A Natural-Gradient Approach for Nonlinear Stochastic Systems with Parameter Uncertainty

Liang Cao
Department of Chemical and Biological Engineering
The University of British Columbia
2360 East Mall, Vancouver, BC V6T 1Z3, Canada
clubc19@mail.ubc.ca

Abstract

Controlling nonlinear stochastic systems with parametric uncertainty is a fundamental challenge in modern control theory. This paper presents a comprehensive theoretical framework for a natural-gradient method applied to polynomial chaos theory. We focus on quadratic regulator problems characterized by both parametric uncertainty and additive stochastic disturbances. We extend existing polynomial chaos approaches from linear systems to general nonlinear dynamics. To achieve this, we develop new mathematical tools to handle the complex interactions between nonlinearity, parameter uncertainty, and noise. The framework provides local convergence guarantees for the proposed natural gradient algorithm. Furthermore, it offers practical computational strategies while carefully characterizing the theoretical limitations in the nonlinear setting.

1 Introduction

Real-world control systems are characterized by inherent nonlinearities, parametric uncertainties, and stochastic disturbances, presenting significant challenges for controller design. While classical linear quadratic regulator (LQR) theory provides elegant solutions for linear systems (Anderson and Moore 1971), the extension to nonlinear stochastic systems with uncertain parameters remains an active area of research with limited practical solutions.

The application of polynomial chaos expansions (PCE) to control problems has emerged as a promising approach for handling parametric uncertainty. Originally developed by Wiener (Wiener 1938) and later generalized by Cameron and Martin (Cameron and Martin 1947), polynomial chaos provides a systematic framework for representing random processes as infinite series of orthogonal polynomials. Modern developments by Ghanem and Spanos (Ghanem and Spanos 1991) and Xiu and Karniadakis (Xiu and Karniadakis 2002) have established PCE as a cornerstone technique in uncertainty quantification.

In the control theory domain, polynomial chaos was applied to linear systems by Fisher and Bhattacharya (Fisher and Bhattacharya 2009), who developed the polynomial chaos theory-based linear quadratic regulator (PCT-LQR) for systems with parametric uncertainty. This work demonstrated that PCE could effectively handle random parameters in linear dynamics while maintaining computational tractability. Subsequently, Hover and Triantafyllou (Hover and Triantafyllou 2006) extended these ideas to stability analysis, showing how polynomial chaos could be used to assess robust stability margins.

However, the linear PCT-LQR framework has several fundamental limitations: (1) it requires the linearization of inherently nonlinear dynamics, potentially missing critical system behaviors; (2) the resulting controllers are often overly conservative due to worst-case assumptions; and (3) the method struggles with systems where nonlinearity and uncertainty interact strongly (Hsu and Bhattacharya 2020). The work by Peng et al. (Peng, Ghanem, and Li 2010) attempted to address nonlinear oscillators using PCE, but their approach was limited to specific system classes and lacked general convergence guarantees.

The optimal control of nonlinear stochastic systems is governed by the Hamilton-Jacobi-Bellman (HJB) equation, whose analytical solution is generally intractable (Fleming and Soner 2006). Various approximation strategies have been developed to overcome this fundamental difficulty. Classical dynamic programming methods (Bertsekas 2017) discretize the state space but suffer from the curse of dimensionality. Some progress includes approximate dynamic programming (Powell 2011) and reinforcement learning methods (Lewis et al. 2012), but these approaches typically do not explicitly handle parametric uncertainty.

H_∞ and robust control techniques (Zhou, Doyle, and Glover 1996; Basar and Bernhard 2008) provide worst-case guarantees but can be overly conservative. Distributionally robust optimization (Wiesemann et al. 2014) offers a middle ground but requires careful choice of uncertainty sets. Model predictive control with robust formulations (Mayne et al. 2011) has shown promise but faces computational challenges for nonlinear systems.

The intersection of stochastic optimal control and parametric uncertainty has received limited attention. Most existing work either treats parameters as known (Yong and Zhou 1999) or assumes they can be estimated online (Duncan and Pasik-Duncan 1990). Dual control methods attempt to balance exploration and exploitation but typically require restrictive assumptions about the uncertainty structure.

Natural gradient methods, pioneered by Amari (Amari 1998) in the context of neural networks, have recently gained attention in control theory. The key insight is to use the Fisher information metric to precondition gradient updates, leading to more efficient optimization in curved parameter spaces.

In reinforcement learning, natural policy gradients (Kakade 2001) have shown superior convergence properties compared to standard gradient methods. Subsequent work by Peters and Schaal (Peters and Schaal 2008) and Bagnell and Schneider (Bagnell and Schneider 2003) established theoretical foundations and practical algorithms. Direct applications of natural gradients to control problems have been limited. Recent work by Fazel et al. (Fazel et al. 2018) analyzed policy gradient methods for LQR problems, while Zhang et al. (Zhang et al. 2020) extended these results to output feedback. However, these approaches focus on known linear systems and do not address parametric uncertainty or nonlinearity.

Existing natural gradient methods in control suffer from several limitations that restrict their applicability to complex real-world systems. These methods are primarily designed for linear systems, which significantly constrains their utility when dealing with the inherent nonlinearities present in most practical control applications. Furthermore, they do not handle parametric uncertainty systematically, leaving practitioners without robust tools for managing the stochastic elements. The convergence analysis for these methods is often limited to convex settings, which fails to address the nonconvex optimization landscapes. Additionally, the computational complexity of these approaches grows rapidly with system dimension, creating scalability issues that prevent their application to high-dimensional systems.

These constraints uncover multiple significant deficiencies in the present understanding, underscoring the necessity for more advanced methodologies. From a theoretical perspective, no unified framework exists for combining polynomial chaos expansions with natural gradient optimization for nonlinear stochastic control problems. Methodologically, existing polynomial chaos methods are predominantly linear, while nonlinear extensions lack rigorous theoretical foundations and convergence guarantees. The computational challenges are equally pronounced, as current approaches either scale poorly with uncertainty dimension in traditional stochastic control or with system dimension in polynomial chaos methods.

This paper addresses these gaps by developing a comprehensive theoretical and computational framework for nonlinear stochastic control with parametric uncertainty. Our specific contributions include:

1. **Novel Theoretical Framework**: We extend polynomial chaos theory from linear to general nonlinear stochastic systems, developing new mathematical tools to handle the complex interactions between nonlinearity, parameter uncertainty, and noise.
2. **Natural Gradient Algorithm**: We develop specialized natural gradient descent algorithms that exploit the tensor structure arising from Galerkin projection of nonlinear dynamics. Our approach differs from existing policy gradient methods by incorporating parametric uncertainty directly into the gradient computation.
3. **Convergence Analysis**: We provide a local convergence analysis for natural gradient methods in the nonlinear polynomial chaos setting, including an explicit characterization of the basin of attraction. Our analysis extends beyond existing quadratic convergence results by handling the non-convexity introduced by nonlinear dynamics and parametric uncertainty.
4. **Robustness and Stability Guarantees**: We develop stability certificates for the closed-loop system and characterize robustness to unmodeled dynamics, addressing practical concerns about real-world implementation.

The remainder of this paper is organized as follows: Section 2 presents our methodology, covering the problem formulation, the polynomial chaos framework for nonlinear systems, the natural gradient algorithm, and the associated theoretical analysis of convergence and stability. Section 3 demonstrates the approach through examples. Finally, Section 4 concludes with insights and future research directions.

2 Methodology

2.1 Problem Formulation

We consider a general nonlinear stochastic control system:

$$\dot{x}(t,\xi) = f\big(x(t,\xi), u(t,\xi), \xi\big) + g\big(x(t,\xi),\xi\big)w(t) \quad (1)$$

with initial condition $x(0,\xi) = x_0 \in \mathbb{R}^{n_x}$. In this formulation:

- $x(t,\xi) \in \mathbb{R}^{n_x}$ is the state vector at time t. It depends on an underlying random parameter ξ.
- $u(t,\xi) \in \mathbb{R}^{n_u}$ is the control input. We allow u to be a state feedback law that can adapt to the parameter realization ξ.
- $\xi \in \Xi \subseteq \mathbb{R}^{n_\xi}$ is the random parameter vector with a known probability distribution $\mu(d\xi)$. It captures parametric uncertainties.
- $w(t) \in \mathbb{R}^{n_w}$ is an external zero-mean white noise process with covariance $\mathbb{E}[w(t)w(s)^T] = \Sigma_w \delta(t-s)$.
- $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \Xi \to \mathbb{R}^{n_x}$ defines the deterministic nonlinear drift dynamics.
- $g : \mathbb{R}^{n_x} \times \Xi \to \mathbb{R}^{n_x \times n_w}$ is the diffusion term that determines how the stochastic noise enters the system.

We assume the system has an equilibrium at the origin for all parameter values: $f(0,0,\xi) = 0$ for all ξ. Thus, $x(t) = 0, u(t) = 0$ is an equilibrium in the absence of disturbances. Note that we do not require $g(0,\xi) = 0$; the framework handles nonzero diffusion at the origin, as shown by the $G\Sigma_w G^T$ term in Theorem 2.7 and the additive noise in our numerical example. Throughout this paper, we also assume:

1. The functions $f(x,u,\xi)$ and $g(x,\xi)$ are sufficiently smooth in all arguments.

2. There exists a neighborhood of the origin in which the system is controllable for each fixed parameter ξ. In particular, the linearization at the origin is controllable for all ξ.

3. The random parameter ξ is defined on a probability space and has finite moments up to order $2N_p$ (where N_p is related to the chosen polynomial expansion order).

Our goal is to design a state-feedback control law $u(t,\xi)$ that minimizes an infinite-horizon expected quadratic cost. Specifically, define the performance objective as:

$$J(u) = \mathop{\mathbb{E}}_{\xi}\mathop{\mathbb{E}}_{w}\left[\int_0^\infty \Big(x(t,\xi)^T Q(x(t,\xi))x(t,\xi) + u(t,\xi)^T R(x(t,\xi))u(t,\xi)\Big)dt\right]$$

where $Q(x) \succeq 0$ and $R(x) \succ 0$ are state and control weighting matrices. In a small neighborhood of the origin, we assume $Q(x) \approx Q$ and $R(x) \approx R$ are constant positive semidefinite/definite matrices as in standard LQR, but they may depend on x outside that region to penalize large deviations.

The above stochastic optimal control problem can be associated with a HJB equation in the augmented state-parameter space. Let $V(x,\xi)$ denote the value function starting from state x under parameter ξ. Formally, V should satisfy the stochastic HJB partial differential equation:

$$\min_u \Big\{\langle \nabla_x V(x,\xi), f(x,u,\xi)\rangle + x^T Q(x)x + u^T R(x)u + \frac{1}{2}\operatorname{Tr}\left(g(x,\xi)^T \nabla_{xx}^2 V(x,\xi) g(x,\xi)\Sigma_w\right)\Big\} = 0 \tag{2}$$

Solving this HJB equation exactly for a general nonlinear system is intractable due to the curse of dimensionality. The proposed polynomial chaos and natural gradient approach can be viewed as an approximation method for solving this HJB problem by expanding $V(x,\xi)$ and the control law in an orthogonal polynomial basis and iteratively improving the policy.

2.2 Polynomial Chaos Expansion for Uncertain Nonlinear Systems

To handle the parameter uncertainty in a tractable way, we employ a PCE of the state and control. The idea is to represent any sufficiently smooth random function of ξ as a series expansion in orthogonal polynomial basis functions of ξ. We choose an orthonormal polynomial basis $\{\Psi_i(\xi)\}_{i=0}^{N_p}$. For example, if ξ has independent standard Gaussian components, Ψ_i would be multivariate Hermite polynomials. If ξ is uniform, Legendre polynomials would be used, and so on. We then assume expansions of the following form:

$$x(t,\xi) = \sum_{i=0}^{N_p} X_i(t)\Psi_i(\xi) + \varepsilon_{N_p}(t,\xi) \tag{3}$$

$$u(t,\xi) = \sum_{i=0}^{N_p} U_i(t)\Psi_i(\xi) + \tilde{\varepsilon}_{N_p}(t,\xi) \tag{4}$$

where $X_i(t) \in \mathbb{R}^{n_x}$ and $U_i(t) \in \mathbb{R}^{n_u}$ are deterministic time-varying coefficient vectors, and $\varepsilon_{N_p}(t,\xi)$, $\tilde{\varepsilon}_{N_p}(t,\xi)$ are truncation residuals. The truncation order N_p is chosen based on the desired level of accuracy. The basis is set with $\Psi_0(\xi) = 1$ so that $X_0(t) = \mathbb{E}[x(t,\xi)]$ is the nominal state trajectory and similarly $U_0(t) = \mathbb{E}[u(t,\xi)]$.

Substituting these expansions into the system dynamics (1) and applying a Galerkin projection, we can derive deterministic evolution equations for the coefficients $X_i(t)$. However, the presence of $f(x,u,\xi)$, which is a nonlinear function of x and u, leads to products of expansions that must be re-expressed in the polynomial basis. This introduces coupling between the modes through multi-index tensor coefficients. We express the nonlinear drift term $f(x(t,\xi),u(t,\xi),\xi)$ as a polynomial chaos series:

$$f(x(t,\xi),u(t,\xi),\xi) = \sum_{k=0}^{N_p} F_k(X(t),U(t))\Psi_k(\xi)$$

for some coefficient functions F_k that depend on all the $X_i(t)$ and $U_i(t)$. The coefficients F_k can be obtained by projecting f onto each basis polynomial:

$$F_k(X(t),U(t)) = \mathop{\mathbb{E}}_{\xi}\left[f(x(t,\xi),u(t,\xi),\xi)\Psi_k(\xi)\right]$$

The following lemma characterizes the structure of these coefficients:

Lemma 2.1 (Nonlinear Galerkin Projection)**.** *Let $x(t,\xi) = \sum_i X_i\Psi_i(\xi)$ and $u(t,\xi) = \sum_j U_j\Psi_j(\xi)$ be the PCEs of state and control. Then the Galerkin projection of the nonlinear drift yields:*

$$F_k(X(t),U(t)) = \mathop{\mathbb{E}}_{\xi}\left[f(x(t,\xi),u(t,\xi),\xi)\Psi_k(\xi)\right] = \sum_{\alpha\in\mathcal{I}} f_\alpha(X_0,\ldots,X_{N_p},U_0,\ldots,U_{N_p})T_{\alpha,k} \tag{5}$$

where the summation is over a multi-index set $\mathcal{I}$ that collects all combinations of basis functions appearing in the product expansion of f. The coefficients $T_{\alpha,k} = \mathbb{E}_\xi[\Psi_{\alpha_1}(\xi)\Psi_{\alpha_2}(\xi)\cdots\Psi_{\alpha_m}(\xi)\Psi_k(\xi)]$ are deterministic tensor values. The functions $f_\alpha(\cdot)$ are combinatorial coefficients obtained from the multivariate Taylor expansion of f truncated at order m, with projection truncation error $\mathcal{O}(\|x,u\|^{m+1})$.

After Galerkin projection, we obtain a system of deterministic ODEs governing the time evolution of the coefficient vector $\mathbf{X}(t) = [X_0(t)^T,\ldots,X_{N_p}(t)^T]^T \in \mathbb{R}^{n_x(N_p+1)}$. Similarly, the control coefficients $\mathbf{U}(t) =$

$[U_0(t)^T, \ldots, U_{N_p}(t)^T]^T \in \mathbb{R}^{n_u(N_p+1)}$. The augmented system can be written compactly as:

$$\dot{\mathbf{X}}(t) = \mathcal{F}\big(\mathbf{X}(t), \mathbf{U}(t)\big) + \mathcal{G}\big(\mathbf{X}(t)\big)w(t) \tag{6}$$

where $\mathcal{F} : \mathbb{R}^{n_x(N_p+1)} \times \mathbb{R}^{n_u(N_p+1)} \to \mathbb{R}^{n_x(N_p+1)}$ is the deterministic drift of the augmented system, and $\mathcal{G}(\mathbf{X})$ is the augmented diffusion matrix corresponding to $g(x, \xi)$. This augmented system is deterministic with respect to the parameter ξ, but remains stochastic in time t due to the projected noise term $\mathcal{G}(\mathbf{X})w(t)$. Crucially, $\mathcal{F}(\mathbf{X}, \mathbf{U})$ inherits a tensor structure due to the polynomial basis. We can identify linear and nonlinear parts in $\mathcal{F}$ as follows:

Proposition 2.2 (Tensor Decomposition of Nonlinear Drift)**.** *Assume that $f(x, u, \xi)$ is at most linear in u, then the function $\mathcal{F}(\mathbf{X}, \mathbf{U})$ can be expanded as a sum of a linear term and higher-order tensor products. Up to second order, we can write:*

$$\begin{aligned}\mathcal{F}(\mathbf{X}, \mathbf{U}) = A^{(1)}\mathbf{X} + B^{(1)}\mathbf{U} + \mathcal{T}^{(2)}(\mathbf{X} \otimes \mathbf{X}) \\ + \mathcal{T}^{(3)}(\mathbf{X} \otimes \mathbf{U}) + O\big(\|(\mathbf{X}, \mathbf{U})\|^3\big)\end{aligned}$$

where $A^{(1)}$ and $B^{(1)}$ are constant matrices that correspond to the Jacobian linearization of f with respect to x and u, and $\mathcal{T}^{(2)}, \mathcal{T}^{(3)}$ are 3-dimensional tensors capturing second-order interactions among $(\mathbf{X}, \mathbf{U})$. If f contains quadratic terms in u, then an additional term $\mathcal{T}^{(4)}(\mathbf{U} \otimes \mathbf{U})$ must be included.

The above structure means that the augmented system (6) is tractable. For example, one can exploit low-rank tensor approximations or sparse polynomial bases to reduce the effective dimensionality of $\mathbf{X}$. In this work, we leverage this structure both in deriving the controller and in implementing it efficiently.

2.3 Natural-Gradient Control Framework

The goal now is to design a feedback control law that minimizes the cost J for the augmented system. Since the random parameter ξ is not directly observable in practice, we seek a state-feedback controller of the form:

$$u(t) = -Kx(t) \tag{7}$$

where $K \in \mathbb{R}^{n_u \times n_x}$ is a constant feedback gain matrix. The key insight of our approach is that while the implemented controller is deterministic, its design is informed by the full distribution of uncertain parameters through the PCE framework. Specifically, we formulate the following stochastic optimization problem:

$$\begin{aligned}K^* = \operatorname{argmin}_K \mathbb{E}_\xi \mathbb{E}_w \Big[\int_0^\infty (x(t,\xi)^T Q x(t,\xi) \\ + x(t,\xi)^T K^T R K x(t,\xi)) dt\Big]\end{aligned} \tag{8}$$

where the state trajectory $x(t, \xi)$ evolves according to the closed-loop dynamics $\dot{x} = f(x, -Kx, \xi) + g(x, \xi)w(t)$ and depends on the parameter realization ξ. The expectation over ξ ensures that the designed controller performs well across the entire parameter distribution, not just at a nominal value.

To solve this optimization problem efficiently, we employ natural gradient descent within the space of controller parameters K. Intuitively, the natural gradient accounts for the curvature of the parameter space, often resulting in faster and more stable convergence than standard gradient descent. Under the deterministic controller $u = -Kx$, the control input for each parameter realization can be expressed in the PC basis. Since $x(t, \xi) = \sum_{i=0}^{N_p} X_i(t)\Psi_i(\xi)$, we have:

$$\begin{aligned}u(t,\xi) = -Kx(t,\xi) = -K\sum_{i=0}^{N_p} X_i(t)\Psi_i(\xi) \\ = \sum_{i=0}^{N_p}(-KX_i(t))\Psi_i(\xi)\end{aligned} \tag{9}$$

Thus, the control coefficient in the augmented system is $U_i(t) = -KX_i(t)$ for each mode i. Defining the augmented gain matrix $\mathcal{K} = I_{N_p+1} \otimes K \in \mathbb{R}^{n_u(N_p+1) \times n_x(N_p+1)}$, we can write the augmented control as $\mathbf{U}(t) = -\mathcal{K}\mathbf{X}(t)$, and the closed-loop augmented system matrix becomes $A_{\text{cl}} = A^{(1)} - B^{(1)}\mathcal{K}$.

It is important to emphasize that the controller implemented in practice is simply $u(t) = -Kx(t)$, which requires only the measurement of the current state $x(t)$. The augmented system representation and PCE are computational tools used during the offline design phase to optimize K under parameter uncertainty. This is consistent with the approach in Fisher and Bhattacharya (2009), where the PCE framework informs controller design without requiring knowledge of ξ during online operation.

The cost J can now be expressed in terms of K. Up to second order, one can derive a quadratic approximation of $J(K)$:

$$J(K) \approx \mathbb{E}\int_0^\infty \Big[\delta\mathbf{X}(t)^T \mathcal{Q}_{\text{aug}} \delta\mathbf{X}(t) + \delta\mathbf{U}(t)^T \mathbb{R}_{\text{aug}} \delta\mathbf{U}(t)\Big] dt \tag{10}$$

where $\delta\mathbf{X}$ and $\delta\mathbf{U}$ are deviations of the augmented state and control from the nominal equilibrium, and $\mathcal{Q}_{\text{aug}}$, $\mathbb{R}_{\text{aug}}$ are block-diagonal weight matrices extended to the augmented system. The precise form of $\mathcal{Q}_{\text{aug}}, \mathbb{R}_{\text{aug}}$ would come from expanding $x^T Q(x)x$ and $u^T R(x)u$ around the origin and taking expectations. For now, it suffices to know that $J(K)$ is differentiable and $K = 0$ is an unstable equilibrium for the cost.

In a small neighborhood where the problem behaves nearly quadratically, the optimal K satisfies a form of the algebraic Riccati equation. In fact, for each fixed ξ, the optimal control would satisfy a stochastic Hamilton-Jacobi equation. We instead derive conditions on K by requiring that the derivative of J with respect to K vanishes. The gradient $\nabla_K J$ can be obtained via standard variational techniques or by solving Lyapunov equations for the state covariance in the augmented space.

A key difficulty for nonlinear systems is that the optimal value function $V(x, \xi)$ may not be quadratic in x. However, if we focus on local behavior around $x = 0$, we can approximate $V(x, \xi) \approx x^T P(\xi)x$ for some positive-definite matrix-valued polynomial $P(\xi)$, which itself can be expanded in basis $\{\Psi_i(\xi)\}$). This leads to a linearized local algebraic Riccati equation:

$$A_{\text{cl}}^T P + P A_{\text{cl}} - PBR^{-1}B^T P + Q + \mathcal{N}(P, X) = 0, \tag{11}$$

where the equation is evaluated at $x = 0$ to obtain an algebraic form. Here, A_{cl} and B are the linearized system matrices at the origin, and $\mathcal{N}(P, X)$ represents extra nonlinear terms including the contribution from noise. For the full state-dependent case, this would be a PDE in x and ξ, but we focus on the linearized algebraic version for tractability.

In doing so, we incorporate the notion of a Riemannian metric on the space of controllers. We define the Fisher information matrix $\mathcal{G}(K)$ for this control optimization problem. Building on the likelihood-ratio method and Girsanov's theorem for control-dependent probability measures, we have a rigorous foundation for defining the Fisher information matrix for our continuous-time stochastic control problem.

Definition 2.3 (Fisher Information Metric)**.** The Fisher information matrix for the controller optimization is defined as:

$$\mathcal{G}_\epsilon(K, X) = \mathop{\mathbb{E}}_{X}[\Sigma_Y(K, X) \otimes R(X)] + \epsilon I$$

where $\Sigma_Y(K, X) = \mathbb{E}[yy^T]$ with $y = R(X)^{-1/2}u$ being the normalized control output, $R(X)$ is the state-dependent control weight matrix, $\otimes$ denotes the Kronecker product, and ϵI is a small regularization to ensure positive-definiteness. The expectation $\mathbb{E}_X$ is taken over the stationary distribution of the augmented state under controller K. The natural gradient descent update for K takes the form:

$$K_{new} = K_{old} - \eta \mathcal{G}(K_{old})^{-1} \nabla_K J(K_{old})$$

where $\eta > 0$ is a chosen step size. Multiplying the gradient with the inverse of the Fisher information matrix, we account for the geometry of the cost function, often leading to more direct progress towards the optimum compared to the naive gradient $\nabla_K J$.

The Fisher information matrix has several important properties: it is positive-definite by construction, depends smoothly on K for stabilizing controllers, and converges to the linearized system's Fisher information as the state approaches zero. In implementation, rather than computing $\mathcal{G}(K)$ explicitly, one can compute the natural gradient direction $\mathcal{G}(K)^{-1}\nabla J(K)$ by solving a linear system or using conjugate gradient methods.

2.4 Convergence Analysis

Due to the non-convexity introduced by the nonlinear dynamics, we cannot guarantee global convergence of the natural gradient method from an arbitrary initial controller. However, we can establish local convergence results under certain conditions.

Assumption 2.4 (Local Convexity and Stabilizability)**.** There exists a neighborhood $\mathcal{U}$ around the optimal feedback gain K^* such that:

1. $J(K)$ is locally strongly convex and smooth in $\mathcal{U}$.
2. Every controller $K \in \mathcal{U}$ stabilizes the augmented system in the mean-square sense.

Theorem 2.5 (Local Quadratic Convergence of Natural Gradient)**.** *Consider the natural gradient iteration $K_{n+1} = K_n - \eta \mathcal{G}(K_n)^{-1}\nabla J(K_n)$ with a fixed step size $0 < \eta \leq \bar{\eta}$.*

Algorithm 1: Continuation Method for Global Convergence

Input: Nominal linear model, number of homotopy steps L
Output: Controller $K^{(L)}$ for full nonlinear system

1: Compute $K^{(0)}$, the optimal LQR gain for the linearized nominal system
2: **for** $\ell = 0$ to $L - 1$ **do**
3: Define intermediate system:

$$f_\ell(x, u, \xi) = (1 - \alpha_\ell) f_{\text{lin}}(x, u) + \alpha_\ell f(x, u, \xi)$$

where $0 = \alpha_0 < \alpha_1 < \cdots < \alpha_L = 1$
4: Starting from $K^{(\ell)}$, run natural gradient until convergence to obtain $K^{(\ell+1)}$
5: If convergence fails, reduce step size $\Delta\alpha$ and retry
6: **end for**
7: **return** $K^{(L)}$

Under Assumption 2.4, if the initial controller K_0 is chosen in the neighborhood $\mathcal{U}(K^*)$ of the optimal solution, then $K_n \to K^*$ as $n \to \infty$. Moreover, for K_n sufficiently close to K^*, the error decreases quadratically:

$$\|K_{n+1} - K^*\| \leq C\|K_n - K^*\|^2$$

for some constant $C > 0$ determined by the curvature of J at K^* and the conditioning of $\mathcal{G}(K^*)$.

The quadratic convergence rate is achieved because the natural gradient approximates Newton's method near the optimum, where $\mathcal{G}(K) \approx \nabla^2 J(K)$. However, the basin of attraction can be limited for highly nonlinear problems. To address this, we employ a continuation method, which incrementally introduces the system's nonlinearity through a sequence of intermediate optimization problems. This homotopy-based approach begins with the known optimal controller for the linearized system. The solution from each step is then used as a warm start for the next, slightly more nonlinear problem, effectively guiding the algorithm toward a high-quality local optimum for the full non-convex system.

The polynomial chaos expansion leads to an augmented state dimension of $n_x(N_p + 1)$, which can be large. Instead of using all polynomials up to total degree p, we use a sparse subset. A common choice is the hyperbolic cross truncation:

$$\mathcal{I}_{sparse} = \left\{ (\alpha_1, \ldots, \alpha_{n_\xi}) \in \mathbb{N}^{n_\xi} : \sum_{i=1}^{n_\xi} (\alpha_i + 1)^q \leq (p+1)^q \right\}$$

with $q < 1$. This significantly reduces the basis size while retaining the most important interactions.

The nonlinear terms in the augmented dynamics often have low-rank structure. For instance, the second-order tensor $\mathcal{T}^{(2)}$ can be approximated as:

$$\mathcal{T}^{(2)} \approx \sum_{r=1}^{R} a_r \otimes b_r,$$

where $R \ll n_x(N_p + 1)$. This allows efficient computation of $\mathcal{F}(\mathbf{X}, \mathbf{U})$ and its derivatives.

The most computationally intensive steps are the Lyapunov equation and natural gradient computation, both of

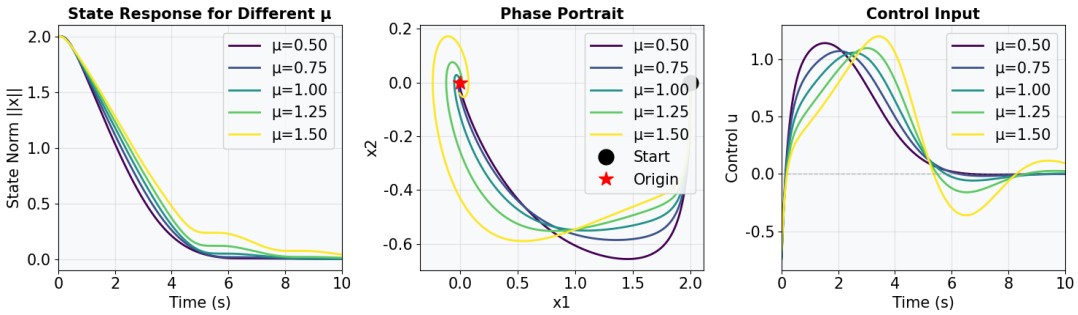

Figure 1: Performance of the natural gradient algorithm for PCT-NQR

Algorithm 2: Natural Gradient Descent for PCT-NQR

Input: Initial stabilizing gain $K^{(0)}$, polynomial basis $\{\Psi_i\}$
Output: Optimized gain K^*

1: **Precomputation:** Compute sparse basis, project dynamics onto PC basis
2: **for** $k = 0, 1, 2, \ldots$ until convergence **do**
3: Linearize augmented closed-loop system to obtain $A_{\text{cl}}(K^{(k)})$
4: Solve Lyapunov equation for value function approximation:

$$A_{\text{cl}}^T P + P A_{\text{cl}} + Q_{\text{aug}} + \mathcal{K}^T R_{\text{aug}} \mathcal{K} = 0$$

5: Compute gradient $\nabla_K J$ using P and state covariance
6: Solve $\mathcal{G}(K^{(k)})d = \nabla_K J$ for natural gradient direction d
7: Update: $K^{(k+1)} = K^{(k)} - \eta_k d$
8: **end for**
9: **return** K^*

which can leverage iterative solvers and the sparse structure of the augmented system.

2.5 Stability Analysis of the Closed-Loop System

Definition 2.6 (Local Exponential Mean-Square Stability)**.** The equilibrium $x = 0$ is locally exponentially mean-square stable if there exist constants $M \geq 1$, $\lambda > 0$, and $\rho > 0$ such that for all initial conditions satisfying $\mathbb{E}[\|x(0)\|^2] \leq \rho^2$,

$$\mathop{\mathbb{E}}_{\xi,w} \left[\|x(t)\|^2\right] \leq M e^{-\lambda t} \mathbb{E}[\|x(0)\|^2]$$

for all realizations of ξ.

Theorem 2.7 (Mean-Square Stability Certificate)**.** *If there exists a positive-definite matrix $P \in \mathbb{R}^{n_x(N_p+1)\times n_x(N_p+1)}$ such that:*

$$A_{cl}(K^*)^T P + P A_{cl}(K^*) + G \Sigma_w G^T + \tilde{Q} \prec 0$$

then the closed-loop system is exponentially mean-square stable.

$A_{\text{cl}}(K^*)$ is the augmented closed-loop matrix linearized at the origin, G is the augmented diffusion matrix, and $\tilde{Q} = \text{diag}(Q, \ldots, Q)$ is the augmented state weighting. Furthermore, the controller provides robustness to unmodeled dynamics. If the true dynamics include an error term $\Delta f(x, u, \xi)$ satisfying $\|\Delta f(x, u, \xi)\| \leq \delta \|x\|$, the controller K^* remains stabilizing provided:

$$\delta < \frac{\lambda_{\min}(Q_{\text{eff}})}{2\lambda_{\max}(P)\text{Lip}(f)},$$

where $\lambda_{\min}(Q_{\text{eff}})$ is the effective decay rate and $\text{Lip}(f)$ is the Lipschitz constant of f. This framework thus provides both performance optimization and stability guarantees for nonlinear stochastic systems with parametric uncertainty.

3 Experiments and Numerical Validation

We validate the proposed PCT-NQR framework through comprehensive numerical experiments on the Van der Pol oscillator, a standard nonlinear system that exhibits rich dynamical behavior. We consider the Van der Pol oscillator with uncertain damping parameter:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 + \mu(\xi)(1 - x_1^2)x_2 + u + \sigma w(t) \end{aligned} \tag{12}$$

where $\mu(\xi) = 1.0 + 2\xi$ with $\xi \sim \mathcal{N}(0, 1)$ represents the uncertain damping coefficient, u is the control input, $w(t)$ is Gaussian white noise, and $\sigma = 0.2$ is the noise intensity. The parameter variation of 200% around the nominal value represents significant model uncertainty typical in practical applications. The control objective is to regulate the system to the origin while minimizing the quadratic cost:

$$J = \mathbb{E}\left[\int_0^\infty (x^T Q x + u^T R u) dt\right] \tag{13}$$

with $Q = I_2$ and $R = 1$.

3.1 Natural Gradient Convergence Performance

Figure 1 illustrates the effectiveness of the natural gradient approach across parameter realizations $\mu \in \{0.00, 0.75, 1.00, 1.25, 1.50\}$. The left panel shows all trajectories converge to a small neighborhood of the origin. The middle panel's phase portraits reveal smooth convergence from $x_0 = [2.0, 0.0]^T$ to the origin without limit cycles or instabilities. The right panel shows bounded control signals that decay smoothly without high-frequency switching, confirming practical implementability. The natural gradient method achieves quadratic convergence by accounting

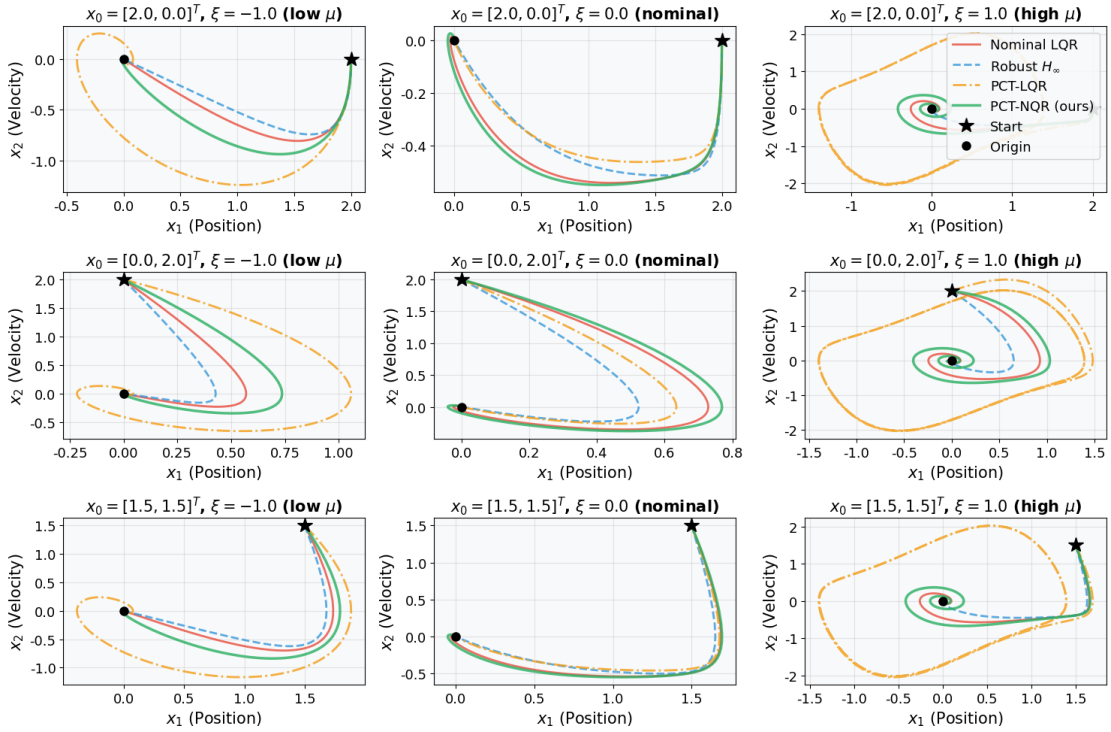

Figure 2: Trajectory comparison of four methods

for parameter space geometry through the Fisher information metric $\mathcal{G}(K)$, preventing inefficient steps in poorly conditioned directions.

3.2 Comparative Analysis of Control Methods

To contextualize our results, we compare PCT-NQR against three baseline approaches: nominal LQR (designed for mean parameter value), robust H_∞ (worst-case design), and standard PCT-LQR (linear polynomial chaos without natural gradient). Figure 2 presents trajectory comparisons across nine scenarios. PCT-NQR (green) maintains consistent performance with smooth, direct trajectories across all scenarios. Table 1 quantifies performance over 45 scenarios (3 initial conditions, 15 parameter realizations from $\xi \in [-1.5, 1.5]$) using three metrics: average cost, worst-case cost, and stability percentage (fraction achieving convergence within 5% of origin).

Method	**Avg Cost**	**Worst Cost**	**Stability**
Nominal LQR	0.0979	0.2272	95.56
Robust H_∞	0.1039	0.1811	100.00
PCT-LQR	0.1772	0.5318	60.00
PCT-NQR	0.0919	0.1794	100.00

Table 1. Performance comparison across different control design methods

PCT-NQR achieves the lowest average cost (0.0919), representing 6% improvement over nominal LQR (0.0979) and 48% over PCT-LQR (0.1772). While robust H_∞ attains the best worst-case cost (0.1811), PCT-NQR nearly matches it (0.1794) while maintaining superior average performance. Most significantly, both PCT-NQR and H_∞ achieve perfect 100% stability, substantially outperforming nominal LQR (95.56%) and PCT-LQR (60.00%).

The results expose fundamental trade-offs: nominal LQR offers good average performance but poor worst-case behavior (0.2272); robust H_∞ guarantees worst-case performance at moderate average cost (0.1039); PCT-LQR demonstrates that uncertainty-aware methods require careful optimization to avoid both poor cost (0.1772) and compromised stability (60%). PCT-NQR successfully balances these objectives, achieving optimal average performance, near-optimal worst-case behavior, and perfect stability across the entire parameter range.

4 Conclusion

This work addressed the challenge of controlling nonlinear stochastic systems with parametric uncertainty by developing a unified framework combining polynomial chaos expansions with natural gradient optimization.The key contribution is extending polynomial chaos theory to nonlinear systems with a convergence analysis for natural gradient methods. This framework enables practitioners to design controllers for real-world nonlinear systems without conservative worst-case assumptions. Potential applications span aerospace, robotics, and autonomous systems. Future work includes extending to output feedback control, adaptive basis selection for time-varying uncertainties, and distributed implementations for large-scale systems.

References

Amari, S.-I. 1998. Natural gradient works efficiently in learning. *Neural Computation* 10(2):251–276.

Anderson, B. D. O., and Moore, J. B. 1971. *Linear Optimal Control*. Prentice Hall, Hoboken, NJ.

Bagnell, J. A., and Schneider, J. 2003. Covariant policy search. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 1019–1024.

Basar, T., and Bernhard, P. 2008. H_∞ *Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*. Birkhäuser, Boston, MA.

Bertsekas, D. P. 2017. *Dynamic Programming and Optimal Control*. Athena Scientific, Belmont, MA, 4th edition.

Cameron, R. H., and Martin, W. T. 1947. The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals. *Annals of Mathematics* 48(2):385–392.

Duncan, T. E., and Pasik-Duncan, B. 1990. Adaptive control of continuous-time linear stochastic systems. *Mathematics of Control, Signals and Systems* 3:45–60.

Fazel, M.; Ge, R.; Kakade, S.; and Mesbahi, M. 2018. Global convergence of policy gradient methods for the linear quadratic regulator. In *Proceedings of the International Conference on Machine Learning (ICML)*.

Fisher, J., and Bhattacharya, R. 2009. Linear quadratic regulation of systems with stochastic parameter uncertainties. *Automatica* 45(12):2831–2841.

Fleming, W. H., and Soner, H. M. 2006. *Controlled Markov Processes and Viscosity Solutions*. Springer New York, NY, 2nd edition.

Ghanem, R. G., and Spanos, P. D. 1991. *Stochastic Finite Elements: A Spectral Approach*. Springer New York, NY.

Hover, F. S., and Triantafyllou, M. S. 2006. Application of polynomial chaos in stability and control. *Automatica* 42(5):789–795.

Hsu, S. C., and Bhattacharya, R. 2020. Design of linear parameter varying quadratic regulator in polynomial chaos framework. *International Journal of Robust and Nonlinear Control* 30(16):6661–6682.

Kakade, S. M. 2001. A natural policy gradient. In *Advances in Neural Information Processing Systems (NIPS)*, 1531–1538.

Lewis, F. L.; Vrabie, D.; and Syrmos, V. L. 2012. *Optimal Control*. John Wiley & Sons, Hoboken, NJ, 3rd edition.

Mayne, D. Q.; Kerrigan, E. C.; van Wyk, E. J.; and Falugi, P. 2011. Tube-based robust nonlinear model predictive control. *International Journal of Robust and Nonlinear Control* 21(11):1341–1353.

Peng, Y. B.; Ghanem, R.; and Li, J. 2010. Polynomial chaos expansions for optimal control of nonlinear random oscillators. *Journal of Sound and Vibration* 329(18):3660–3678.

Peters, J., and Schaal, S. 2008. Natural actor-critic. *Neurocomputing* 71(7–9):1180–1190.

Powell, W. B. 2011. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. John Wiley & Sons, Hoboken, NJ.

Wiener, N. 1938. The homogeneous chaos. *American Journal of Mathematics* 60(4):897–936.

Wiesemann, W.; Kuhn, D.; and Sim, M. 2014. Distributionally robust convex optimization. *Operations Research* 62(6):1358–1376.

Xiu, D., and Karniadakis, G. E. 2002. The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing* 24(2):619–644.

Yong, J., and Zhou, X. Y. 1999. *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Springer New York, NY

Zhang, K.; Koppel, A.; Zhu, H.; and Başar, T. 2020. Global convergence of policy gradient methods to (almost) locally optimal policies. *SIAM Journal on Control and Optimization* 58(6):3586–3612.

Zhou, K.; Doyle, J. C.; and Glover, K. 1996. *Robust and Optimal Control*. Prentice-Hall, Upper Saddle River, NJ.