establishing the necessary mathematical foundations. Section 4 presents the main results through a rigorous three-step derivation. It details the spectral decomposition of the spatial domain, the dynamic programming derivation of the IDRE, and the Spectral Galerkin projection that yields the computable, finite-dimensional MRDE. Section 5 illustrates the theoretical framework with a numerical example. It demonstrates the practical computation of the MRDE and visualizes the resulting optimal control alongside the expected physical state. Finnaly, conclusion the contributions of the paper and discusses prommising avenues for future research, including integration with data-driven Reinforcement Learning techniques.

2. Setting of the problem

We consider an optimal control problem for a parabolic system defined on the one-dimensional spatial domain $[0, L]$ over a finite time horizon $[0, T]$. The state of the system is denoted by $y(t, x)$, and it is driven by a distributed control input $u(t, x)$. The physical dynamics are governed by the following linear parabolic partial differential equation (PDE):

$$\begin{cases} \frac{\partial y(t,x)}{\partial t} - a\frac{\partial^2 y(t,x)}{\partial x^2} = u(t,x), & t \in (0,T],\ x \in (0,L), \\ y(t,0) = y(t,L) = 0, & t \in [0,T], \\ y(0,x) = y_0(x), & x \in [0,L], \end{cases} \tag{1}$$

where $y_0 \in L^2(0,L)$ is a known initial state profile. The control variable is restricted to the space of square-integrable functions, $u \in L^2(Q_T)$, where $Q_T = (0,T) \times (0,L)$ denotes the space-time cylinder.

The parameter a represents the true physical diffusion coefficient of the medium. To ensure the mathematical problem is well-posed and physically realistic, a must satisfy the uniform ellipticity condition: $a \geq \gamma_3$, where $\gamma_3 > 0$ is a fixed physical constant.

If the parameter a were perfectly known to the controller, the performance of the control strategy u would be evaluated using the standard deterministic Linear Quadratic Regulator (LQR) cost functional:

$$J(u) = \int_0^T \left(\int_0^L y^2(t,x)\,dx + \gamma_1 \int_0^L u^2(t,x)\,dx \right) dt + \gamma_2 \int_0^L y^2(T,x)\,dx, \tag{2}$$

where $\gamma_1 > 0$ and $\gamma_2 > 0$ are fixed penalty weights for the control effort and the final state deviation, respectively.

However, in this study, the exact numerical value of a is unknown. To handle this parametric uncertainty while strictly satisfying the structural constraints of the PDE, we rely on the averaged optimal control formulation introduced in [1]. We model our knowledge of the unknown system parameter as a known probability measure π over the admissible parameter space

$$\mathcal{A} = \{a \in \mathbb{R} : a \geq \gamma_3\}.$$

For any specific realization $a \in \mathcal{A}$, let $y(a;t,x)$ denote the unique state trajectory corresponding to that coefficient. We transform the deterministic problem by taking the expectation of the quadratic cost with respect to the measure π:

$$\begin{aligned} J_\pi(u) &= \mathbb{E}_{a\sim\pi}[J(u)] \\ &= \int_{\mathcal{A}} \left[\int_0^T \left(\int_0^L y^2(a;t,x)\,dx + \gamma_1 \int_0^L u^2(t,x)\,dx \right) dt \right. \\ &\qquad \left. + \gamma_2 \int_0^L y^2(a;T,x)\,dx \right] d\pi(a). \end{aligned} \tag{3}$$

For the specific scope of this article, we model the probability measure π using a Shifted Gamma distribution. Specifically, we assume that a random variable X follows a standard Gamma distribution, $X \sim \text{Gamma}(\alpha, \beta)$, and define our system parameter as $a = X + \gamma_3$. This specific choice of continuous distribution rigorously enforces the uniform ellipticity constraint $a \in [\gamma_3, \infty)$ while enabling analytical tractability for the subsequent dynamic programming solution.

3. Preliminary results

In this section, we introduce the foundational mathematical concepts and established properties that are strictly necessary to formulate and resolve the averaged optimal control problem.

Statement 1 (Bellman's Principle of Optimality, [8])**.** *An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.*

For a dynamical system evolving over a time horizon $[0, T]$, let $S[t, y]$ denote the optimal value functional representing the minimum accumulated cost from an initial state y at time t to the terminal time T. Dynamic Programming considers a family of optimal control problems where the value function represents the minimum of the cost functional [3]. The principle of optimality dictates that for any small time increment $\Delta t > 0$, the optimal control policy must satisfy the recursive relationship:

$$S[t, y] = \min_{u} \left\{ \text{Cost}(t, t + \Delta t) + S[t + \Delta t, y + \Delta y] \right\}, \tag{4}$$

where Δy is the state trajectory increment resulting from the applied control u over the interval $[t, t + \Delta t]$. This principle forms the basis for deriving the Hamilton-Jacobi-Bellman (HJB) equation, whose solution provides the optimal feedback control [3, 7].

Definition 1. Let $S[t, y]$ be a continuous value functional defined over a state profile $y \in L_2(\mathcal{A}, \pi)$. Assuming S possesses a Fréchet differential with respect to the state variable, the linear principal part of the functional's increment corresponding to a small state perturbation Δy is denoted as $dS[t, y, \Delta y]$. By the Riesz representation theorem, this differential can be uniquely represented as an inner product with a functional gradient $v(a, t) \in L_2(\mathcal{A}, \pi)$:

$$dS[t, y, \Delta y] = \int_{\mathcal{A}} v(a, t) \Delta y(a; t) \, d\pi(a). \tag{5}$$

Definition 2. To facilitate the pointwise evaluation of the state within double integrals over the parameter space, we introduce the measure-specific Dirac delta distribution, denoted by $\delta_\pi(a - b)$. It is defined strictly via its sifting property under the probability measure π. For any continuous test function f defined on $\mathcal{A}$, it satisfies:

$$\int_{\mathcal{A}} f(b) \delta_\pi(a - b) \, d\pi(b) = f(a). \tag{6}$$

Lemma 1 (Gautschi W., [11])**.** *A fundamental property of any system of orthonormal polynomials is that multiplication by the independent variable is governed by a three-term recurrence relation.*

Let $\{\psi_i(a)\}_{i=0}^{N}$ be a sequence of orthonormal polynomials associated with the probability measure π. Operation from lemma 1 can be exactly represented by a symmetric, tridiagonal Jacobi matrix Λ:

$$a\psi_i(a) = \sum_{k=0}^{N} \Lambda_{ki}\psi_k(a). \tag{7}$$

4. Main results

To render the infinite-dimensional averaged optimal control problem (1), (3) computationally tractable, we propose a rigorous three-step mathematical transformation. First, we apply spectral decomposition to resolve the spatial operators, decoupling the partial differential equation into an infinite system of parametric ordinary differential equations (ODEs). Second, we apply the Bellman optimality principle to these decoupled systems, deriving the exact optimal feedback law governed by a singular Integro-Differential Riccati Equation (IDRE). Third, we eliminate the stochastic parametric uncertainty by projecting the continuous IDRE into a deterministic, finite-dimensional matrix format using a Spectral Galerkin approach based on orthogonal polynomials.

4.1. Spectral decomposition

To transform the partial differential equation (1) into a computationally manageable form, we employ a spectral decomposition method [13]. We represent the parameter-dependent state $y(a;t,x)$ and the control $u(t,x)$ as series expansions using the eigenfunctions of the Dirichlet-Laplacian operator on the interval $[0,L]$. These eigenfunctions are given by

$$\phi_k(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{k\pi x}{L}\right), \quad k = 1, 2, \dots . \tag{8}$$

The set $\{\phi_k\}_{k\geq 1}$ forms an orthonormal basis for $L^2(0,L)$. This allows us to express the state and control as follows:

$$y(a;t,x) = \sum_{k=1}^{\infty} y_k(a;t)\phi_k(x), \quad u(t,x) = \sum_{k=1}^{\infty} u_k(t)\phi_k(x), \tag{9}$$

where $y_k(a;t)$ and $u_k(t)$ are the time-dependent spectral coefficients. Note that while the state coefficients depend on the specific realization of the unknown parameter a, the control coefficients depend only on time, reflecting the fact that the controller must act without perfect knowledge of a.

By substituting the expansions (9) into (1), and utilizing the fact that $-\frac{\partial^2\phi_n}{\partial x^2} = \lambda_n\phi_n$, where $\lambda_n = \left(\frac{n\pi}{L}\right)^2$, we obtain:

$$\sum_{n=1}^{\infty} \left(\dot{y}_n(a;t) + a\lambda_n y_n(a;t)\right)\phi_n(x) = \sum_{n=1}^{\infty} u_n(t)\phi_n(x). \tag{10}$$

To isolate the dynamics of a specific spectral component k, we multiply the equation by $\phi_k(x)$ and integrate over $[0, L]$:

$$\sum_{n=1}^{\infty} (\dot{y}_n(a;t) + a\lambda_n y_n(a;t)) \int_0^L \phi_n(x)\phi_k(x)\,dx = \sum_{n=1}^{\infty} u_n(t) \int_0^L \phi_n(x)\phi_k(x)\,dx. \tag{11}$$

Using the orthonormality property, the integral $\int_0^L \phi_n(x)\phi_k(x)\,dx$ equals the Kronecker delta δ_{nk} (which is 1 if $n = k$ and 0 otherwise). This causes all terms in the sums to vanish except for $n = k$, yielding a set of decoupled parametric ordinary differential equations:

$$\dot{y}_k(a;t) + a\lambda_k y_k(a;t) = u_k(t), \quad k = 1, 2, \ldots . \tag{12}$$

The initial spectral coefficients $y_{0,k}$ are determined by projecting the initial state $y_0(x)$ onto the orthonormal basis:

$$y_k(a;0) = y_{0,k} = \int_0^L y_0(x)\phi_k(x)\,dx. \tag{13}$$

The averaged cost functional (3) can also be expressed in terms of the spectral coefficients. Using the orthonormality of the basis $\{\phi_k\}$, the spatial integrals simplify via Parseval's identity:

$$\|y(a;t,\cdot)\|^2_{L^2(0,L)} = \int_0^L \left(\sum_{k=1}^{\infty} y_k(a;t)\phi_k(x)\right)^2 dx = \sum_{k=1}^{\infty} y_k^2(a;t), \tag{14}$$

$$\|u(t,\cdot)\|^2_{L^2(0,L)} = \int_0^L \left(\sum_{k=1}^{\infty} u_k(t)\phi_k(x)\right)^2 dx = \sum_{k=1}^{\infty} u_k^2(t). \tag{15}$$

Applying this to all terms in (3), the expected cost functional becomes an integral over the system's spectral components:

$$J_\pi(u) = \int_{\mathcal{A}} \left[\int_0^T \sum_{k=1}^{\infty} \left(y_k^2(a;t) + \gamma_1 u_k^2(t)\right) dt + \gamma_2 \sum_{k=1}^{\infty} y_k^2(a;T)\right] d\pi(a). \tag{16}$$

After interchanging summation and integration, the global averaged cost functional completely decomposes into a sum of independent component-wise contributions:

$$J_\pi(u) = \sum_{k=1}^{\infty} \underbrace{\int_{\mathcal{A}} \left[\int_0^T \left(y_k^2(a;t) + \gamma_1 u_k^2(t)\right) dt + \gamma_2 y_k^2(a;T)\right] d\pi(a)}_{J_{\pi,k}(u_k)}. \tag{17}$$

An essential consequence of this spectral decomposition is the separability of the control problem under uncertainty. Since the parametric state dynamics (12), initial values (13), and the averaged cost functional (17) are decoupled across the spectral components, the global optimization problem can be reduced to an infinite set of independent sub-problems:

$$\min_u J_\pi(u) = \sum_{k=1}^{\infty} \min_{u_k} J_{\pi,k}(u_k). \tag{18}$$

4.2. Dynamic programming

Following the spectral decomposition, the averaged problem decouples into a separate optimal control problem for each individual component:

$$\begin{cases} \dot{y}_k(a;t) + a\lambda_k y_k(a;t) = u_k(t), \\ y_k(a;0) = y_{0,k}, \end{cases} \tag{19}$$

where the initial coefficient $y_{0,k}$ is computed via the projection

$$y_{0,k} = \int_0^L y_0(x)\phi_k(x)\,dx.$$

The corresponding cost functional is defined as:

$$J_k(u_k) = \int_{\mathcal{A}} \left[\int_0^T \left(y_k^2(a;t) + \gamma_1 u_k^2(t) \right) dt + \gamma_2 y_k^2(a;T) \right] d\pi(a). \tag{20}$$

For notational simplicity, index k will be omitted throughout the remainder of this section.

In accordance with Bellman's Principle of Optimality (Statement 1), we introduce the value functional:

$$S[t,y] = \min_u \left\{ \int_{\mathcal{A}} \gamma_2 y^2(a;T)\,d\pi(a) + \int_t^T \int_{\mathcal{A}} \left[y^2(a;\tau) + \gamma_1 u^2(\tau) \right] d\pi(a)\,d\tau \right\}, \tag{21}$$

where $t \in [0,T]$ is an arbitrary moment in time.

Introducing a small time increment Δt such that $t' = t + \Delta t$, and denoting the corresponding state trajectory perturbation by Δy, evaluated pointwise as $\Delta y(a;t) = y(a;t') - y(a;t)$, the value functional at the perturbed state evaluates to:

$$S[t', y + \Delta y] = S[t + \Delta t, y + \Delta y]. \tag{22}$$

Assuming that the value functional S is continuously differentiable with respect to time t and possesses a Fréchet differential with respect to the state variable y, we obtain:

$$\begin{aligned} S[t', y + \Delta y] &= S[t,y] + \frac{\partial S[t, y + \Delta y]}{\partial t}\Delta t + \Phi(t, y, \Delta y) \\ &\quad + o(\Delta t) + \omega(t, y, \Delta y) = S[t,y] + \frac{\partial S[t,y]}{\partial t}\Delta t + \Phi(t, y, \Delta y) \\ &\quad + \left[\frac{\partial S[t, y + \Delta y]}{\partial t} - \frac{\partial S[t,y]}{\partial t} \right] \Delta t + o(\Delta t) + \omega(t, y, \Delta y). \end{aligned} \tag{23}$$

Here, $\Phi(t, y, \Delta y)$ represents the Fréchet differential of the functional S with respect to the state y, evaluated at the point (t,y) acting on the increment Δy.

Applying the mean value theorem (Lagrange's formula) to the time derivative of the functional, we can express the difference as:

$$\begin{aligned} \frac{\partial S[t, y + \Delta y]}{\partial t} - \frac{\partial S[t,y]}{\partial t} &= \Phi_1(t, y + \theta \Delta y, \Delta y) \\ &= \Phi_1(t, y, \Delta y) + \omega_1(t, y, \Delta y), \end{aligned} \tag{24}$$

where $\theta \in (0,1)$. Substituting this intermediate result back into the expansion, we obtain the refined expression for the value functional at the perturbed state:

$$S[t+\Delta t, y+\Delta y] = S[t,y] + \frac{\partial S[t,y]}{\partial t}\Delta t + dS[t,y,\Delta y] \\ + o(\Delta t) + \omega_2(t,y,\Delta y), \tag{25}$$

where $dS[t,y,\Delta y]$ represents the total Fréchet differential, and the remainder terms satisfy the standard asymptotic decay conditions:

$$\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0 \quad \text{and} \quad \lim_{\|\Delta y\| \to 0} \frac{\omega_2(t,y,\Delta y)}{\|\Delta y\|} = 0. \tag{26}$$

Taking into account the definition of the value functional in (21), we can partition the time horizon $[t,T]$ into sub-intervals $[t, t+\Delta t]$ and $[t+\Delta t, T]$. By applying the additive property of integrals, we have:

$$\begin{aligned} S[t,y] &= \min_{u \in L^2(t,T)} \left\{ \int_{\mathcal{A}} \gamma_2 y^2(a;T)\, d\pi(a) \right. \\ &\quad + \int_t^{t+\Delta t} \int_{\mathcal{A}} \left[y^2(a;\tau) + \gamma_1 u^2(\tau) \right] d\pi(a)\, d\tau \\ &\quad \left. + \int_{t+\Delta t}^{T} \int_{\mathcal{A}} \left[y^2(a;\tau) + \gamma_1 u^2(\tau) \right] d\pi(a)\, d\tau \right\} \\ &= \min_{u \in L^2(t,t+\Delta t)} \left\{ \int_t^{t+\Delta t} \int_{\mathcal{A}} \left[y^2(a;\tau) + \gamma_1 u^2(\tau) \right] d\pi(a)\, d\tau \right. \\ &\quad + \min_{u \in L^2(t+\Delta t,T)} \left[\int_{\mathcal{A}} \gamma_2 y^2(a;T)\, d\pi(a) \right. \\ &\qquad \left.\left. + \int_{t+\Delta t}^{T} \int_{\mathcal{A}} \left[y^2(a;s) + \gamma_1 u^2(s) \right] d\pi(a)\, ds \right] \right\} \\ &= \min_{u \in L^2(t,t+\Delta t)} \left\{ \int_t^{t+\Delta t} \int_{\mathcal{A}} \left[y^2(a;\tau) + \gamma_1 u^2(\tau) \right] d\pi(a)\, d\tau \right. \\ &\quad \left. + S[t+\Delta t, y+\Delta y] \right\}. \end{aligned} \tag{27}$$

Substituting the expanded form of the value functional $S[t+\Delta t, y+\Delta y]$ from (25) into the recursive Bellman relationship (27), and rearranging the terms, we deduce that:

$$-\frac{\partial S[t,y]}{\partial t}\Delta t = \min_{u \in L^2(t,t+\Delta t)} \left\{ \int_t^{t+\Delta t} \int_{\mathcal{A}} \left[y^2(a;\tau) + \gamma_1 u^2(\tau) \right] d\pi(a)\, d\tau \right. \\ \left. + dS[t,y,\Delta y] + o(\Delta t) + \omega_2(t,y,\Delta y) \right\}. \tag{28}$$

Assuming the state trajectory $y(a;\cdot)$ is continuous in time and square-integrable over the parameter space, the increment $\Delta y(a;t)$ resides in the weighted space $L_2(\mathcal{A},\pi)$ for all $t \in [0,T]$. Consequently, applying the Riesz representation established in Definition 1, the Fréchet differential can be expressed as an inner product:

$$dS[t,y,\Delta y] = \int_{\mathcal{A}} v(a,t)\Delta y(a;t)\, d\pi(a), \tag{29}$$

where $v(a,t) \in L_2(\mathcal{A},\pi)$ is the functional gradient of S evaluated at (t,y), defined for almost all $t \in [0,T]$. Substituting this integral representation into (28), we obtain:

$$\begin{aligned} -\frac{\partial S[t,y]}{\partial t}\Delta t = \min_{u\in L^2(t,t+\Delta t)} \Bigg\{ &\int_t^{t+\Delta t} \int_{\mathcal{A}} \left[y^2(a;\tau) + \gamma_1 u^2(\tau)\right] d\pi(a)\, d\tau \\ &+ \int_{\mathcal{A}} v(a,t)\Delta y(a;t)\, d\pi(a) + o(\Delta t) + \omega_2(t,y,\Delta y)\Bigg\}. \end{aligned} \tag{30}$$

Utilizing the state dynamics given in (19), the state increment can be linearly approximated to first order as:

$$\Delta y(a;t) \approx \dot{y}(a;t)\Delta t = \left(-a\lambda y(a;t) + u(t)\right)\Delta t. \tag{31}$$

It is important to note that the state increment $\Delta y(a;t)$ is governed strictly by the temporal evolution of the system. Because the dynamic programming principle evaluates the trajectory of the infinite-dimensional state profile $y(\cdot;t)$ over the time interval $[t, t+\Delta t]$, the spatial parameter $a \in \mathcal{A}$ effectively acts as a fixed coordinate index. Consequently, no spatial variation with respect to a is required during this step.

Substituting this approximation (31) into (30), we obtain:

$$\begin{aligned} -\frac{\partial S[t,y]}{\partial t}\Delta t = \min_{u\in L^2(t,t+\Delta t)} \Bigg\{ &\int_t^{t+\Delta t} \int_{\mathcal{A}} \left[y^2(a;\tau) + \gamma_1 u^2(\tau)\right] d\pi(a)\, d\tau \\ + \int_{\mathcal{A}} v(a,t) &\left(-a\lambda y(a;t) + u(t)\right)\Delta t\, d\pi(a) + o(\Delta t) + \omega_2(t,y,\Delta y)\Bigg\}. \end{aligned} \tag{32}$$

Dividing by Δt and passing to the limit as $\Delta t \to 0$ in (32), we derive the Hamilton-Jacobi-Bellman (HJB) equation:

$$\begin{aligned} -\frac{\partial S[t,y]}{\partial t} = \min_u \Bigg\{ &\int_{\mathcal{A}} \left[y^2(a;t) + \gamma_1 u^2(t)\right]\, d\pi(a) \\ &+ \int_{\mathcal{A}} v(a,t) \left(-a\lambda y(a;t) + u(t)\right)\, d\pi(a)\Bigg\}, \end{aligned} \tag{33}$$

where the equality holds for almost all $t \in [0,T]$. In subsequent equations of the form (33), we will use the standard equality sign to imply equality almost everywhere.

Equation (33) constitutes the desired Hamilton-Jacobi-Bellman equation for the optimal control problem under consideration. Because $v(a,t)$ represents the gradient of the value functional S, (33) is inherently an equation in functional derivatives. It follows directly from the definition of the functional S in (21) that $S \geq 0$, and evaluating it at the boundary time T yields:

$$S[T, y] = \int_{\mathcal{A}} \gamma_2 y^2(a; T)\, d\pi(a). \tag{34}$$

Thus, the optimal control problem is now reduced to finding the control policy u and the value functional S that satisfy the Bellman equation (33) subject to the terminal condition (34), ensuring that the functional S remains non-negative.

Isolating the terms within the minimization operator of (33) that depend explicitly on the control effort $u(t)$, we obtain the following expression:

$$\int_{\mathcal{A}} \left(\gamma_1 u^2(t) + v(a,t)u(t)\right)\, d\pi(a). \tag{35}$$

To determine the optimal control $u(t)$ that minimizes the objective, we differentiate the expression in (35) with respect to $u(t)$ and apply the first-order necessary condition for optimality:

$$2\gamma_1 u(t) + \int_{\mathcal{A}} v(a,t)\, d\pi(a) = 0. \tag{36}$$

Solving (36) for the optimal control $u(t)$ yields:

$$u(t) = -\frac{1}{2\gamma_1} \int_{\mathcal{A}} v(a,t)\, d\pi(a). \tag{37}$$

Substituting the optimal control policy (37) into the Hamilton-Jacobi-Bellman equation (33), we obtain the purely state-dependent HJB equation:

$$\begin{aligned} &-\frac{\partial S[t,y]}{\partial t} \\ &\quad = \int_{\mathcal{A}} \left[y^2(a;t) - v(a,t)a\lambda y(a;t)\right]\, d\pi(a) - \frac{1}{4\gamma_1}\left(\int_{\mathcal{A}} v(a,t)\, d\pi(a)\right)^2. \end{aligned} \tag{38}$$

Given the linear dynamics and quadratic cost, we postulate that the value functional assumes a purely quadratic form characterized by a spatially symmetric kernel $P(a,b,t) = P(b,a,t)$:

$$S[t,y] = \int_{\mathcal{A}} \int_{\mathcal{A}} P(a,b,t) y(a;t) y(b;t)\, d\pi(a)\, d\pi(b). \tag{39}$$

Taking the partial derivative of S with respect to time affects only the kernel, as the state y is treated as an independent variable when evaluating the partial time derivative of the functional:

$$\frac{\partial S[t,y]}{\partial t} = \int_{\mathcal{A}} \int_{\mathcal{A}} \frac{\partial P(a,b,t)}{\partial t} y(a;t) y(b;t)\, d\pi(a)\, d\pi(b). \tag{40}$$

To determine the gradient function $v(a,t)$, we first compute the Fréchet differential with respect to a small state perturbation $\Delta y(a;t)$. Substituting the perturbed state $y + \Delta y$ into S and isolating the linear, first-order terms yields:

$$dS[t, y, \Delta y] = \int_{\mathcal{A}} \int_{\mathcal{A}} P(a, b, t) \big[y(a; t)\Delta y(b; t) + y(b; t)\Delta y(a; t) \big] \, d\pi(a) \, d\pi(b). \quad (41)$$

Because the kernel $P(a, b, t)$ is symmetric with respect to a and b, both terms inside the bracket evaluate to identical double integrals. We can therefore combine them:

$$dS[t, y, \Delta y] = 2 \int_{\mathcal{A}} \int_{\mathcal{A}} P(a, b, t) y(b; t) \Delta y(a; t) \, d\pi(a) \, d\pi(b). \quad (42)$$

By definition, the Riesz representation of the Fréchet differential is given by:

$$dS[t, y, \Delta y] = \int_{\mathcal{A}} v(a, t) \Delta y(a; t) \, d\pi(a). \quad (43)$$

By isolating the term multiplying $\Delta y(a; t) \, d\pi(a)$ in (42) and comparing it with (43), we extract the gradient:

$$v(a, t) = 2 \int_{\mathcal{A}} P(a, b, t) y(b; t) \, d\pi(b). \quad (44)$$

We now substitute the time derivative (40) and the gradient (44) back into the purely state-dependent Hamilton-Jacobi-Bellman equation (38). Our goal is to express all terms as double integrals over the measure spaces $d\pi(a) \, d\pi(b)$ by performing three sequential transformations.

First, we represent the squared state term using the measure-specific Dirac delta distribution $\delta_\pi(a - b)$. Applying its sifting property (Definition 2) yields:

$$\int_{\mathcal{A}} y^2(a; t) \, d\pi(a) = \int_{\mathcal{A}} \int_{\mathcal{A}} \delta_\pi(a - b) y(a; t) y(b; t) \, d\pi(a) \, d\pi(b). \quad (45)$$

Second, we substitute $v(a, t)$ into the cross-coupling term. We can symmetrize the linear multiplier $a\lambda$ into $\frac{1}{2}\lambda(a + b)$ by exploiting the symmetry of both the kernel $P(a, b, t) = P(b, a, t)$ and the state product $y(a; t)y(b; t)$. This transformation results in

$$\begin{aligned} - \int_{\mathcal{A}} & v(a, t) a\lambda y(a; t) \, d\pi(a) \\ &= -2 \int_{\mathcal{A}} \int_{\mathcal{A}} P(a, b, t) a\lambda y(a; t) y(b; t) \, d\pi(a) \, d\pi(b) \\ &= - \int_{\mathcal{A}} \int_{\mathcal{A}} \lambda(a + b) P(a, b, t) y(a; t) y(b; t) \, d\pi(a) \, d\pi(b). \end{aligned} \quad (46)$$

Third, we expand the squared integral term by introducing the macroscopic aggregated gain $R(a, t)$, which represents the projection of the kernel over the parameter space. It is defined as

$$R(a, t) = \int_{\mathcal{A}} P(a, b, t) \, d\pi(b). \quad (47)$$

The integral of the gradient over $\mathcal{A}$ simplifies to $2\int_{\mathcal{A}} R(b,t)y(b;t)\,d\pi(b)$. By squaring this integral and absorbing the $-\frac{1}{4\gamma_1}$ coefficient, we obtain

$$-\frac{1}{4\gamma_1}\left(\int_{\mathcal{A}} v(a,t)\,d\pi(a)\right)^2 = -\frac{1}{\gamma_1}\int_{\mathcal{A}}\int_{\mathcal{A}} R(a,t)R(b,t)y(a;t)y(b;t)\,d\pi(a)\,d\pi(b). \tag{48}$$

Finally, substituting (45), (46), and (48) into the HJB equation and collecting all terms under the common double integral, we obtain the unified integral representation:

$$\begin{aligned} -\int_{\mathcal{A}}\int_{\mathcal{A}} &\frac{\partial P(a,b,t)}{\partial t} y(a;t)y(b;t)\,d\pi(a)\,d\pi(b) \\ = \int_{\mathcal{A}}\int_{\mathcal{A}} &\left[\delta_\pi(a-b) - \lambda(a+b)P(a,b,t) - \frac{1}{\gamma_1}R(a,t)R(b,t)\right] \\ \times\, &y(a;t)y(b;t)\,d\pi(a)\,d\pi(b). \end{aligned} \tag{49}$$

Because the equality in (49) must hold true for any arbitrary state trajectory $y(\cdot;t)$, the corresponding kernels must be identical. This yields the following Integro-Differential Riccati Equation:

$$-\frac{\partial P(a,b,t)}{\partial t} = \delta_\pi(a-b) - \lambda(a+b)P(a,b,t) - \frac{1}{\gamma_1}R(a,t)R(b,t). \tag{50}$$

To determine the appropriate boundary condition for this kernel, we can rewrite the terminal cost from (34) using the measure-specific Dirac delta function as a double integral:

$$S[T,y] = \int_{\mathcal{A}}\int_{\mathcal{A}} \gamma_2\delta_\pi(a-b)y(a;T)y(b;T)\,d\pi(a)\,d\pi(b). \tag{51}$$

By comparing this expression with our postulated quadratic form (39) evaluated at the terminal time $t = T$, we establish the identity:

$$\int_{\mathcal{A}}\int_{\mathcal{A}} P(a,b,T)y(a;T)y(b;T)\,d\pi(a)\,d\pi(b) = \int_{\mathcal{A}}\int_{\mathcal{A}} \gamma_2\delta_\pi(a-b)y(a;T)y(b;T)\,d\pi(a)\,d\pi(b). \tag{52}$$

Since this relationship must hold for any terminal state profile $y(\cdot;T)$, we extract the terminal condition for the Riccati kernel:

$$P(a,b,T) = \gamma_2\delta_\pi(a-b). \tag{53}$$

4.3. Spectral Galerkin approach

Solving the infinite-dimensional Integro-Differential Riccati Equation (50) directly is computationally prohibitive, primarily due to the continuous parameter space $\mathcal{A}$ and the presence of the singular source term $\delta_\pi(a-b)$. We employ a Spectral Galerkin approach to transform the continuous infinite-dimensional problem into a finite-dimensional standard Matrix Riccati Differential Equation

(MRDE). We approximate the kernel $P(a, b, t)$ using a Generalized Polynomial Chaos (gPC) expansion [10]. Following the Spectral Galerkin framework [9], let $\{\psi_i(a)\}_{i=0}^N$ be a set of orthogonal polynomials associated with the probability measure π. Because the unknown parameter follows a Shifted Gamma distribution, these correspond to Shifted Generalized Laguerre polynomials [10].

The polynomials satisfy the orthogonality condition:

$$\int_{\mathcal{A}} \psi_i(a)\psi_j(a)\, d\pi(a) = \delta_{ij}, \tag{54}$$

where δ_{ij} is the Kronecker delta, defined as $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. Furthermore, since π is a probability measure, the zero-th polynomial is exactly one, $\psi_0(a) = 1$.

We postulate that the Riccati kernel can be approximated by a truncated double series expansion up to degree N:

$$P(a, b, t) \approx \sum_{i=0}^{N}\sum_{j=0}^{N} C_{ij}(t)\psi_i(a)\psi_j(b), \tag{55}$$

where $C(t) \in \mathbb{R}^{(N+1)\times(N+1)}$ is a symmetric matrix of time-dependent coefficients.

Substituting (55) into the aggregated gain $R(a, t)$ defined in (47), and utilizing the fact that $d\pi(b) = \psi_0(b)\, d\pi(b)$, we obtain:

$$R(a, t) = \int_{\mathcal{A}} \left[\sum_{i=0}^{N}\sum_{j=0}^{N} C_{ij}(t)\psi_i(a)\psi_j(b)\right] \psi_0(b)\, d\pi(b). \tag{56}$$

Due to the orthogonality condition (54), the integration over b eliminates all terms except for $j = 0$. The integral therefore collapses entirely, yielding

$$R(a, t) = \sum_{i=0}^{N} C_{i0}(t)\psi_i(a). \tag{57}$$

This demonstrates that $R(a, t)$ is entirely determined by the first column of the coefficient matrix $C(t)$.

We now substitute the expansion (55) back into the original equation (50). To isolate the coefficient matrix $C(t)$, we apply the Galerkin projection by multiplying the entire equation by the test function $\psi_m(a)\psi_n(b)$ and integrating over both spatial parameters a and b with respect to the measure π.

We will evaluate these projected terms individually.

First, substituting the series expansion into the time derivative and interchanging the summation and integration operations yields:

$$
\begin{aligned}
&-\int_{\mathcal{A}}\int_{\mathcal{A}} \frac{\partial P(a,b,t)}{\partial t}\psi_m(a)\psi_n(b)\,d\pi(a)\,d\pi(b) \\
&= -\sum_{i=0}^{N}\sum_{j=0}^{N}\dot{C}_{ij}(t)\left(\int_{\mathcal{A}}\psi_i(a)\psi_m(a)\,d\pi(a)\right)\left(\int_{\mathcal{A}}\psi_j(b)\psi_n(b)\,d\pi(b)\right) \\
&= -\sum_{i=0}^{N}\sum_{j=0}^{N}\dot{C}_{ij}(t)\delta_{im}\delta_{jn} = -\dot{C}_{mn}(t). \qquad (58)
\end{aligned}
$$

Second, the singular Dirac delta term seamlessly evaluates to the identity matrix I, as shown by

$$
\begin{aligned}
&\int_{\mathcal{A}}\int_{\mathcal{A}}\delta_\pi(a-b)\psi_m(a)\psi_n(b)\,d\pi(a)\,d\pi(b) \qquad (59)\\
&= \int_{\mathcal{A}}\psi_m(a)\left(\int_{\mathcal{A}}\delta_\pi(a-b)\psi_n(b)\,d\pi(b)\right)d\pi(a) \\
&= \int_{\mathcal{A}}\psi_m(a)\psi_n(a)\,d\pi(a) = \delta_{mn}.
\end{aligned}
$$

Third, the linear coupling term involves multiplication by the independent variables a and b. We separate the integral into two symmetric components and substitute the series expansion. As established in Lemma 1, multiplication by the independent variable is governed by a three-term recurrence relation, which can be exactly represented by a symmetric, tridiagonal Jacobi matrix Λ. This allows us to compactly express the multiplication as in formulla (7), where the tridiagonal structure of Λ guarantees that at most three terms are non-zero for any degree i.

Evaluating the first component corresponding to the variable a, we group the integrals and apply the recurrence substitution to obtain

$$
\begin{aligned}
&\int_{\mathcal{A}}\int_{\mathcal{A}} aP(a,b,t)\psi_m(a)\psi_n(b)\,d\pi(a)\,d\pi(b) \\
&= \sum_{i=0}^{N}\sum_{j=0}^{N}C_{ij}(t)\left(\int_{\mathcal{A}}a\psi_i(a)\psi_m(a)\,d\pi(a)\right)\left(\int_{\mathcal{A}}\psi_j(b)\psi_n(b)\,d\pi(b)\right) \\
&= \sum_{i=0}^{N}\sum_{j=0}^{N}C_{ij}(t)\left(\int_{\mathcal{A}}\sum_{k=0}^{N}\Lambda_{ki}\psi_k(a)\psi_m(a)\,d\pi(a)\right)\delta_{jn} \\
&= \sum_{i=0}^{N}C_{in}(t)\left(\sum_{k=0}^{N}\Lambda_{ki}\delta_{km}\right) = \sum_{i=0}^{N}\Lambda_{mi}C_{in}(t) = \big(\Lambda C(t)\big)_{mn}.
\end{aligned}
$$

By exact mathematical symmetry, evaluating the second component corresponding to the variable b yields the transposed multiplication $(C(t)\Lambda)_{mn}$. Combining both evaluated components and applying the constant multiplier

$-\lambda$ produces the final projected linear term:

$$-\lambda \int_{\mathcal{A}} \int_{\mathcal{A}} (a+b) P(a,b,t) \psi_m(a) \psi_n(b)\, d\pi(a)\, d\pi(b)$$
$$= -\lambda \big(\Lambda C(t) + C(t) \Lambda \big)_{mn}. \tag{60}$$

Finally, we evaluate the quadratic term. We substitute the collapsed series representation of the aggregated gain from (57) into the integral for both $R(a,t)$ and $R(b,t)$. By exploiting the symmetry of the coefficient matrix ($C_{j0} = C_{0j}$), rearranging the summations, and applying the orthogonality condition (54), the integrals completely decouple and collapse:

$$\begin{aligned}
&\int_{\mathcal{A}} \int_{\mathcal{A}} R(a,t) R(b,t) \psi_m(a) \psi_n(b)\, d\pi(a)\, d\pi(b) \\
&= \int_{\mathcal{A}} \int_{\mathcal{A}} \left(\sum_{i=0}^{N} C_{i0}(t) \psi_i(a) \right) \left(\sum_{j=0}^{N} C_{0j}(t) \psi_j(b) \right) \psi_m(a) \psi_n(b)\, d\pi(a)\, d\pi(b) \\
&= \sum_{i=0}^{N} \sum_{j=0}^{N} C_{i0}(t) C_{0j}(t) \left(\int_{\mathcal{A}} \psi_i(a) \psi_m(a)\, d\pi(a) \right) \left(\int_{\mathcal{A}} \psi_j(b) \psi_n(b)\, d\pi(b) \right) \\
&= \sum_{i=0}^{N} \sum_{j=0}^{N} C_{i0}(t) C_{0j}(t) \delta_{im} \delta_{jn} = C_{m0}(t) C_{0n}(t).
\end{aligned} \tag{61}$$

Assembling these projected components yields a finite-dimensional, standard Matrix Riccati Differential Equation (MRDE) governing the coefficient matrix $C(t)$:

$$-\dot{C}(t) = I - \lambda \Lambda C(t) - \lambda C(t) \Lambda - \frac{1}{\gamma_1} C(t) E_{00} C(t), \tag{62}$$

where I is the $(N+1) \times (N+1)$ identity matrix and E_{00} is a sparse matrix of the same dimension with $(E_{00})_{ij} = 1$ if $i = j = 0$ and zero otherwise.

To complete the formulation of the optimal control problem, we must establish the terminal boundary condition for the Matrix Riccati Differential Equation (62). We apply the same Galerkin projection to the continuous terminal condition $P(a,b,T) = \gamma_2 \delta_\pi(a-b)$.

Substituting the series expansion $C_{mn}(T)$ on the left-hand side and evaluating the integral of the Dirac delta distribution on the right-hand side using the sifting property (Definition 2), we obtain

$$\begin{aligned}
C_{mn}(T) &= \int_{\mathcal{A}} \int_{\mathcal{A}} P(a,b,T) \psi_m(a) \psi_n(b)\, d\pi(a)\, d\pi(b) \\
&= \gamma_2 \int_{\mathcal{A}} \int_{\mathcal{A}} \delta_\pi(a-b) \psi_m(a) \psi_n(b)\, d\pi(a)\, d\pi(b) \\
&= \gamma_2 \int_{\mathcal{A}} \psi_m(a) \left(\int_{\mathcal{A}} \delta_\pi(a-b) \psi_n(b)\, d\pi(b) \right) d\pi(a) \\
&= \gamma_2 \int_{\mathcal{A}} \psi_m(a) \psi_n(a)\, d\pi(a) = \gamma_2 \delta_{mn}.
\end{aligned} \tag{63}$$

Because the Kronecker delta δ_{mn} defines the entries of the identity matrix, this element-wise equality translates directly to the finite-dimensional matrix boundary condition

$$C(T) = \gamma_2 I. \tag{64}$$

Equation (62), combined with the terminal condition (64), constitutes a standard, finite-dimensional Matrix Riccati Differential Equation (MRDE). Because the system is now strictly finite-dimensional and well-posed, the coefficient matrix $C(t)$ can be efficiently computed by integrating the MRDE backward in time from $t = T$ to $t = 0$. As extensively documented in the literature for Riccati systems [12], this can be achieved using standard numerical integration techniques, such as explicit Runge-Kutta methods or backward differentiation formulas (BDF) for stiffer configurations.

Once the time-dependent matrix $C(t)$ is computed and stored offline, the optimal feedback control law is immediately accessible. Recalling that the aggregated macroscopic gain is determined entirely by the first column of the coefficient matrix, $R(a,t) = \sum_{i=0}^{N} C_{i0}(t)\psi_i(a)$, the optimal control input $u(t)$ applied to the system at any time t is explicitly given by

$$\begin{aligned} u(t) &= -\frac{1}{\gamma_1}\int_{\mathcal{A}} R(a,t)y(a;t)\,d\pi(a) \\ &= -\frac{1}{\gamma_1}\int_{\mathcal{A}} \left(\sum_{i=0}^{N} C_{i0}(t)\psi_i(a)\right) y(a;t)\,d\pi(a). \end{aligned} \tag{65}$$

This formulation completely circumvents the need to solve the singular, infinite-dimensional Integro-Differential Riccati Equation, reducing the optimal control problem to the evaluation of a finite polynomial expansion.

5. Numerical Example

To demonstrate the proposed theoretical framework, we present a numerical simulation of the averaged linear quadratic parabolic optimal control problem.

We consider the spatial domain of length $L = \pi$ over a finite time horizon $T = 1$. The initial state profile is chosen as $y_0(x) = \sin(x)$. The penalty weights in the cost functional are set to $\gamma_1 = 1$ for the control effort and $\gamma_2 = 1$ for the terminal state deviation.

The physical diffusion coefficient a is subjected to parametric uncertainty and is modeled as a Shifted Gamma random variable. To strictly satisfy the uniform ellipticity condition, we define $a = X + \gamma_3$, where the physical bound is $\gamma_3 = 1$, and the stochastic component X follows a standard Gamma distribution $X \sim \text{Gamma}(\alpha, \beta)$ with shape parameter $\alpha = 2$ and rate parameter $\beta = 1$.

Applying the spectral decomposition to the spatial domain, the initial state $y_0(x)$ is projected onto the orthonormal basis $\phi_k(x) = \sqrt{2/\pi}\sin(kx)$. Due to the orthogonality of the sine functions, only the first spectral component is non-zero:

$$y_{0,1} = \int_0^\pi \sin(x)\sqrt{\frac{2}{\pi}}\sin(x)dx = \sqrt{\frac{\pi}{2}}. \tag{66}$$

Consequently, the entire infinite-dimensional state dynamics simplify to a single active component corresponding to the eigenvalue $\lambda_1 = 1^2 = 1$.

To project the stochastic parameter space into a finite-dimensional deterministic form, we apply the Spectral Galerkin approach using Shifted Generalized Laguerre polynomials.

We truncate the polynomial chaos expansion at degree $N = 3$.

The multiplication by the stochastic parameter a is governed by the three-term recurrence relation, yielding the specific symmetric, tridiagonal 4×4 Jacobi matrix Λ:

$$\Lambda = \begin{pmatrix} 3 & -\sqrt{2} & 0 & 0 \\ -\sqrt{2} & 5 & -\sqrt{6} & 0 \\ 0 & -\sqrt{6} & 7 & -2\sqrt{3} \\ 0 & 0 & -2\sqrt{3} & 9 \end{pmatrix}. \tag{67}$$

By substituting Λ, $\lambda_1 = 1$, and the penalty weights into the projected system, the original infinite-dimensional Integro-Differential Riccati Equation is reduced to the following computationally tractable 4×4 Matrix Riccati Differential Equation:

$$-\dot{C}(t) = I - \Lambda C(t) - C(t)\Lambda - C(t)E_{00}C(t), \tag{68}$$

subject to the terminal boundary condition $C(1) = I$.

The matrix equation (68) is integrated backward in time using a standard Runge-Kutta method.

The resulting time-dependent coefficients are then used to synthesize the optimal control law and simulate the forward state dynamics. Because the system operates under parametric uncertainty, we evaluate the expected behavior of the system.

The computational results are visualized in Fig. 1.

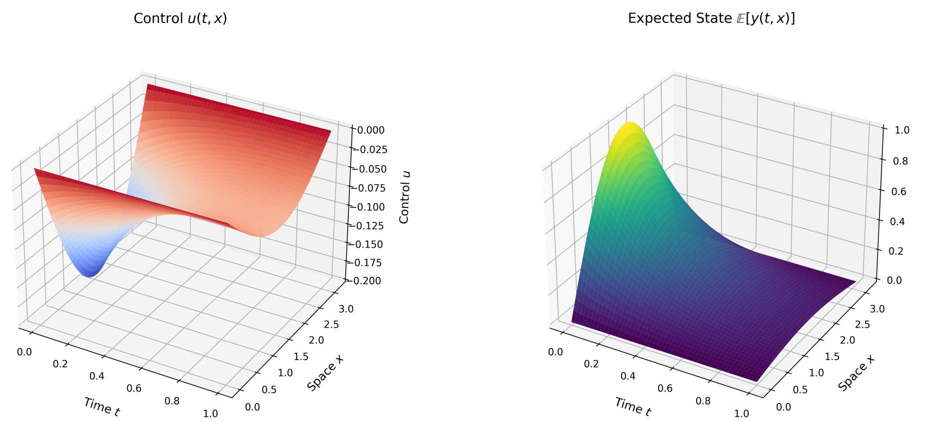

FIGURE 1. Computational results of the parabolic optimal control problem under parametric uncertainty.

As shown, the pre-computed optimal feedback control counteracts the initial state profile and decays as it approaches the terminal time $T = 1$. Consequently, the expected physical state is smoothly dissipated, successfully satisfying the objective of the LQR cost functional.

Conclusion

In this paper, we addressed the averaged Linear Quadratic Regulator (LQR) problem for a parabolic partial differential equation subject to parametric uncertainty in its diffusion dynamic.

The primary contribution of this work is the development of a rigorous, computationally tractable framework that systematically reduces an infinite-dimensional stochastic control problem into a finite-dimensional deterministic one.

This was achieved through a sequential, three-stage decomposition strategy. First, we employed spectral decomposition to eliminate the spatial dimension, projecting the PDE onto an orthonormal basis to obtain a decoupled infinite system of parametric ordinary differential equations. Second, by applying the Bellman optimality principle and dynamic programming, we established that the optimal value functional is governed by an infinite-dimensional Integro-Differential Riccati Equation (IDRE), extending classical optimal control approaches where the feedback is obtained via a standard Riccati equation. Finally, to overcome the computational intractability of the continuous parameter space and the singular Dirac delta distribution, we applied a Spectral Galerkin approach utilizing a Generalized Polynomial Chaos (gPC) expansion.

By projecting the IDRE onto a basis of Shifted Generalized Laguerre polynomials, the problem was successfully collapsed into a standard, finite-dimensional Matrix Riccati Differential Equation (MRDE). This formulation completely circumvents the need for complex numerical PDE solvers or real-time stochastic sampling. Because the system's projection matrices are highly sparse and tridiagonal, the MRDE can be solved highly efficiently in advance. The resulting coefficient matrix provides an exact, analytically derived optimal feedback control law that requires only a simple polynomial evaluation during actual system operation.

Future research will focus on extending this framework to multidimensional spatial domains and systems with multiple uncertain parameters.

Additionally, using this mathematically rigorous pre-computed solution as a baseline to guide real-time, data-driven adaptive control algorithms — such as Reinforcement Learning [5] — presents a highly promising avenue for managing complex systems operating in heavily uncertain environments.

Bridging this theoretical foundation with model-based learning algorithms could further enhance robustness without the need for exhaustive real-time reconstruction of the dynamics [3, 7].

The authors declare that there is no conflict of interest concerning the publication of this article.

References

1. Kapustian O., Laptiev O., Makarovych A. Averaging of Linear Quadratic Parabolic Optimal Control Problem. *Axioms*. 2025. Vol. 14. No. 7. P. 512.
2. Pesare A., Palladino M., Falcone M. Convergence results for an averaged LQR problem with applications to reinforcement learning. *Mathematics of Control, Signals, and Systems*. 2021. Vol. 33. No. 3. P. 379–411.
3. Alla A., Pacifico A., Palladino M., Pesare A. Online identification and control of PDEs via Reinforcement Learning methods. *Advances in computational mathematics*. 2024. Vol. 50, No. 4. https://doi.org/10.1007/s10444-024-10167-y
4. Sutton R.S., Barto A.G., Williams R.J. Reinforcement learning is direct adaptive optimal control. *IEEE Control Systems*. 1992. Vol. 12. No. 2. P. 19–22.
5. Sutton R.S., Barto A.G. Reinforcement Learning: An Introduction, 2nd ed. MIT Press, Cambridge, MA, 2018.
6. Recht B. A tour of reinforcement learning: The view from continuous control. *Annual Review of Control, Robotics, and Autonomous Systems*. 2019. Vol. 2. P. 253–279.
7. Pacifico A., Pesare A., Falcone M. A new algorithm for the LQR problem with partially unknown dynamics. *Large-Scale scientific computing*. Cham, 2022. P. 322–330.
8. Kirk D.E. Optimal control theory: an introduction. Dover Publications, 2004. 464 p.
9. Ghanem R.G., Spanos P.D. Stochastic finite elements: a spectral approach. Mineola, N.Y : Dover Publications, 2003. 222 p.
10. Xiu D., Karniadakis G.E. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing*. 2002. Vol. 24. No. 2. P. 614–644.
11. Gautschi W. Orthogonal Polynomials: Computation and Approximation. Oxford University Press, 2004.
12. Abou-Kandil H., Freiling G., Ionescu V., Jank G. Matrix Riccati Equations in Control and Systems Theory. Birkhäuser Basel, 2003.
13. Kapustyan E.A., Nakonechnyi A.G. Optimal bounded control synthesis for a parabolic boundary-value problem with fast oscillatory coefficients. *Journal of automation and information sciences*. 1999. Vol. 31, no. 12. P. 33–44.

Received: 27.02.2026 / Revised: 25.03.2026 /
Accepted: 15.04.2026 / Published: 24.04.2026