Koopman Control via Nyström Kernel Approximation

Muthana Alsaadi

College of Information Engineering
University of Al-Nahrain

September 2025

Table of Contents I

Table of Contents II

Table of Contents III

Table of Contents I

Table of Contents II

Table of Contents III

Introduction: Why Can't a Robot Fold a T-Shirt?

Problem

Why can a robot build a car with precision, but can't fold a t-shirt?

Defining the Challenge: Nonlinear Chaos

The problem isn't the robot's strength or precision; it's the nature of the object—it's an unpredictable floppy mess.

Linear Systems (Simple)	Nonlinear Systems (Chaos)
Predictable. Analogous to pushing a ball down a ramp.	Unpredictable. Analogous to a piece of cloth falling through the air. A tiny input can lead to a massive unexpected output.

The Breakthrough Solution & Key Wins

- We combine the Koopman operator framework with kernel methods.
- We introduce the Nyström approximation to handle the computational complexity, acting as a "good enough shortcut".

The Big Deal

The approximated Riccati operator converges at rate $\mathbf{m^{-1/2}}$, and the LQR objective converges at rate $\mathbf{m^{-1}}$, where m is the random subspace size.

Summary Diagram

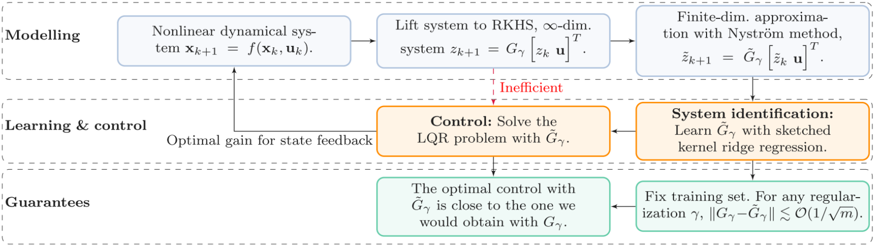

Figure: Given controls and trajectories of a nonlinear system, kernels build a linear, data-driven model. Kernel inversions are made tractable via Nyström.

Table of Contents I

Table of Contents II

Table of Contents III

Systems Notation (Formalizing the Chaos)

- Discrete dynamics: $x_{t+1} = f(x_t, u_t)$.
- Affine-in-input focus: $f(x, u) = g(x) + Bu$.
- Augmented state w: combining state x and input u.

The Koopman Operator (The Magic Trick)

- The Koopman operator transforms our view of the system (not the system itself).
- It lifts behavior via nonlinear observables ξ.
- In the lifted space the dynamics appear linear.

Definition

$(K\xi)(w_t) = \xi(w_{t+1})$.

Reproducing Kernel Hilbert Spaces (RKHS) (The Infinite Stage)

- To get a perfect linear viewpoint, observables live in an RKHS H.
- RKHS can be infinite-dimensional and act as universal approximators.

Challenge

Computers cannot handle infinity.

Linear Quadratic Regulator (LQR) (The Engineer's Tool)

- Once dynamics are linear (in the lifted space), apply LQR.
- Minimize a quadratic objective over an infinite horizon.
- Has an analytic solution: state-feedback gain K via DARE.

Quiz 1 (Prerequisites Check)

Q1

How does the Koopman operator turn nonlinear dynamics into linear ones?

Q2

Why do standard control methods fail on truly nonlinear systems like soft cloth?

Table of Contents I

Table of Contents II

Table of Contents III

The RKHS: The Challenge of the “Infinite Stage” I

- The Koopman operator works in an RKHS (H_1) that contains observables $\psi(x)$ needed to linearize the behavior.

The Cheat Sheet Analogy

The RKHS is like an infinite “cheat sheet” containing every possible feature—too big to compute exactly.

The RKHS: The Challenge of the “Infinite Stage” II

Problem

Modeling perfectly would require processing infinite information. We need a finite surrogate.

Formalizing the Lift

We lift with $\phi(w)$ into H_1 (keeping u linear) and seek an operator G_γ such that $z_{t+1} = G_\gamma \phi(w_t)$.

Kernel Methods and The Nyström Approximation: Making Infinity Finite I

- **Kernel Methods:** Represent the infinite space H_1 while retaining universality.
- **Nyström Shortcut:** A randomized approximation to compute efficiently.

Kernel Methods and The Nyström Approximation: Making Infinity Finite II

The Random Survey Analogy

Select m landmark data pairs ("representative voters"): $\tilde{x}_{\text{in}}, \tilde{x}_{\text{out}}$.

- **Dimensionality Reduction:** Project dynamics onto finite, data–dependent subspaces to get an approximate operator $\widetilde{G}_\gamma$.

Vector Dynamics and Efficiency: The Computable Map I

- Even with Nyström, $z_{t+1} = G_\gamma \phi(w_t)$ lives in function space; we need vectors.
- **The** $m \times m$ **Breakthrough:** Obtain a finite autoregressive model for coordinates $\tilde{z}$ that only needs inverting an $m \times m$ matrix.

Vector Dynamics and Efficiency: The Computable Map II

Rubik's Cube Analogy

Nyström reduces a "billion-layer" cube to an $m \times m$ puzzle—huge computational savings.

- **State Reconstruction:** Reconstruct x from z using a regularized least-squares estimate to define C.

Table of Contents I

Table of Contents II

Table of Contents III

The Two LQR Problems (The Ideal vs. The Real) I

Recap

We now have a finite, vector-valued representation of the dynamics; we can apply LQR.

LQR Goal

Minimize a quadratic cost penalizing state error and control effort.

1. **Exact Dynamics LQR (Ideal):** Solve with the original (infinite-dimensional) RKHS dynamics—perfect but intractable.
2. **Approximated Dynamics LQR (Real):** Solve with Nyström-approximated vector dynamics, yielding the stabilizing gain K.

The Two LQR Problems (The Ideal vs. The Real) II

God-like Controller Analogy

The ideal LQR “knows everything”; the approximated LQR is the fast, effective controller we can actually build.

Defining the Final Control Input (Putting the Gain to Work)

- The approximated LQR gives the optimal gain K.
- Apply K back to the true state x_k of the nonlinear system.
- Final control law: $\mathbf{u_k} = \mathbf{K}\,\mathbf{\Pi}_{\mathrm{out},\mathbf{x}}\,\psi(\mathbf{x_k})$.

Steering Wheel Analogy

Compute K in the simple (linear) surrogate; feed it to the real nonlinear system to steer it.

The Full Pipeline (A 3-Step Recipe for Chaos Control)

1. **Sample Landmarks:** Select m input/output landmarks from training data.
2. **Compute the Map:** Use landmarks to compute efficient vector dynamics (only invert an $m \times m$ matrix).
3. **Find the Steer:** Use these dynamics to solve LQR and get the optimal gain K.

Quiz 2 (Control Loop Check)

Q1

Why can problems (22) and (24) be treated as equivalent for LQR, though one is in function space and the other in vectors?

Q2

With $n = 10{,}000$ and $m = 100$ landmarks, what dictates the complexity of inverting the matrix for (18b)?

Table of Contents I

Table of Contents II

Table of Contents III

Setting the Stage: Necessary Assumptions

To guarantee the LQR problem is well-posed, we need standard stability conditions:

1. **Bounded Kernel:** The RKHS kernel k is bounded.
2. **Stabilizability:** The approximated dynamics can be driven to zero with suitable feedback.
3. **Detectability:** Unstable modes are observable through the LQR cost.
4. **Technical Riccati Condition:** $\sigma_{\min}(P) \geq 1$ (often satisfied by rescaling Q and R).

Accuracy of the Transition Operator G_γ

Theorem 5 (Convergence Rate for $G_\gamma - \widetilde{G}_\gamma$)

This bounds the error between the true infinite-dimensional dynamics (G_γ) and the Nyström shortcut ($\widetilde{G}_\gamma$).

The Result

The error converges proportional to $\mathbf{m^{-1/2}}$.

Rule Book Analogy

If we quadruple m, the operator error roughly halves. The rate applies to components too: $\|A - \widetilde{A}\|$ and $\|B - \widetilde{B}\|$.

Convergence of the Riccati Operator $\widetilde{P}$ I

Recap

The Riccati operator P solves the Discrete-Time Algebraic Riccati Equation (DARE) and determines the LQR gain K.

Lemma 6 (Convergence Rate for $\widetilde{P} - P$)

This shows that the error in the optimal control strategy ($\widetilde{P}$ vs. P) is also bounded by $\mathcal{O}(\epsilon)$.

Convergence of the Riccati Operator $\widetilde{P}$ II

The Result

Since ϵ is $m^{-1/2}$, the Riccati operator converges at rate $\mathbf{m^{-1/2}}$.

Significance

This confirms that as our model improves at a predictable rate, the optimal control strategy derived from it improves at the same rate.

The Ultimate Test: Convergence of the LQR Objective Function $\hat{J}$ I

The Setup

We compare the true optimal cost (J, using perfect dynamics) with the cost ($\hat{J}$) achieved when applying the Nyström-derived control gain $\widetilde{K}$ to the exact dynamics.

Theorem 8 (Convergence Rate for $\hat{J} - J$)

The error between the costs is bounded proportionally to $g(\epsilon)^2$.

The Ultimate Test: Convergence of the LQR Objective Function $\hat{J}$ II

The Big Result

The final LQR objective function (total cost of effort and deviation) converges faster, at rate $\mathbf{m^{-1}}$.

Why Faster?

The error in cost is proportional to the square of the operator error (where $g(\epsilon) = m^{-1/2}$). This quadratic relationship means the performance measure improves much quicker as we increase m.

Quiz 3 (Theory Check)

Q1

If we double the number of Nyström landmarks (m), how much does the theoretical error on the LQR objective cost (J) decrease? (Answer: $O(m^{-1})$, so it approximately halves the error.)

Q2

What technical adjustment must be made to the LQR weights (Q, R) to ensure Assumption 4 ($\sigma_{\min}(P) \geq 1$) is fulfilled?

Table of Contents I

Table of Contents II

Table of Contents III

- Full Presentation

Recap of the Big Wins (The Summary)

1. **Brilliant Combination:** Successfully paired the Koopman framework, reproducing kernels, and the Nyström method to design linear control for nonlinear systems.
2. **Computational Efficiency:** The Nyström approximation achieved huge computational savings while preserving accuracy.
3. **Theoretical Guarantee:** Provided the first theoretical proof that this approach is stable and reliable for control, with the LQR objective cost converging at rate $\mathbf{m^{-1}}$.
4. **Proven Performance:** Tamed hard problems, including dynamic cloth manipulation.
5. **Open Science:** The implementation is publicly available and open-source.

Future Frontiers (Where Do We Go Next?)

This research significantly improves our ability to give robots the tools they need to handle soft, chaotic, and complex things.

Possibilities

What new frontiers could smart robots finally be ready to take on?

1. Delicate manufacturing?
2. Autonomous surgery?
3. Creative fields like art and sculpture?

Final Quiz (Assessment)

Q1

List the convergence rates derived in the theoretical analysis for (1) the transition operator $\widetilde{G}_\gamma$ and (2) the final LQR objective cost J.

Q2

Based on the numerical results, which feature representation (Nyström, Splines, or Eigenfunctions) proved most robust for the Duffing oscillator when the number of features m was small?

Table of Contents I

Table of Contents II

Table of Contents III

- Full Presentation

Reference Paper

Edoardo Caldarelli, Antoine Chatalic, Adrià Colomé, Cesare Molinari, Carlos Ocampo-Martinez, Carme Torras, Lorenzo Rosasco, Linear quadratic control of nonlinear systems with Koopman operator learning and the Nyström method, Automatica, Volume 177, 2025, 112302, ISSN 0005-1098,
https://doi.org/10.1016/j.automatica.2025.112302

Figure: QR code redirecting to the scientific journal referenced.

Full Presentation

Get the full Beamer LaTeX code from my GitHub repository:

Figure: QR code linking to the repository with the full presentation source.