EXPERIMENT DESIGN WITH CONTROL GUARANTEES

DANIEL Y ENRIQUE

Abstract. We study experiment design for continuous-time linear time-invariant systems with the goal of certifying controllability and stabilizability directly from data. We develop a continuous-time, data-driven analogue of the Hautus (PBH) test based on residual cross-moments, provide derivative-free and Itô-compatible estimators, and prove high-probability error bounds under an Itô noise model. Finally, we derive sharp input designs that optimally condition the relevant Gramians under L^2 energy and H^1 smoothness budgets.

Contents

†Chair for Dynamics, Control, Machine Learning, and Numerics (Alexander von Humboldt Professorship), Department of Mathematics, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany.

2020 *Mathematics Subject Classification.* 35B40, 45K05, 74F05.

Key words and phrases. Experiment design, Controllability, Persistent excitation, Linear systems, System identification.

1. Introduction

Modern system identification and data-driven control rely on experiments that generate *informative* data, so that the learned model supports reliable closed-loop guarantees. In particular, for linear systems, a fundamental prerequisite for many synthesis and certification tasks is that the data reveals the system's *controllable subspace* and, ideally, certifies controllability [3, 8]. However, learning objectives that focus only on trajectory matching (prediction or simulation error) do *not* automatically preserve control-theoretic properties such as controllability or stabilizability, since distinct systems can generate similar trajectories on a finite horizon while having different reachable dynamics. This gap motivates experiment design criteria that directly encode control-relevant identifiability.

The classical notion of *persistent excitation* formalizes the idea that the input must excite all relevant degrees of freedom so that system parameters (or behaviors) become identifiable [13]. In the behavioral framework, Willems' Fundamental Lemma states that, for controllable LTI systems, all finite-length trajectories can be parameterized from a single measured trajectory provided the input is persistently exciting of sufficiently high order [13]. This principle underpins a large class of methods for data-driven simulation and control and has been extended in several directions, including multiple datasets and rank-based characterizations [12, 10]. Recent developments also highlight a converse viewpoint: if one seeks a single experiment design that works *uniformly* for broad classes of controllable systems ("universal" inputs), then persistent excitation is not merely sufficient but essentially necessary at the right order [7]. For continuous-time systems, related identifiability conditions and continuous-time variants of Willems-type results have also been investigated [6, 5].

Alongside trajectory-based learning, there is a growing literature on *control-theoretic tests from data*, including controllability/stabilizability certificates that avoid explicit system identification. In discrete time, informativity-based tests provide purely data-dependent rank conditions for control properties [10, 2]. More recently, Mishra et al. developed algebraic data-driven tests for controllability of LTI systems from batches of measurements [4]. These results collectively suggest that controllability can, in principle, be certified directly from suitably designed experiments, even when the system matrices are unknown.

Scope of this paper. In this work, we revisit experiment design for LTI systems from the perspective of certifying the property of *controllability*. Consider an unknown continuous-time LTI system with state matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ and input matrix $\mathbf{B} \in \mathbb{R}^{n\times m}$, governed by

$$\dot{x}(t) = \mathbf{A}x(t) + \mathbf{B}u(t), \qquad x(0) = x_0.$$

The controllable subspace is characterized by the Kalman controllability matrix

$$\mathbf{C} := [\mathbf{B}, \mathbf{AB}, \dots, \mathbf{A}^{n-1}\mathbf{B}] \in \mathbb{R}^{n\times nm},$$

and $\mathrm{rank}(\mathbf{C})$ determines the dimension of the reachable set from the origin [3, 1]. An equivalent characterization is given in terms of the classical Hautus (Popov–Belevitch–Hautus, PBH) test [9, Theorem 3.13], which characterizes controllability and stabilizability as follows:

$$\begin{aligned} &(\mathbf{A},\mathbf{B}) \text{ is controllable} \quad &\Leftrightarrow \quad &\mathrm{rank}[\mathbf{A}-\lambda\mathbf{I},\ \mathbf{B}] = n \quad \forall \lambda \in \mathbb{C}, \\ &(\mathbf{A},\mathbf{B}) \text{ is stabilizable} \quad &\Leftrightarrow \quad &\mathrm{rank}[\mathbf{A}-\lambda\mathbf{I},\ \mathbf{B}] = n \quad \forall \lambda \in \mathbb{C} \text{ with } \Re(\lambda) \geq 0. \end{aligned} \tag{1}$$

The data-driven Hautus test estimates the rank of the matrix using the following identity:

$$[\mathbf{A}-\lambda\mathbf{I},\ \mathbf{B}] \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} = \dot{x}(t) - \lambda x(t)$$

Therefore, assuming access to $u(\cdot)$ and $x(\cdot)$ on a finite horizon $[0, T]$, one can estimate the rank of the Hautus matrix from data without explicitly identifying $(\mathbf{A}, \mathbf{B})$. To recover the rank property *from data*, one can form the cross-moment, defined by

$$\mathbf{H}_\lambda(T) := \int_0^T (\dot{x}(t) - \lambda x(t)) z(t)^\top \, \mathrm{d}t = [\mathbf{A} - \lambda \mathbf{I},\ \mathbf{B}] \underbrace{\int_0^T z(t) z(t)^\top \, \mathrm{d}t}_{\mathbf{S}_Z(T)} \in \mathbb{C}^{n \times (n+m)},$$

where $z(t) := [x(t); u(t)] \in \mathbb{R}^{n+m}$ is the stacked state-input vector. The idea is that, if the Gramian $\mathbf{S}_Z(T)$ is invertible and well-conditioned, then the rank of $\mathbf{H}_\lambda(T)$ reveals the rank of the Hautus matrix. However, the quality of this estimate depends strongly on the choice of input u. So, the main question we address in this paper is: *how to design inputs u that reliably reveal the controllability rank from data?*

1.1. Main Contributions.

- **A data-driven controllability test.** We provide a *continuous-time* and *data-driven* analogue of the Hautus test:

$$(\mathbf{A}, \mathbf{B}) \text{ is controllable} \quad \Leftrightarrow \quad \operatorname{rank}(\mathbf{H}_\lambda(T)) = n \quad \forall \lambda \in \mathbb{C},$$

  assuming $\mathbf{S}_Z(T)$ is invertible. It suffices to check a finite set $\lambda \in \sigma(\mathbf{K})$, where

$$\mathbf{K} := \Big(\int_0^T x(t) x(t)^* \, \mathrm{d}t \Big)^{-1} \Big(\int_0^T x(t) \dot{x}(t)^* \, \mathrm{d}t \Big).$$

  We further develop a derivative-free formulation using Fourier methods.
- **Input design for conditioning the Hautus test.** We provide sharp, model-agnostic designs for conditioning the input, which improves the data-driven test.
  (1) Under an L^2 energy budget $\|u\|_{L^2} \le 1$, the best possible conditioning is achieved by any input with m orthonormal time functions (Proposition 2.3), yielding

$$\int_0^T u(t) u(t)^\top \, \mathrm{d}t = \frac{1}{m} \mathbf{I}_m.$$

  (2) Under an H^1 smoothness budget $\|u\|_{H^1} \le 1$, the optimal design is given by

$$u(t) = \sqrt{\alpha}\, Q \begin{bmatrix} \psi_0(t) \\ \vdots \\ \psi_{m-1}(t) \end{bmatrix}, \qquad \text{where} \qquad \psi_k(t) = \begin{cases} \frac{1}{\sqrt{T}}, & k = 0, \\ \sqrt{\frac{2}{T}} \cos\Big(\frac{k\pi t}{T}\Big), & k \ge 1. \end{cases}$$

  with $Q \in \mathbb{R}^{m \times m}$ an orthogonal matrix and α normalization term (Proposition 2.4).
- **Concentration bounds under Itô noise.** Consider the noisy LTI model given by

$$\mathrm{d}x(t) = \Big(\mathbf{A} x(t) + \mathbf{B} u(t) \Big) \, \mathrm{d}t + \boldsymbol{\beta} \, \mathrm{d}W(t),$$

  where W is a Brownian motion. In Proposition 3.1, we prove

$$|\sigma_{\min}(\widetilde{\mathbf{P}}_\lambda(T)) - \sigma_{\min}(\mathbf{P}_\lambda)| = \mathcal{O}\left(\tfrac{1}{\sqrt{T}}\right),$$

  where $\mathbf{P}_\lambda := [\mathbf{A} - \lambda \mathbf{I},\ \mathbf{B}]$ and $\widetilde{\mathbf{P}}_\lambda(T)$ is the data-driven estimate from the noisy model. The estimation error of the Hautus test decays as the observation horizon T increases.

Together, these results provide a principled experiment-design toolkit for learning models that preserve controllability structure and are therefore suitable for safe downstream control.

Outline. The remainder of the paper is organized as follows. Section 2 introduces the deterministic residual cross-moment $\mathbf{H}_\lambda(T)$, proves the continuous-time data-driven Hautus test,

and provides model-agnostic input-conditioning results; it also presents an operator formulation that reduces checking all $\lambda \in \mathbb{C}$ to finitely many candidate values. Section 3 develops the Itô cross-moment estimator and concentration bounds, and discusses a derivative-free Fourier approximation. Section 4 reports numerical experiments. Appendix A contains the proofs.

2. Deterministic Framework for Controllability Certification

We assume access to $u(\cdot)$ and the resulting state $x(\cdot)$ on a finite horizon $[0, T]$. When $\dot{x}$ is not directly available or is too noisy to estimate reliably, we rely on the cross-moment formulation in Section 3, which only uses increments of x in the stochastic setting and avoids forming $\dot{x}$ in the deterministic setting.

2.1. Noise-free cross-moment formulation.

We consider the continuous-time Linear Time-Invariant (LTI) system

$$\begin{aligned} \dot{x}(t) &= \mathbf{A}x(t) + \mathbf{B}u(t), \\ x(0) &= x_0, \end{aligned} \tag{2}$$

where the state $x(t) \in \mathbb{R}^n$, the input $u(t) \in \mathbb{R}^m$, and the system matrices $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times m}$ are unknown.

For input design it is convenient to work with a time-domain moment formulation based on the residual and the stacked state–input signal. Fix a horizon $T > 0$ and let $u \in L^2(0, T; \mathbb{R}^m)$ be an input generating an absolutely continuous state trajectory $x : [0, T] \to \mathbb{R}^n$ with $x, \dot{x} \in L^2(0, T; \mathbb{R}^n)$. For $\lambda \in \mathbb{C}$ define the cross-moment

$$\mathbf{H}_\lambda(T) := \int_0^T (\dot{x}(t) - \lambda x(t))z(t)^\top \, \mathrm{d}t \in \mathbb{C}^{n\times(n+m)}.$$

where $z(t) := [x(t); u(t)] \in \mathbb{R}^{n+m}$ is the stacked state-input vector. Define the stacked Gramian

$$\mathbf{S}_Z(T) := \int_0^T z(t)z(t)^\top \, \mathrm{d}t \in \mathbb{R}^{(n+m)\times(n+m)}.$$

Finally, define the Hautus matrix

$$\mathbf{P}_\lambda := \begin{bmatrix} \mathbf{A} - \lambda\mathbf{I} & \mathbf{B} \end{bmatrix} \in \mathbb{C}^{n\times(n+m)}.$$

Note that $\mathbf{S}_Z(T) \succeq 0$ by construction. Moreover, for every admissible input u,

$$\mathbf{H}_\lambda(T) = \mathbf{P}_\lambda \mathbf{S}_Z(T). \tag{3}$$

If $\mathbf{S}_Z(T)$ is invertible, then right-multiplication by $\mathbf{S}_Z(T)$ preserves row rank and yields the continuous-time Hautus test.

Theorem 2.1 (Continuous Hautus Test). *If there exists u such that $\operatorname{rank}(\mathbf{H}_\lambda(T)) = n$ for all $\lambda \in \mathbb{C}$, then $(\mathbf{A}, \mathbf{B})$ is controllable. Moreover, if $\mathbf{S}_Z(T)$ is invertible, then the converse holds*

$$\begin{aligned} (\mathbf{A}, \mathbf{B}) \text{ is controllable} \quad &\Leftrightarrow \quad \operatorname{rank}(\mathbf{H}_\lambda(T)) = n \text{ for all } \lambda \in \mathbb{C}, \\ (\mathbf{A}, \mathbf{B}) \text{ is stabilizable} \quad &\Leftrightarrow \quad \operatorname{rank}(\mathbf{H}_\lambda(T)) = n \text{ for all } \lambda \in \mathbb{C} \text{ with } \Re(\lambda) \geq 0. \end{aligned}$$

Proof. See Appendix A.1. □

This result justifies the use of the margin $\sigma_{\min}(\mathbf{H}_\lambda(T))$ as a data-driven certificate of controllability and stabilizability. Moreover, $\mathbf{S}_Z(T)$ plays a key role in the conditioning of the test (see Section 3). Therefore, input design should seek to maximize $\sigma_{\min}(\mathbf{S}_Z(T))$ under suitable constraints.

2.2. **Best conditioning of control inputs.** The singular-value margins of the data-driven Hautus test are controlled—up to the model-dependent factor $\sigma_{\min}(\mathbf{P}_\lambda)$—by the smallest singular value of the stacked Gramian $\mathbf{S}_Z(T)$. Since $\mathbf{S}_Z(T)$ depends on the (unknown) response $x(\cdot)$, a natural model-agnostic surrogate is to ensure that the *input* Gramian is well conditioned, so that the input directions are persistently excited and the inversion of $\mathbf{S}_Z(T)$ is numerically stable. Define the input Gramian

$$\mathbf{S}_U(u) := \int_0^T u(t)u(t)^\top \, \mathrm{d}t \in \mathbb{R}^{m\times m}.$$

Remark 2.2. *The matrix $\mathbf{S}_U(u)$ is a principal submatrix of $\mathbf{S}_Z(T)$, hence*

$$\sigma_{\min}(\mathbf{S}_Z(T)) \le \sigma_{\min}(\mathbf{S}_U(u)).$$

Thus, even though $\sigma_{\min}(\mathbf{S}_Z(T))$ depends on the state response, and therefore on $(\mathbf{A},\mathbf{B})$ and the initial condition, choosing inputs with well-conditioned $\mathbf{S}_U(u)$ is a natural baseline when seeking a large $\sigma_{\min}(\mathbf{S}_Z(T))$ without model knowledge. This idea motivates the input designs in the next results.

Under an L^2 energy budget, the best possible conditioning corresponds to spreading the energy isotropically across the m input channels.

Proposition 2.3 (Best conditioning under $\|u\|_{L^2} \le 1$)**.** *Let us define the following optimization problem:*

$$\max_{u\in L^2(0,T;\mathbb{R}^m)} \lambda_{\min}\Big(\mathbf{S_U(u)}\Big) \qquad \textit{s.t.} \quad \|u\|^2_{L^2(0,T;\mathbb{R}^m)} \le 1,$$

then

$$\lambda_{\min}(\mathbf{S}_U(u)) \le \frac{1}{m}.$$

Moreover, there exist u with $\|u\|_{L^2(0,T)} \le 1$ such that $\mathbf{S}_U(u) = \frac{1}{m}\mathbf{I}_m$.

Proof. See Appendix A.5. □

In applications one may also constrain input *smoothness*, for instance by an H^1 budget

$$\|u\|^2_{H^1(0,T)} := \int_0^T \Big(\|u(t)\|_2^2 + \|\dot u(t)\|_2^2\Big)\,\mathrm{d}t \le 1.$$

This penalizes high-frequency excitation and therefore reduces the best achievable isotropic conditioning compared to the L^2-only case.

Proposition 2.4 (Best conditioning under $\|u\|_{H^1} \le 1$)**.** *The following optimization problem*

$$\max_{u\in H^1(0,T;\mathbb{R}^m)} \lambda_{\min}\Big(\mathbf{S_U(u)}\Big) \qquad \textit{s.t.} \quad \|u\|^2_{H^1(0,T;\mathbb{R}^m)} \le 1,$$

admits the upper bound

$$\lambda_{\min}(\mathbf{S}_U(u)) \le \frac{1}{\displaystyle\sum_{k=0}^{m-1}\left(1+\left(\tfrac{k\pi}{T}\right)^2\right)} = \frac{1}{m + \dfrac{\pi^2}{T^2}\cdot\dfrac{(m-1)m(2m-1)}{6}}.$$

Moreover, the bound is tight if $\{\psi_k\}_{k\ge0}$ are the Neumann eigenfunctions on $[0,T]$ given by

$$\psi_0(t) := \frac{1}{\sqrt{T}}, \qquad \psi_k(t) := \sqrt{\frac{2}{T}}\cos\!\left(\frac{k\pi t}{T}\right)\ (k\ge1),$$

and $Q \in \mathbb{R}^{m\times m}$ is orthogonal, then with

$$u(t) = \sqrt{\alpha}\, Q \begin{bmatrix} \psi_0(t) \\ \vdots \\ \psi_{m-1}(t) \end{bmatrix}, \quad \textit{where} \quad \alpha := \Big(\sum_{k=0}^{m-1} \big(1 + (\tfrac{k\pi}{T})^2\big) \Big)^{-1}$$

one has $\|u\|^2_{H^1(0,T)} = 1$ and $\mathbf{S}_U(u) = \alpha\, \mathbf{I}_m$.

Proof. See Appendix A.6. □

Remark 2.5. *As $T \to \infty$ the derivative penalty vanishes and $\alpha \to 1/m$, recovering the L^2-budget optimum in Proposition 2.3. For short horizons (or large m), the optimal H^1-budget design suppresses high-frequency components and yields a smaller isotropic eigenvalue α.*

2.3. **An operator formulation and finitely many candidate λ.** The main inconvenience of Theorem 2.1 is that the rank condition must hold for all $\lambda \in \mathbb{C}$. To address this, we now develop an operator formulation that allows us to identify finitely many candidate λ where rank failure can occur. Let $x : [0,T] \to \mathbb{R}^n$ be absolutely continuous with $x, \dot{x} \in L^2(0,T;\mathbb{R}^n)$ and define operators

$$\mathcal{X}, \dot{\mathcal{X}} : L^2(0,T) \to \mathbb{C}^n, \qquad \mathcal{X}(\varphi) := \int_0^T x(t)\varphi(t)\,\mathrm{d}t, \quad \dot{\mathcal{X}}(\varphi) := \int_0^T \dot{x}(t)\varphi(t)\,\mathrm{d}t.$$

For $\varphi \in L^2(0,T)$ and $w \in \mathbb{C}^n$,

$$\langle \mathcal{X}(\varphi), w\rangle_{\mathbb{C}^n} = \Big(\int_0^T x(t)\varphi(t)\,\mathrm{d}t \Big)^* w = \int_0^T \overline{\varphi(t)}\, x(t)^* w\,\mathrm{d}t = \langle \varphi,\, x(\cdot)^* w\rangle_{L^2(0,T)}.$$

Repeating the same process for $\dot{\mathcal{X}}$ yields

$$(\mathcal{X}^* w)(t) = x(t)^* w, \qquad (\dot{\mathcal{X}}^* w)(t) = \dot{x}(t)^* w \qquad \text{in } L^2(0,T),$$

For $\lambda \in \mathbb{C}$, define the pencil of operators

$$\mathcal{P}(\lambda) := \dot{\mathcal{X}} - \lambda \mathcal{X}.$$

Since the codomain is finite-dimensional, $\operatorname{rank}(\mathcal{P}(\lambda)) := \dim(\operatorname{range}(\mathcal{P}(\lambda))) \le n$, and *full rank* means surjectivity onto $\mathbb{C}^n$. Therefore, for $w \in \mathbb{C}^n$,

$$\mathcal{P}(\lambda)\mathcal{P}(\lambda)^* w = \mathcal{P}(\lambda)\big(\dot{\mathcal{X}}^* w - \overline{\lambda}\mathcal{X}^* w\big) = \int_0^T \big(\dot{x}(t) - \lambda x(t)\big)\big(\dot{x}(t) - \lambda x(t)\big)^* w\,\mathrm{d}t =: \mathbf{G}_\lambda(u)\, w.$$

Hence,

$$\mathcal{P}(\lambda)\mathcal{P}(\lambda)^* = \mathbf{G}_\lambda(u) \qquad \text{as operators } \mathbb{C}^n \to \mathbb{C}^n. \tag{4}$$

Since $\mathcal{P}(\lambda) : L^2(0,T) \to \mathbb{C}^n$ has finite-dimensional codomain, it is surjective if and only if $\mathcal{P}(\lambda)\mathcal{P}(\lambda)^*$ is invertible on $\mathbb{C}^n$. By (4), for every $\lambda \in \mathbb{C}$,

$$\operatorname{rank}(\mathcal{P}(\lambda)) = n \quad \Leftrightarrow \quad \mathbf{G}_\lambda(u) \text{ is invertible.} \tag{5}$$

When $\mathbf{S}_Z(T) \succ 0$, the cross-moment representation further gives

$$\mathbf{G}_\lambda(u) = \mathbf{H}_\lambda(T)\mathbf{S}_Z(T)^{-1}\mathbf{H}_\lambda(T)^*, \tag{6}$$

so $\mathbf{G}_\lambda(u)$ is invertible if and only if $\mathbf{H}_\lambda(T)$ has full row rank. Thus, in the data-driven Hautus test it suffices to understand for which λ the matrix $\mathbf{H}_\lambda(T)$ can lose rank. The next result shows that, under a mild nondegeneracy assumption, rank failure can only occur at finitely many *candidate* values of λ.

Theorem 2.6 (Finite candidate set for $\lambda \in \mathbb{C}$). *Assume* $\operatorname{rank}(\mathcal{X}) = n$, *equivalently*

$$\mathcal{X}\mathcal{X}^* = \int_0^T x(t)x(t)^*\,\mathrm{d}t \in \mathbb{C}^{n\times n} \quad \text{is invertible.}$$

This is a data-richness condition: it requires, in particular, that the trajectory does not remain in a strict subspace of $\mathbb{R}^n$ *on* $[0,T]$. *Define*

$$\mathbf{K} := (\mathcal{X}\mathcal{X}^*)^{-1}\mathcal{X}\dot{\mathcal{X}}^* = \Big(\int_0^T x(t)x(t)^*\,\mathrm{d}t\Big)^{-1}\Big(\int_0^T x(t)\dot{x}(t)^*\,\mathrm{d}t\Big) \in \mathbb{C}^{n\times n}.$$

Then for every $\lambda \in \mathbb{C}$,

$$\operatorname{rank}(\mathcal{P}(\lambda)) < n \quad \Longrightarrow \quad \lambda \in \sigma(\mathbf{K}).$$

In particular, the set of λ *for which* $\operatorname{rank}(\mathcal{P}(\lambda)) < n$ *is contained in* $\sigma(\mathbf{K})$.

Proof. See Appendix A.3. □

Suppose there existed $\lambda_0 \in \mathbb{C}$ with $\operatorname{rank}\Big(\mathbf{H}_{\lambda_0}(T)\Big) < n$. Since $\mathbf{S}_Z(T) \succ 0$, the matrix (6) is singular, hence (5) implies $\operatorname{rank}(\mathcal{P}(\lambda_0)) < n$. Therefore, the above theorem yields $\lambda_0 \in \sigma(\mathbf{K})$, which implies the following result.

Corollary 2.7 (Finite checking for the data-driven test). *Assume* $\operatorname{rank}(\mathcal{X}) = n$ *and* $\mathbf{S}_Z(T) \succ 0$, *and let* $\mathbf{K}$ *be as in Theorem 2.6. Define the candidate set* $\Lambda := \sigma(\mathbf{K})$, *hence* $|\Lambda| \le n$. *If*

$$\operatorname{rank}\Big(\mathbf{H}_\lambda(T)\Big) = n \qquad \text{for all } \lambda \in \Lambda,$$

then $\operatorname{rank}\Big(\mathbf{H}_\lambda(T)\Big) = n$ *for all* $\lambda \in \mathbb{C}$.

Lemma 2.8 (A quantitative lower bound for $\sigma_{\min}(\mathbf{H}_\lambda(T))$ via $\mathbf{K}$). *Assume* $\operatorname{rank}(\mathcal{X}) = n$, $\mathbf{S}_Z(T) \succ 0$, *and let* $\mathbf{K}$ *be as in Theorem 2.6. Then for every* $\lambda \in \mathbb{C}$,

$$\sigma_{\min}\Big(\mathbf{H}_\lambda(T)\Big) \ge \sqrt{\sigma_{\min}\Big(\mathbf{S}_Z(T)\Big)}\,\frac{\sigma_{\min}(\mathbf{K} - \overline{\lambda}\,\mathbf{I})}{\|(\mathcal{X}\mathcal{X}^*)^{-1}\mathcal{X}\|}.$$

In particular, if $\lambda \notin \sigma(\mathbf{K})$, *then* $\mathbf{H}_\lambda(T)$ *has full row rank.*

Proof. See Appendix A.4. □

3. Stochastic model and Itô residual

The deterministic framework of Section 2 assumes access to $\dot{x}$ or relies on forming derivatives from sampled data. In practice, measurements are corrupted by noise, making derivative estimation unreliable. In this section, we extend the framework to handle additive stochastic disturbances and estimate $\mathbf{P}_\lambda$—and thus the rank of the data-driven residual cross-moment—from cross-moments between the state and input signals, using only increments of x. This is particularly natural in a stochastic setting, where $\dot{x}$ does not exist pointwise. Assume that the measured state is an Itô process satisfying the linear SDE

$$\begin{aligned} \mathrm{d}x(t) &= \Big(\mathbf{A}x(t) + \mathbf{B}u(t)\Big)\,\mathrm{d}t + \boldsymbol{\beta}\,\mathrm{d}W(t), \\ x(0) &= x_0, \end{aligned} \tag{7}$$

where W is a q-dimensional standard Brownian motion and $\boldsymbol{\beta} \in \mathbb{R}^{n\times q}$ is constant. As before, define the stacked signal $z(t) := [x(t); u(t)] \in \mathbb{R}^{n+m}$ and the Hautus matrix

$$\mathbf{P}_\lambda := \begin{bmatrix} \mathbf{A} - \lambda\mathbf{I} & \mathbf{B} \end{bmatrix} \in \mathbb{C}^{n\times(n+m)}.$$

For $\lambda \in \mathbb{C}$, define the Itô residual differential

$$\mathrm{d}m_\lambda(t) := \mathrm{d}x(t) - \lambda x(t)\,\mathrm{d}t = \mathbf{P}_\lambda z(t)\,\mathrm{d}t + \boldsymbol{\beta}\,\mathrm{d}W(t). \tag{8}$$

Cross-moment factorization. Define the (random) stacked Gramian and cross-moment matrix

$$\widetilde{\mathbf{S}}_Z(T) := \int_0^T z(t)z(t)^\top \,\mathrm{d}t \in \mathbb{R}^{(n+m)\times(n+m)}, \tag{9}$$

$$\widetilde{\mathbf{H}}_\lambda(T) := \int_0^T \mathrm{d}m_\lambda(t)\, z(t)^\top = \mathbf{P}_\lambda \widetilde{\mathbf{S}}_Z(T) + \boldsymbol{\beta} \int_0^T \mathrm{d}W(t)\, z(t)^\top \in \mathbb{C}^{n\times(n+m)}. \tag{10}$$

The last term is a matrix-valued martingale with $\mathbb{E}[\int_0^T \mathrm{d}W(t)\, z(t)^\top] = 0$, hence

$$\mathbb{E}\Big[\widetilde{\mathbf{H}}_\lambda(T) - \mathbf{P}_\lambda \widetilde{\mathbf{S}}_Z(T)\Big] = 0.$$

Let us estimate $\mathbf{P}_\lambda$. Assume that $\widetilde{\mathbf{S}}_Z(T)$ is invertible and define

$$\widetilde{\mathbf{P}}_\lambda(T) := \widetilde{\mathbf{H}}_\lambda(T)\widetilde{\mathbf{S}}_Z(T)^{-1}. \tag{11}$$

Then (10) implies the exact error identity

$$\widetilde{\mathbf{P}}_\lambda(T) - \mathbf{P}_\lambda = \boldsymbol{\beta}\left(\int_0^T \mathrm{d}W(t)\, z(t)^\top\right)\widetilde{\mathbf{S}}_Z(T)^{-1}. \tag{12}$$

We want to characterize the statistical rate of convergence of $\widetilde{\mathbf{P}}_\lambda(T)$ to $\mathbf{P}_\lambda$ as $T \to \infty$. Introduce the normalized quantities $\bar{\mathbf{S}}_Z(T) := \frac{1}{T}\widetilde{\mathbf{S}}_Z(T)$ and $\bar{\mathbf{H}}_\lambda(T) := \frac{1}{T}\widetilde{\mathbf{H}}_\lambda(T)$ so that $\widetilde{\mathbf{P}}_\lambda(T) = \bar{\mathbf{H}}_\lambda(T)\bar{\mathbf{S}}_Z(T)^{-1}$. The Itô isometry yields (componentwise) the bound

$$\mathbb{E}\left\|\frac{1}{T}\int_0^T \mathrm{d}W(t)\, z(t)^\top\right\|_F^2 = \frac{q}{T^2}\mathbb{E}\left[\int_0^T \|z(t)\|_2^2 \,\mathrm{d}t\right]. \tag{13}$$

If $\mathbb{E}[\int_0^T \|z(t)\|_2^2 \,\mathrm{d}t] = \mathcal{O}(T)$ and $\|\bar{\mathbf{S}}_Z(T)^{-1}\|_2 = \mathcal{O}_{\mathbb{P}}(1)$, then (12) and (13) imply

$$\|\widetilde{\mathbf{P}}_\lambda(T) - \mathbf{P}_\lambda\|_2 = \mathcal{O}_{\mathbb{P}}(T^{-1/2}). \tag{14}$$

The $\mathcal{O}_{\mathbb{P}}(T^{-1/2})$ statement above can be strengthened to an explicit bound holding with probability $1-\delta$ and depending only on the (random) conditioning of $\bar{\mathbf{S}}_Z(T)$.

Proposition 3.1 (Cross-moment error bound)**.** *Assume that $\widetilde{\mathbf{S}}_Z(T) \succ 0$. Then for every $\delta \in (0,1)$, with probability at least $1-\delta$,*

$$\|\widetilde{\mathbf{P}}_\lambda(T) - \mathbf{P}_\lambda\|_2 \le \frac{\|\boldsymbol{\beta}\|_2}{\sqrt{T\,\sigma_{\min}(\bar{\mathbf{S}}_Z(T))}}\left(\sqrt{q} + \sqrt{n+m} + \sqrt{2\log(1/\delta)}\right) = \mathcal{O}(T^{-1/2}). \tag{15}$$

Moreover, this implies the uniform bound $\sup_{\lambda\in\mathbb{C}} \|\widetilde{\mathbf{P}}_\lambda(T) - \mathbf{P}_\lambda\|_2$ since the right-hand side of (12) does not depend on λ. Finally, by Weyl's inequality,

$$\left|\sigma_{\min}\Big(\widetilde{\mathbf{P}}_\lambda(T)\Big) - \sigma_{\min}(\mathbf{P}_\lambda)\right| \le \|\widetilde{\mathbf{P}}_\lambda(T) - \mathbf{P}_\lambda\|_2,$$

Proof. See Appendix A.2. □

Remark 3.2. *The above result provides statistical guarantees for the Hautus test. In particular,*

$$\begin{aligned}
\sigma_{\min}(\mathbf{P}_\lambda) &\ge \sigma_{\min}\Big(\widetilde{\mathbf{P}}_\lambda(T)\Big) - \left|\sigma_{\min}(\mathbf{P}_\lambda) - \sigma_{\min}\Big(\widetilde{\mathbf{P}}_\lambda(T)\Big)\right| \\
&\ge \sigma_{\min}\Big(\widetilde{\mathbf{P}}_\lambda(T)\Big) - \|\widetilde{\mathbf{P}}_\lambda(T) - \mathbf{P}_\lambda\|_2 \\
&\ge \sigma_{\min}\Big(\widetilde{\mathbf{P}}_\lambda(T)\Big) - \mathcal{O}(T^{-1/2}).
\end{aligned}$$

3.1. **A Fourier-domain approximation without derivatives.** The cross-moment in (10) can also be approximated in the frequency domain without forming $\dot{x}$. For each $\omega \in \mathbb{R}$, define the Fourier transform of the residual increment

$$\widehat{m}_\lambda(\mathrm{i}\omega) := \int_0^T e^{-\mathrm{i}\omega t}\,\mathrm{d}m_\lambda(t) = \int_0^T e^{-\mathrm{i}\omega t}\,\mathrm{d}x(t) - \lambda\widehat{x}(\mathrm{i}\omega), \qquad \widehat{x}(\mathrm{i}\omega) := \int_0^T x(t)e^{-\mathrm{i}\omega t}\,\mathrm{d}t.$$

Since $t \mapsto e^{-\mathrm{i}\omega t}$ is of bounded variation, Itô integration by parts yields

$$\int_0^T e^{-\mathrm{i}\omega t}\,\mathrm{d}x(t) = x(T)e^{-\mathrm{i}\omega T} - x(0) + \mathrm{i}\omega\int_0^T x(t)e^{-\mathrm{i}\omega t}\,\mathrm{d}t,$$

hence

$$\widehat{m}_\lambda(\mathrm{i}\omega) = x(T)e^{-\mathrm{i}\omega T} - x(0) + (\mathrm{i}\omega - \lambda)\widehat{x}(\mathrm{i}\omega). \tag{16}$$

Thus, $\widehat{m}_\lambda(\mathrm{i}\omega)$ can be computed from an FFT of x plus the boundary terms $x(0), x(T)$. Define also $\widehat{z}(\mathrm{i}\omega) := \int_0^T z(t)e^{-\mathrm{i}\omega t}\,\mathrm{d}t$. If x is absolutely continuous so that $\mathrm{d}m_\lambda(t) = (\dot{x}(t) - \lambda x(t))\,\mathrm{d}t$, then Parseval gives the deterministic identities

$$\mathbf{H}_\lambda(T) = \int_{\mathbb{R}} \widehat{m}_\lambda(\mathrm{i}\omega)\widehat{z}(\mathrm{i}\omega)^*\,\frac{\mathrm{d}\omega}{2\pi}, \qquad \mathbf{S}_Z(T) = \int_{\mathbb{R}} \widehat{z}(\mathrm{i}\omega)\widehat{z}(\mathrm{i}\omega)^*\,\frac{\mathrm{d}\omega}{2\pi}.$$

In the Itô setting, the unwindowed energy $\int_{\mathbb{R}} \|\widehat{m}_\lambda(\mathrm{i}\omega)\|_2^2\,\mathrm{d}\omega$ is not finite, so frequency-domain computations should be interpreted with an explicit cutoff/windowing when needed.

4. Numerical Experiments

We validate the theoretical results through numerical experiments implemented in PyTorch. All simulations use Euler–Maruyama time stepping to integrate controlled linear SDEs of the form

$$\mathrm{d}x(t) = (\mathbf{A}x(t) + \mathbf{B}u(t))\,\mathrm{d}t + \boldsymbol{\beta}\,\mathrm{d}W(t),$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$ is the drift matrix, $\mathbf{B} \in \mathbb{R}^{n\times m}$ the input matrix, $\boldsymbol{\beta} \in \mathbb{R}^{n\times q}$ the noise intensity, and W a q-dimensional standard Brownian motion. Unless stated otherwise, the input is a deterministic multi-sinusoid with each channel given by

$$u_i(t) = \sin(\omega_i t), \qquad i = 1, \ldots, m,$$

where $\omega_1, \ldots, \omega_m \in \mathbb{R}$ are distinct frequencies chosen to avoid aliasing.

4.1. **Validation of the Data-Driven Hautus Test.** Our first experiment validates the data-driven Hautus test by comparing the estimated matrix $\widetilde{\mathbf{P}}_\lambda$ with the true Hautus matrix $\mathbf{P}_\lambda = [\mathbf{A} - \lambda\mathbf{I},\ \mathbf{B}]$. We simulate a system with $n = 6$ states, $m = 3$ inputs, and $q = 2$ noise dimensions over a horizon $T = 100$ with time step $\mathrm{d}t = 0.05$. The drift matrix $\mathbf{A}$ is constructed to be Hurwitz stable, with all eigenvalues having real parts bounded by -0.1.

For each test value $\lambda \in \mathbb{C}$, we compute both time-domain and FFT-based estimates via the cross-moment formulation (11).

Figure 1 shows the element-wise comparison between the true and estimated matrices for a representative $\lambda = 0.3 + 1.3\mathrm{i}$. Both methods yield accurate estimates, with estimation errors $\|\widetilde{\mathbf{P}}_\lambda - \mathbf{P}_\lambda\|_2$ on the order of 10^{-2}. To assess variability, we run 50 independent trajectories and examine the distribution of errors across the eigenvalues of $\mathbf{A}$ and additional random test points. Figure 2 (left) summarizes the signed error $\sigma_{\min}(\mathbf{P}_\lambda) - \sigma_{\min}(\widetilde{\mathbf{P}}_\lambda)$, showing near-zero bias and comparable variance for time-domain and FFT estimators.

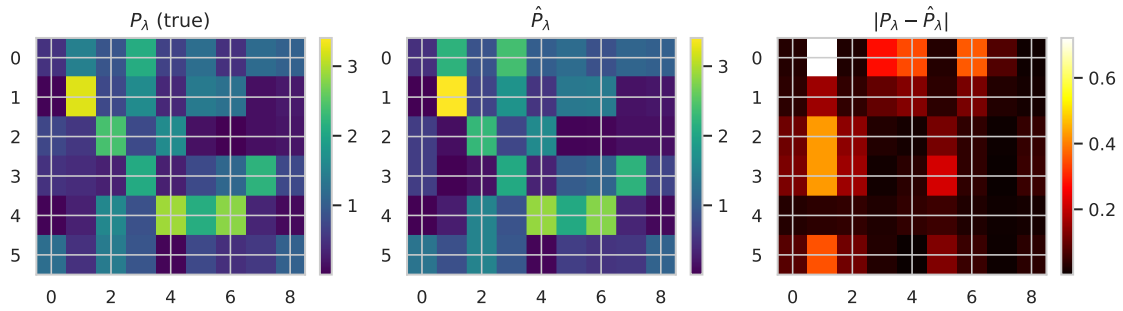
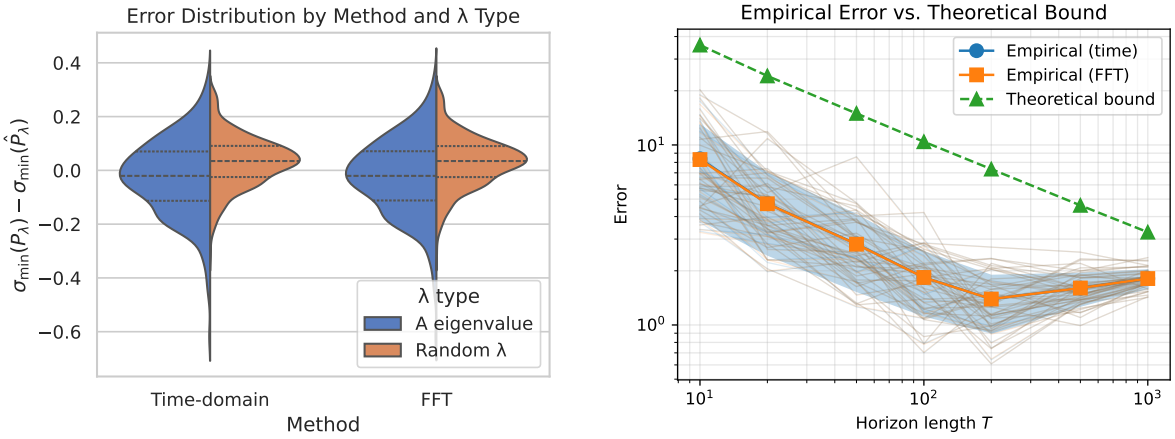


FIGURE 1. Element-wise comparison of the true Hautus matrix $\mathbf{P}_\lambda$ (left) and its time-domain estimate $\widetilde{\mathbf{P}}_\lambda$ (center), with absolute difference (right). The small difference magnitudes confirm the accuracy of the cross-moment estimator.

4.2. **Error Convergence Rate.** Proposition 3.1 predicts that the estimation error $\|\widetilde{\mathbf{P}}_\lambda - \mathbf{P}_\lambda\|_2$ decays at the rate $\mathcal{O}(T^{-1/2})$ as the observation horizon T increases. We validate this theoretical prediction by simulating systems with $n = 5$ states, $m = 3$ inputs, and $q = 2$ noise dimensions across horizon lengths $T \in \{10, 20, 50, 100, 200, 500, 1000\}$. For each value of T, we run 50 Monte Carlo trials with independent noise realizations and random test eigenvalues λ.

Figure 2 (right) presents the results on a log-log scale. The empirical mean errors for both time-domain and FFT methods closely follow the $T^{-1/2}$ rate. The shaded region indicates ± 1 standard deviation across trials. We also overlay the theoretical bound from Proposition 3.1 (computed with $\delta = 0.05$), which provides a valid upper envelope for the empirical errors. A linear regression on the log-transformed data yields an estimated slope of approximately -0.5, confirming the theoretical rate.



(A) Error distribution across 50 trajectories (eigenvalues of $\mathbf{A}$ and random test points).

(B) Mean error versus horizon T (log-log) with ± 1 std shading and the theoretical bound.

FIGURE 2. Summary of estimation accuracy: variability across trials (left) and $T^{-1/2}$ convergence (right).

4.3. **Finite Candidate Eigenvalue Check.** A key practical advantage of the cross-moment formulation is that controllability can be verified by testing only a finite set of candidate eigenvalues rather than all $\lambda \in \mathbb{C}$. Specifically, the operator formulation developed in Section 2.3 shows that rank deficiency of the pencil $\mathcal{P}(\lambda) = \dot{\mathcal{X}} - \lambda\mathcal{X}$ can only occur when λ belongs to the spectrum of the data-driven matrix

$$\mathbf{K} := \mathbf{S}_X^{-1}\mathbf{M}_X, \quad \text{where} \quad \mathbf{S}_X := \int_0^T x(t)x(t)^\top \,\mathrm{d}t, \quad \mathbf{M}_X := \int_0^T x(t)\,\mathrm{d}x(t)^\top.$$

In the noise-free case $\mathrm{d}x(t) = \dot{x}(t)\,\mathrm{d}t$, so $\mathbf{M}_X$ reduces to $\int_0^T x(t)\dot{x}(t)^\top \,\mathrm{d}t$; under Itô noise, $\mathbf{M}_X$ is estimated directly from state increments and includes an additional martingale term.

We simulate a controllable system with $n = 6$, $m = 3$, $q = 2$ over a horizon $T = 100$. The candidate eigenvalues $\sigma(\mathbf{K})$ are computed from the trajectory data and compared with the true eigenvalues $\sigma(\mathbf{A})$. For each candidate λ_i, we compute the estimated minimum singular value $\sigma_{\min}(\widetilde{\mathbf{P}}_{\lambda_i})$ and compare with the true value $\sigma_{\min}(\mathbf{P}_{\lambda_i})$.

The results show excellent agreement between estimated and true singular values across all candidates. The candidates $\sigma(\mathbf{K})$ cluster near the true eigenvalues $\sigma(\mathbf{A})$ in the complex plane, validating the theoretical prediction that the data-driven eigenvalue estimates converge to the true system eigenvalues. Since $\sigma_{\min}(\widetilde{\mathbf{P}}_\lambda) > 0$ for all candidates, the system is correctly identified as controllable.

4.4. **Comparison of Time-Domain and FFT Methods.** We compare two computational approaches for evaluating the cross-moments: (i) the *time-domain method*, which computes $\widetilde{\mathbf{H}}_\lambda(T)$ via direct time stepping, and (ii) the *FFT method*, which leverages frequency-domain representations (via (16) and Parseval-type identities).

For a system with $n = 8$ states, $m = 4$ inputs, $q = 2$ noise dimensions, and horizon $T = 100$ with time step $\mathrm{d}t = 0.02$ (yielding 5001 sample points), we test both methods across various values of λ. Both methods achieve comparable accuracy, with estimation errors typically differing by less than 10%.

The FFT method requires a frequency cutoff $\omega_{\max}$ to truncate the integral. Figure 3a shows the effect of $\omega_{\max}$ on the estimation error for $\lambda = 0.3 + 0.8\mathrm{i}$. As expected, larger cutoffs improve accuracy, with diminishing returns beyond $\omega_{\max} \approx 100$. For compactness, we place the cutoff summary next to a representative 3D phase portrait (Figure 3b).

In terms of computational efficiency, the FFT method can be advantageous when evaluating many λ values, since $\widehat{x}(\mathrm{i}\omega)$ and $\widehat{z}(\mathrm{i}\omega)$ can be computed once and reused across all λ.

4.5. **Trajectory Visualization.** To provide intuition for the controlled stochastic dynamics, Figure 3b displays representative phase portraits for a 9-dimensional system with $m = 4$ inputs and $q = 2$ noise dimensions, simulated over a horizon $T = 50$. The drift matrix $\mathbf{A}$ has eigenvalues with real parts near -0.05, corresponding to slowly decaying modes.

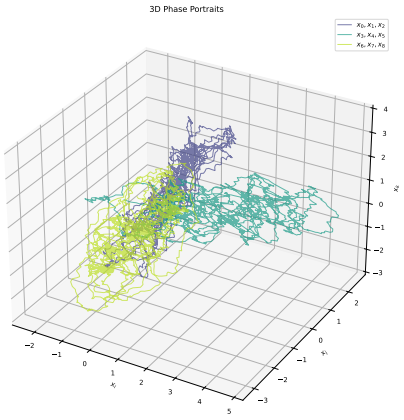
The phase portraits demonstrate that the sinusoidal input, combined with the stochastic forcing, provides sufficient excitation to explore the state space. This persistent excitation ensures that the stacked Gramian $\widetilde{\mathbf{S}}_Z(T)$ is well-conditioned, which is crucial for accurate estimation of the Hautus matrix and reliable controllability certification.

References

[1] Chi-Tsong Chen, ed. *Linear System Theory and Design*. 3rd ed. The Oxford Series in Electrical and Computer Engineering. New York: Oxford University Press, 1999. 334 pp. ISBN: 978-0-19-511777-6.

$\omega_{\max}$	$\sigma_{\min}(\widetilde{\mathbf{P}}_\lambda)$	$\|\widetilde{\mathbf{P}}_\lambda - \mathbf{P}_\lambda\|_2$
10	0.412	0.089
20	0.428	0.051
50	0.435	0.027
100	0.438	0.018
200	0.439	0.015
∞ (all)	0.439	0.014

(A) Effect of frequency cutoff $\omega_{\max}$ on FFT method accuracy for $\lambda = 0.3 + 0.8\mathrm{i}$.



(B) Three-dimensional phase portraits for (x_0, x_1, x_2), (x_3, x_4, x_5), and (x_6, x_7, x_8).

FIGURE 3. Frequency cutoff sensitivity and trajectory visualization. Left: accuracy of the FFT-based estimator improves with larger cutoff values, with diminishing returns beyond $\omega_{\max} \approx 100$. Right: representative phase portraits for a 9-dimensional controlled stochastic system.

[2] Andrea Iannelli et al. "Design of Input for Data-Driven Simulation with Hankel and Page Matrices". In: *2021 60th IEEE Conference on Decision and Control (CDC)*. 2021 60th IEEE Conference on Decision and Control (CDC). Dec. 2021, pp. 139–145. DOI: 10.1109/CDC45484.2021.9683709. URL: https://ieeexplore.ieee.org/document/9683709 (visited on 12/15/2025).

[3] R. E. Kalman. "Mathematical Description of Linear Dynamical Systems". In: *Journal of the Society for Industrial and Applied Mathematics Series A Control* 1.2 (Jan. 1963), pp. 152–192. ISSN: 0887-4603. DOI: 10.1137/0301010. URL: http://epubs.siam.org/doi/10.1137/0301010 (visited on 12/08/2025).

[4] Vikas Kumar Mishra et al. "Data-Driven Tests for Controllability". In: *IEEE Control Systems Letters* 5.2 (Apr. 2021), pp. 517–522. ISSN: 2475-1456. DOI: 10.1109/LCSYS.2020.3003770. URL: https://ieeexplore.ieee.org/document/9121339/ (visited on 12/11/2025).

[5] P. Rapisarda et al. "A "Fundamental Lemma" for Continuous-Time Systems, with Applications to Data-Driven Simulation". In: *Systems & Control Letters* 179 (Sept. 2023), p. 105603. ISSN: 01676911. DOI: 10.1016/j.sysconle.2023.105603. URL: https://linkinghub.elsevier.com/retrieve/pii/S0167691123001500 (visited on 01/20/2026).

[6] P. Rapisarda et al. "A Persistency of Excitation Condition for Continuous-Time Systems". In: *IEEE Control Systems Letters* 7 (2023), pp. 589–594. ISSN: 2475-1456. DOI: 10.1109/LCSYS.2022.3205550. URL: https://ieeexplore.ieee.org/document/9882335/ (visited on 01/20/2026).

[7] Amir Shakouri et al. "A New Perspective on Willems' Fundamental Lemma: Universality of Persistently Exciting Inputs". In: *IEEE Control Systems Letters* 9 (2025), pp. 583–588. ISSN: 2475-1456. DOI: 10.1109/LCSYS.2025.3576276. arXiv: 2503.12489 [math]. URL: http://arxiv.org/abs/2503.12489 (visited on 12/08/2025).

[8] Leonard M. Silverman et al. "Controllability, Observability and Stability of Linear Systems". In: *SIAM Journal on Control* 5.1 (1967), pp. 121–136. DOI: 10.1137/0305010. URL: https://epubs.siam.org/doi/10.1137/0306010.

[9] Harry L. Trentelman et al. *Control Theory for Linear Systems*. Springer, 2001. ISBN: 978-1-85233-316-4. URL: https://research.utwente.nl/en/publications/control-theory-for-linear-systems/ (visited on 12/17/2025).

[10] Henk J. van Waarde et al. "Data Informativity: A New Perspective on Data-Driven Analysis and Control". In: *IEEE Transactions on Automatic Control* 65.11 (Nov. 2020), pp. 4753–4768. ISSN: 1558-2523. DOI: 10.1109/TAC.2020.2966717. URL: https://ieeexplore.ieee.org/document/8960476/ (visited on 12/11/2025).

[11] Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge University Press, 2018. ISBN: 978-1-108-41519-4.

[12] Henk J. van Waarde et al. *Willems' Fundamental Lemma for State-space Systems and Its Extension to Multiple Datasets*. May 7, 2020. DOI: 10.48550/arXiv.2002.01023. arXiv: 2002.01023 [math]. URL: http://arxiv.org/abs/2002.01023 (visited on 12/08/2025). Pre-published.

[13] Jan C. Willems et al. "A Note on Persistency of Excitation". In: *Systems & Control Letters* 54.4 (Apr. 2005), pp. 325–329. ISSN: 01676911. DOI: 10.1016/j.sysconle.2004.09.003. URL: https://linkinghub.elsevier.com/retrieve/pii/S0167691104001434 (visited on 12/08/2025).

Appendix A. Proofs

A.1. Proof of the continuous-time Hautus test.

Proof. If $(\mathbf{A}, \mathbf{B})$ is not controllable, then by the (continuous-time) Hautus test there exist $\lambda \in \mathbb{C}$ and $v \neq 0$ such that $v^* \mathbf{P}_\lambda = 0$. Using (3), this gives $v^* \mathbf{H}_\lambda(T) = v^* \mathbf{P}_\lambda \mathbf{S}_Z(T) = 0$, hence $\operatorname{rank}(\mathbf{H}_\lambda(T)) < n$. This proves the contrapositive of the first implication in Theorem 2.1.

Conversely, assume $(\mathbf{A}, \mathbf{B})$ is controllable and $\mathbf{S}_Z(T)$ is invertible. Then $\mathbf{P}_\lambda$ has full row rank for every $\lambda \in \mathbb{C}$ and (3) implies

$$\operatorname{rank}(\mathbf{H}_\lambda(T)) = \operatorname{rank}(\mathbf{P}_\lambda \mathbf{S}_Z(T)) = \operatorname{rank}(\mathbf{P}_\lambda) = n \qquad \forall \lambda \in \mathbb{C},$$

which proves the converse. □

A.2. Proof of the cross-moment error bound.

Proof. Let $\mathbf{M}(T) := \int_0^T \mathrm{d}W(t)\, z(t)^\top \in \mathbb{R}^{q \times (n+m)}$. Conditional on the path $\{z(t)\}_{t \in [0,T]}$, $\mathbf{M}(T)$ is a centered Gaussian matrix whose rows are independent and satisfy

$$\mathbb{E}\Big[\mathbf{M}(T)_{i,:}^\top \mathbf{M}(T)_{i,:} \mid z\Big] = \int_0^T z(t) z(t)^\top \,\mathrm{d}t = \widetilde{\mathbf{S}}_Z(T), \qquad i = 1, \dots, q.$$

Therefore, conditional on z, we have the distributional identity $\mathbf{M}(T) \overset{d}{=} \mathbf{G} \widetilde{\mathbf{S}}_Z(T)^{1/2}$ where $\mathbf{G} \in \mathbb{R}^{q \times (n+m)}$ has i.i.d. $\mathcal{N}(0,1)$ entries. Using (12),

$$\widetilde{\mathbf{P}}_\lambda(T) - \mathbf{P}_\lambda \overset{d}{=} \boldsymbol{\beta}\, \mathbf{G} \widetilde{\mathbf{S}}_Z(T)^{-1/2},$$

conditional on z, hence

$$\|\widetilde{\mathbf{P}}_\lambda(T) - \mathbf{P}_\lambda\|_2 \leq \frac{\|\boldsymbol{\beta}\|_2}{\sqrt{\sigma_{\min}(\widetilde{\mathbf{S}}_Z(T))}}\, \|\mathbf{G}\|_2.$$

The standard Gaussian matrix bound $\mathbb{P}\Big(\|\mathbf{G}\|_2 \geq \sqrt{q} + \sqrt{n+m} + t\Big) \leq e^{-t^2/2}$ (valid for all $t \geq 0$; see, e.g., [11, Theorem 4.4.5]) yields (15) by choosing $t = \sqrt{2\log(1/\delta)}$ and using $\sigma_{\min}(\widetilde{\mathbf{S}}_Z(T)) = T\sigma_{\min}(\bar{\mathbf{S}}_Z(T))$. □

A.3. **Proof of the finite candidate set result.**

Proof. Fix $\lambda \in \mathbb{C}$ and assume $\operatorname{rank}(\mathcal{P}(\lambda)) < n$. Since the codomain is $\mathbb{C}^n$, the left nullspace is nontrivial, so there exists $0 \neq w \in \mathbb{C}^n$ with $\mathcal{P}(\lambda)^* w = 0$, i.e. $\dot{\mathcal{X}}^* w = \overline{\lambda}\mathcal{X}^* w$. Left-multiply by $(\mathcal{X}\mathcal{X}^*)^{-1}\mathcal{X}$ to obtain

$$\mathbf{K}w = (\mathcal{X}\mathcal{X}^*)^{-1}\mathcal{X}\dot{\mathcal{X}}^* w = \overline{\lambda}\,(\mathcal{X}\mathcal{X}^*)^{-1}\mathcal{X}\mathcal{X}^* w = \overline{\lambda}\,w,$$

so $\overline{\lambda} \in \sigma(\mathbf{K})$. Since x is real-valued, the matrices $\mathcal{X}\mathcal{X}^*$ and $\mathcal{X}\dot{\mathcal{X}}^*$ are real, hence $\mathbf{K} \in \mathbb{R}^{n\times n}$. For real matrices, eigenvalues come in conjugate pairs, so $\sigma(\mathbf{K})$ is closed under complex conjugation. Therefore, $\lambda = \overline{\overline{\lambda}} \in \sigma(\mathbf{K})$. □

A.4. **Proof of the quantitative margin bound.**

Proof. Fix $w \in \mathbb{C}^n$. Since $w^*(\dot{\mathcal{X}} - \lambda\mathcal{X})$ is a bounded linear functional on $L^2(0,T)$, we have

$$\big\|w^*(\dot{\mathcal{X}} - \lambda\mathcal{X})\big\| = \big\|(\dot{\mathcal{X}}^* - \overline{\lambda}\,\mathcal{X}^*)w\big\|_{L^2(0,T)}.$$

Moreover,

$$(\mathbf{K} - \overline{\lambda}\,\mathbf{I})w = (\mathcal{X}\mathcal{X}^*)^{-1}\mathcal{X}(\dot{\mathcal{X}}^* - \overline{\lambda}\,\mathcal{X}^*)w,$$

hence

$$\|(\mathbf{K} - \overline{\lambda}\,\mathbf{I})w\| \leq \|(\mathcal{X}\mathcal{X}^*)^{-1}\mathcal{X}\|\,\big\|(\dot{\mathcal{X}}^* - \overline{\lambda}\,\mathcal{X}^*)w\big\|_{L^2(0,T)}.$$

Using (4), $\|(\dot{\mathcal{X}}^* - \overline{\lambda}\,\mathcal{X}^*)w\|^2_{L^2(0,T)} = w^*\mathbf{G}_\lambda(u)w$, so

$$w^*\mathbf{G}_\lambda(u)w \geq \frac{\|(\mathbf{K} - \overline{\lambda}\,\mathbf{I})w\|^2}{\|(\mathcal{X}\mathcal{X}^*)^{-1}\mathcal{X}\|^2}.$$

Since $\mathbf{S}_Z(T) \succ 0$, (6) gives

$$\mathbf{G}_\lambda(u) = \mathbf{H}_\lambda(T)\mathbf{S}_Z(T)^{-1}\mathbf{H}_\lambda(T)^*.$$

Therefore, for every $w \in \mathbb{C}^n$,

$$w^*\mathbf{G}_\lambda(u)w = w^*\mathbf{H}_\lambda(T)\mathbf{S}_Z(T)^{-1}\mathbf{H}_\lambda(T)^* w \leq \frac{1}{\sigma_{\min}(\mathbf{S}_Z(T))}\,\|\mathbf{H}_\lambda(T)^* w\|_2^2,$$

i.e.,

$$\|\mathbf{H}_\lambda(T)^* w\|_2^2 \geq \sigma_{\min}(\mathbf{S}_Z(T))\,w^*\mathbf{G}_\lambda(u)w.$$

Combining with the lower bound on $w^*\mathbf{G}_\lambda(u)w$ above and taking the minimum over $\|w\|_2 = 1$ gives

$$\sigma_{\min}\Big(\mathbf{H}_\lambda(T)\Big)^2 = \min_{\|w\|_2=1}\|\mathbf{H}_\lambda(T)^* w\|_2^2 \geq \sigma_{\min}(\mathbf{S}_Z(T))\,\frac{\sigma_{\min}(\mathbf{K} - \overline{\lambda}\,\mathbf{I})^2}{\|(\mathcal{X}\mathcal{X}^*)^{-1}\mathcal{X}\|^2},$$

which yields the claimed bound after taking square roots. The final claim follows since $\sigma_{\min}(\mathbf{K} - \overline{\lambda}\,\mathbf{I}) > 0$ whenever $\lambda \notin \sigma(\mathbf{K})$. □

A.5. **Proof of the L^2-budget isotropic design.**

Proof. Since $\mathbf{S}_U(u) \succeq 0$, we have $\lambda_{\min}(\mathbf{S}_U(u)) \le \frac{1}{m}\operatorname{tr}(\mathbf{S}_U(u))$. Moreover,

$$\begin{aligned}\operatorname{tr}(\mathbf{S}_U(u)) &= \int_0^T \operatorname{tr}\Big(u(t)u(t)^\top\Big)\,\mathrm{d}t \\ &= \int_0^T \|u(t)\|_2^2\,\mathrm{d}t \\ &= \|u\|_{L^2(0,T)}^2 \le 1,\end{aligned}$$

which gives the upper bound. For achievability, pick an orthonormal set $\{\varphi_i\}_{i=1}^m \subset L^2(0,T)$ and define $u(t) := \frac{1}{\sqrt{m}}[\varphi_1(t)\ \cdots\ \varphi_m(t)]^\top$; then $\mathbf{S}_U(u) = \frac{1}{m}\mathbf{I}_m$. □

A.6. **Proof of the H^1-budget isotropic design and characterization of optimizers.**

Proof. Expand u in the orthonormal Neumann basis $\{\psi_k\}_{k\ge0}$ as $u(t) = \sum_{k\ge0}\psi_k(t)a_k$ with coefficients $a_k \in \mathbb{R}^m$. Then

$$\mathbf{S}_U(u) = \sum_{k\ge0} a_k a_k^\top, \qquad \|u\|_{H^1(0,T)}^2 = \sum_{k\ge0}\left(1 + \left(\tfrac{k\pi}{T}\right)^2\right)\|a_k\|_2^2.$$

Let $w_k := 1 + (\frac{k\pi}{T})^2$ and $W := \operatorname{diag}(w_0, w_1, \dots)$, and define $A := [a_0\ a_1\ \cdots] \in \mathbb{R}^{m\times\infty}$, so that $\mathbf{S}_U(u) = AA^\top$ and $\|u\|_{H^1(0,T)}^2 = \operatorname{tr}(AWA^\top)$. Write the compact SVD of A as

$$A = \mathbf{S}_U(u)^{1/2}R, \qquad RR^\top = \mathbf{I}_m,$$

so that $\Pi := R^\top R$ is a rank-m orthogonal projector. Then

$$\begin{aligned}\|u\|_{H^1(0,T)}^2 &= \operatorname{tr}\Big(\mathbf{S}_U(u)^{1/2}RWR^\top\mathbf{S}_U(u)^{1/2}\Big) \\ &= \operatorname{tr}\Big(\mathbf{S}_U(u)\,RWR^\top\Big) \\ &\ge \lambda_{\min}(\mathbf{S}_U(u))\ \operatorname{tr}(RWR^\top) \\ &= \lambda_{\min}(\mathbf{S}_U(u))\ \operatorname{tr}(W\Pi).\end{aligned}$$

Minimizing $\operatorname{tr}(W\Pi)$. Since W is diagonal with strictly increasing entries w_k, we have

$$\min_{\substack{\Pi=\Pi^\top=\Pi^2\\ \operatorname{rank}(\Pi)=m}} \operatorname{tr}(W\Pi) \;=\; \sum_{k=0}^{m-1} w_k,$$

and, moreover, the minimizer is unique: $\Pi_\star = \operatorname{diag}(\mathbf{I}_m, 0, 0, \dots)$, i.e., projection onto the span of the first m coordinate vectors. Therefore

$$\|u\|_{H^1(0,T)}^2 \;\ge\; \lambda_{\min}(\mathbf{S}_U(u))\sum_{k=0}^{m-1} w_k.$$

Under the budget $\|u\|_{H^1(0,T)}^2 \le 1$ this yields

$$\lambda_{\min}(\mathbf{S}_U(u)) \;\le\; \alpha \;:=\; \Big(\sum_{k=0}^{m-1} w_k\Big)^{-1} \;=\; \Big(\sum_{k=0}^{m-1}\Big(1 + \Big(\tfrac{k\pi}{T}\Big)^2\Big)\Big)^{-1}.$$

Tightness and characterization of all optimizers. Suppose the bound is tight, i.e., $\|u\|_{H^1(0,T)}^2 \le 1$ and $\lambda_{\min}(\mathbf{S}_U(u)) = \alpha$. Then equality must hold throughout the chain of inequalities above. In particular:

(i) Low-mode support. Equality in $\mathrm{tr}(W\Pi) \geq \sum_{k=0}^{m-1} w_k$ forces $\Pi = \Pi_\star$ by uniqueness of the minimizer. Equivalently, the row space of R lies in the span of the first m coordinates, which implies that A has no columns beyond the first m:

$$a_k = 0 \quad \text{for all } k \geq m, \qquad \text{hence} \qquad u(t) = \sum_{k=0}^{m-1} \psi_k(t) a_k.$$

(ii) Isotropy of the Gramian. Equality in

$$\mathrm{tr}\Big(\mathbf{S}_U(u)\, RWR^\top\Big) \;\geq\; \lambda_{\min}(\mathbf{S}_U(u))\; \mathrm{tr}(RWR^\top)$$

forces $\mathbf{S}_U(u)$ to be isotropic. Indeed, since $RWR^\top \succ 0$ (because $w_k > 0$ and R has full row rank), write $\mathbf{S}_U(u) = U\,\mathrm{diag}(s_1, \ldots, s_m)U^\top$ with $s_m = \lambda_{\min}(\mathbf{S}_U(u))$. Then

$$\mathrm{tr}\Big((\mathbf{S}_U(u) - s_m \mathbf{I}_m)\, RWR^\top\Big) \geq 0,$$

with equality if and only if $\mathbf{S}_U(u) - s_m\mathbf{I}_m = 0$ since $RWR^\top \succ 0$. Hence $\mathbf{S}_U(u) = \lambda_{\min}(\mathbf{S}_U(u))\mathbf{I}_m = \alpha\mathbf{I}_m$.

(iii) Explicit form of all maximizers. From (i), A effectively reduces to an $m \times m$ matrix $A_m = [a_0 \;\cdots\; a_{m-1}]$. From (ii), we have $A_m A_m^\top = \mathbf{S}_U(u) = \alpha\mathbf{I}_m$, so $A_m = \sqrt{\alpha}\, Q$ for some $Q \in O(m)$. Defining

$$\Psi(t) := \begin{bmatrix} \psi_0(t) \\ \vdots \\ \psi_{m-1}(t) \end{bmatrix} \in \mathbb{R}^m,$$

we obtain the complete optimizer family

$$u(t) = \sqrt{\alpha}\, Q\, \Psi(t), \qquad Q \in O(m),$$

which satisfies $\|u\|^2_{H^1(0,T)} = \alpha \sum_{k=0}^{m-1} w_k = 1$ and $\mathbf{S}_U(u) = \alpha\mathbf{I}_m$. □