

EXPERIMENT DESIGN WITH CONTROL GUARANTEES

DANIEL Y ENRIQUE

ABSTRACT. We study deterministic experiment design for continuous-time linear time-invariant systems with the goal of certifying controllability properties directly from data. We develop a continuous-time, data-driven analogue of the Hautus (PBH) test based on residual Gramians and a derivative-free cross-moment formulation, and we provide basic 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

1. Introduction	2
1.1. Main Contributions	3
2. Problem Statement: Data-Driven Controllability Certification	4
2.1. Noise-free cross-moment formulation	5
2.2. Stochastic model and Itô residual	5
2.3. A Fourier-domain approximation without derivatives.	7
2.4. An operator formulation and finitely many candidate λ	7
2.5. Best conditioning of control inputs	8
3. Numerical Experiments	10
3.1. Validation of the Data-Driven Hautus Test	10
3.2. Error Convergence Rate	10
3.3. Finite Candidate Eigenvalue Check	11
3.4. Comparison of Time-Domain and FFT Methods	11
3.5. Trajectory Visualization	12
References	12
Appendix A. Proofs	14
A.1. Proof of the continuous-time Hautus test	14
A.2. Proof of the cross-moment error bound	14
A.3. Proof of the finite candidate set result	14
A.4. Proof of Corollary 2.4	14
A.5. Proof of the quantitative margin bound	15
A.6. Proof of the L^2 -budget isotropic design	15
A.7. Proof of the H^1 -budget isotropic design	16

[†]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

Motivation: learning models that are *control-relevant*. 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. 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.

Persistent excitation and universality in data-driven control. 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 [15]. 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 [15]. 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 [14, 12]. 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 [9]. For continuous-time systems, related identifiability conditions and continuous-time variants of Willems-type results have also been investigated [8, 7].

Offline vs. online experiment design. Most classical guidance concerns *offline* experiment design, where one selects a fixed open-loop excitation signal before collecting data. Offline designs are attractive because they are universal and simple to implement, but they can be conservative and sample-inefficient. This has motivated *online* (adaptive) experiment design methods that adjust the input in real time based on the observed data to accelerate identifiability and reduce the amount of data required [11, 3]. In fact, sharp sample-efficiency results have recently been obtained for the length of an informative trajectory needed for linear system identification [1]. These developments emphasize that experiment design should be treated as a first-class component of safe data-driven control, rather than a preprocessing step.

Controllability certification from data. 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 [12, 4]. More recently, Mishra et al. developed algebraic data-driven tests for controllability of LTI systems from batches of measurements [6]. 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

$$(1) \quad \dot{x}(t) = \mathbf{A}x(t) + \mathbf{B}u(t), \quad x(0) = x_0.$$

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 $\text{rank}(\mathbf{C})$ determines the dimension of the reachable set from the origin [5, 2]. An equivalent characterization is given in terms of the classical Hautus (Popov–Belevitch–Hautus, PBH) test [10, Theorem 3.13], which characterizes controllability and stabilizability as follows:

$$(2) \quad \begin{aligned} (\mathbf{A}, \mathbf{B}) \text{ is controllable} &\Leftrightarrow \text{rank}[\mathbf{A} - \lambda\mathbf{I}, \mathbf{B}] = n \quad \forall \lambda \in \mathbb{C}, \\ (\mathbf{A}, \mathbf{B}) \text{ is stabilizable} &\Leftrightarrow \text{rank}[\mathbf{A} - \lambda\mathbf{I}, \mathbf{B}] = n \quad \forall \lambda \in \mathbb{C} \text{ with } \Re(\lambda) \geq 0. \end{aligned}$$

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

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

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

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

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 well-conditioned, then the rank of $\mathbf{H}_\lambda(u)$ reveals the rank of the Hautus matrix. However, the quality of this estimate depends 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 continuous-time data-driven Hautus margin.** We provide a *continuous-time* and *data-driven* analogue of the Hautus test (2):

$$(3) \quad (\mathbf{A}, \mathbf{B}) \text{ is controllable} \Leftrightarrow \text{rank}(\mathbf{H}_\lambda(u)) = n \quad \forall \lambda \in \mathbb{C}.$$

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

$$\mathbf{K} := \left(\int_0^T x(t)x(t)^* dt \right)^{-1} \left(\int_0^T x(t)\dot{x}(t)^* dt \right).$$

We further develop a derivative-free formulation using Fourier transform.

- **Input design for conditioning the Hautus margin.** We provide sharp, model-agnostic designs for conditioning the input, which improves the data-driven test (3):
 (1) Under an L^2 energy budget $\|u\|_{L^2} \leq 1$, the best possible conditioning is achieved by any input with m orthonormal time functions (Proposition 2.7), yielding

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

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

$$u(t) = \sqrt{\alpha} Q \begin{bmatrix} \psi_0(t) \\ \vdots \\ \psi_{m-1}(t) \end{bmatrix}, \quad \text{where} \quad \psi_k(t) := \sqrt{\frac{2}{T}} \cos\left(\frac{k\pi t}{T}\right),$$

with $Q \in \mathbb{R}^{m \times m}$ an orthogonal matrix and α a normalization term (Proposition 2.8).

- **Concentration bounds under Itô noise.** Consider the Itô noise model

$$dx(t) = (\mathbf{A}x(t) + \mathbf{B}u(t)) dt + \beta dW(t),$$

where W is a q -dimensional standard Brownian motion and $\beta \in \mathbb{R}^{n \times q}$ is the noise intensity matrix. In Proposition 2.2, we prove that with probability at least $1 - \delta$,

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

where $\mathbf{P}_\lambda := [\mathbf{A} - \lambda \mathbf{I}, \mathbf{B}]$, $\hat{\mathbf{P}}_\lambda(T)$ is the data-driven estimate and $\bar{\mathbf{S}}_Z(T) := \frac{1}{T} \int_0^T z(t) z(t)^\top dt$ is its normalized Gramian.

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

2. PROBLEM STATEMENT: DATA-DRIVEN CONTROLLABILITY CERTIFICATION

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

$$(4) \quad \dot{x}(t) = \mathbf{A}x(t) + \mathbf{B}u(t), \quad x(0) = x_0,$$

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. Given an initial condition $x(0) \in \mathbb{R}^n$, the state trajectory admits the closed-form solution

$$(5) \quad x(t) = e^{\mathbf{A}t} x(0) + \int_0^t e^{\mathbf{A}(t-\tau)} \mathbf{B} u(\tau) d\tau.$$

A fundamental property of the system is **controllability**: the ability to steer the state to any point in \mathbb{R}^n using an appropriate input signal. More precisely, the reachable (controllable) subspace from the origin is the image of the Kalman controllability matrix

$$(6) \quad \mathbf{C} := [\mathbf{B}, \mathbf{A}\mathbf{B}, \mathbf{A}^2\mathbf{B}, \dots, \mathbf{A}^{n-1}\mathbf{B}] \in \mathbb{R}^{n \times nm}.$$

Let $r := \text{rank}(\mathbf{C})$ denote the dimension of this subspace [5, 2]; the pair (\mathbf{A}, \mathbf{B}) is said to be *controllable* if and only if $r = n$. Our objective is to design input signals $u(t)$ such that the observed trajectories reveal the controllability rank r from finitely many state snapshots. When only output measurements $y(t) = \mathbf{C}_y x(t)$ are available (with a known output matrix $\mathbf{C}_y \in \mathbb{R}^{p \times n}$), the same rank analysis applies to the output-restricted controllability matrix $\mathbf{C}_y \mathbf{C}$.

A natural route to certifying controllability from data is via terminal-state information under structural assumptions, such as a fixed initial condition and structured input families. In this section, we develop an alternative approach based on rank conditions that are equivalent to controllability and stabilizability.

Recall that for the continuous-time LTI system (4), the classical Hautus (Popov–Belevitch–Hautus, PBH) test [10, Theorem 3.13] characterizes controllability and stabilizability. Here we state a continuous-time version based on time-domain moments of the state and input signals.

Data model. We assume access to $u(\cdot)$ and the resulting state $x(\cdot)$ on a finite horizon $[0, T]$ (either as a continuous-time record or via sufficiently fine sampling so that the integrals below can be approximated numerically). When \dot{x} is not directly available or is too noisy to estimate reliably, we rely on the cross-moment formulation in Section 2.2, which only uses increments of x in the stochastic setting and avoids forming \dot{x} in the deterministic setting.

Outline. The remainder of this section is organized as follows. In Section 2.1, we introduce a residual $\mathbf{H}_\lambda(u)$ and show that having full row rank for all $\lambda \in \mathbb{C}$ is equivalent to controllability under a mild data-richness condition. Section 2.2 develops a derivative-free cross-moment

formulation—applicable in both deterministic and Itô settings—that enables estimating $\mathbf{H}_\lambda(u)$ and a corresponding Hautus margin directly from measured data. In Section 2.4, we present an operator formulation that reduces checking all $\lambda \in \mathbb{C}$ to a finite candidate set. Finally, Section 2.5 discuss optimal input design principles for conditioning the input Gramian under L^2 and H^1 budgets, respectively.

2.1. Noise-free cross-moment formulation. 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 $x : [0, T] \rightarrow \mathbb{R}^n$ with $x, \dot{x} \in L^2(0, T; \mathbb{R}^n)$. For $\lambda \in \mathbb{C}$ define the residual signal and its cross-moment

$$m_\lambda(t) := \dot{x}(t) - \lambda x(t) \in \mathbb{C}^n, \quad \mathbf{H}_\lambda(u) := \int_0^T m_\lambda(t) z(t)^* dt \in \mathbb{C}^{n \times (n+m)}.$$

Define the stacked signal and its Gramian

$$z(t) := \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} \in \mathbb{R}^{n+m}, \quad \mathbf{S}_Z(u) := \int_0^T z(t) z(t)^\top dt \in \mathbb{R}^{(n+m) \times (n+m)}.$$

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

$$m_\lambda(t) = (\mathbf{A} - \lambda \mathbf{I})x(t) + \mathbf{B}u(t) = \mathbf{P}_\lambda z(t) \quad \text{for a.e. } t \in (0, T),$$

and therefore

$$(7) \quad \mathbf{H}_\lambda(u) = \mathbf{P}_\lambda \mathbf{S}_Z(u).$$

If $\mathbf{S}_Z(u)$ is invertible, then right-multiplication by $\mathbf{S}_Z(u)$ preserves row rank, so (7) yields

$$\text{rank}(\mathbf{H}_\lambda(u)) = \text{rank}(\mathbf{P}_\lambda), \quad \forall \lambda \in \mathbb{C}.$$

Using this, we can now state the continuous-time Hautus test.

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

$$(\mathbf{A}, \mathbf{B}) \text{ is controllable} \iff \text{rank}(\mathbf{H}_\lambda(u)) = n \text{ for all } \lambda \in \mathbb{C}.$$

Proof. See Appendix A.1. □

2.2. Stochastic model and Itô residual. Another alternative to direct state-derivative measurements is to estimate \mathbf{P}_λ , and hence $\mathbf{H}_\lambda(u)$, 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

$$dx(t) = (\mathbf{A}x(t) + \mathbf{B}u(t)) dt + \beta dW(t), \quad t \in [0, T],$$

where W is a q -dimensional standard Brownian motion and $\beta \in \mathbb{R}^{n \times q}$ is constant. For $\lambda \in \mathbb{C}$, define the Itô residual differential

$$(8) \quad dy_\lambda(t) := dx(t) - \lambda x(t) dt = \mathbf{P}_\lambda z(t) dt + \beta dW(t).$$

Cross-moment factorization. Define the (matrix-valued) cross-moment

$$(9) \quad \mathbf{H}_\lambda(T) := \int_0^T dy_\lambda(t) z(t)^\top \in \mathbb{C}^{n \times (n+m)}$$

$$(10) \quad = \mathbf{P}_\lambda \mathbf{S}_Z(u) + \beta \int_0^T dW(t) z(t)^\top.$$

The last term is a matrix-valued martingale with $\mathbb{E}[\int_0^T dW(t) z(t)^\top] = 0$, hence $\mathbf{H}_\lambda(T) - \mathbf{P}_\lambda \mathbf{S}_Z(u)$ has zero mean. If x is absolutely continuous and $\beta = 0$, then $dy_\lambda(t) = m_\lambda(t) dt$. If moreover $\mathbf{S}_Z(u)$ is invertible, then the deterministic residual Gramian

$$\mathbf{G}_\lambda(u) := \int_0^T m_\lambda(t) m_\lambda(t)^* dt = \mathbf{P}_\lambda \mathbf{S}_Z(u) \mathbf{P}_\lambda^*$$

admits the cross-moment representation

$$(11) \quad \mathbf{G}_\lambda(u) = \mathbf{H}_\lambda(T) \mathbf{S}_Z(u)^{-1} \mathbf{H}_\lambda(T)^*.$$

Let us estimate \mathbf{P}_λ . Assume that $\mathbf{S}_Z(u)$ is invertible and define

$$(12) \quad \hat{\mathbf{P}}_\lambda(T) := \mathbf{H}_\lambda(T) \mathbf{S}_Z(u)^{-1}.$$

Then (10) implies the exact error identity

$$(13) \quad \hat{\mathbf{P}}_\lambda(T) - \mathbf{P}_\lambda = \beta \left(\int_0^T dW(t) z(t)^\top \right) \mathbf{S}_Z(u)^{-1}.$$

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

$$(14) \quad \mathbb{E} \left\| \frac{1}{T} \int_0^T dW(t) z(t)^\top \right\|_F^2 = \frac{q}{T^2} \mathbb{E} \left[\int_0^T \|z(t)\|_2^2 dt \right].$$

If $\mathbb{E}[\int_0^T \|z(t)\|_2^2 dt] = \mathcal{O}(T)$ and $\|\bar{\mathbf{S}}_Z(T)^{-1}\|_2 = \mathcal{O}_{\mathbb{P}}(1)$, then (13) and (14) imply

$$(15) \quad \|\hat{\mathbf{P}}_\lambda(T) - \mathbf{P}_\lambda\|_2 = \mathcal{O}_{\mathbb{P}}(T^{-1/2}).$$

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

$$(16) \quad \|\hat{\mathbf{P}}_\lambda(T) - \mathbf{P}_\lambda\|_2 \leq \frac{\|\beta\|_2}{\sqrt{T} \sigma_{\min}(\bar{\mathbf{S}}_Z(T))} \left(\sqrt{q} + \sqrt{n+m} + \sqrt{2 \log(1/\delta)} \right).$$

Proof. See Appendix A.2. □

Moreover, this implies the uniform bound $\sup_{\lambda \in \mathbb{C}} \|\hat{\mathbf{P}}_\lambda(T) - \mathbf{P}_\lambda\|_2$ since the right-hand side of (13) does not depend on λ . Finally, by Weyl's inequality,

$$|\sigma_{\min}(\hat{\mathbf{P}}_\lambda(T)) - \sigma_{\min}(\mathbf{P}_\lambda)| \leq \|\hat{\mathbf{P}}_\lambda(T) - \mathbf{P}_\lambda\|_2,$$

which provides a simple statistical analogue of the deterministic Hautus margin. In particular, Proposition 2.2 yields a corresponding $(1 - \delta)$ bound on the singular-value deviation.

2.3. A Fourier-domain approximation without derivatives. The cross-moment $\mathbf{H}_\lambda(T)$ in (9) 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

$$\widehat{dy}_\lambda(i\omega) := \int_0^T e^{-i\omega t} dy_\lambda(t) = \int_0^T e^{-i\omega t} dx(t) - \lambda \widehat{x}(i\omega), \quad \widehat{x}(i\omega) := \int_0^T x(t) e^{-i\omega t} dt.$$

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

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

hence

$$(17) \quad \widehat{dy}_\lambda(i\omega) = x(T)e^{-i\omega T} - x(0) + (i\omega - \lambda)\widehat{x}(i\omega).$$

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

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

In the Itô setting, the unwind energy $\int_{\mathbb{R}} \|\widehat{dy}_\lambda(i\omega)\|_2^2 d\omega$ is not finite (it corresponds to the L^2 -energy of a derivative that does not exist), so frequency-domain computations should be interpreted with an explicit cutoff/windowing when needed.

2.4. 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 $x : [0, T] \rightarrow \mathbb{R}^n$ be absolutely continuous with $x, \dot{x} \in L^2(0, T; \mathbb{R}^n)$ and define operators

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

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, $\text{rank}(\mathcal{P}(\lambda)) := \dim(\text{range}(\mathcal{P}(\lambda))) \leq n$, and *full rank* means surjectivity onto \mathbb{C}^n . For $\varphi \in L^2(0, T)$ and $w \in \mathbb{C}^n$,

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

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

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

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

$$\mathcal{P}(\lambda) \mathcal{P}(\lambda)^* w = \mathcal{P}(\lambda) (\dot{\mathcal{X}}^* w - \lambda \mathcal{X}^* w) = \int_0^T m_\lambda(t) m_\lambda(t)^* w dt =: \mathbf{G}_\lambda(u) w.$$

Hence,

$$(18) \quad \mathcal{P}(\lambda) \mathcal{P}(\lambda)^* = \mathbf{G}_\lambda(u) \quad \text{as operators } \mathbb{C}^n \rightarrow \mathbb{C}^n.$$

Since $\mathcal{P}(\lambda) : L^2(0, T) \rightarrow \mathbb{C}^n$ has finite-dimensional codomain, it is surjective if and only if $\mathcal{P}(\lambda) \mathcal{P}(\lambda)^*$ is invertible on \mathbb{C}^n . By (18), for every $\lambda \in \mathbb{C}$,

$$(19) \quad \text{rank}(\mathcal{P}(\lambda)) = n \quad \Leftrightarrow \quad \mathbf{G}_\lambda(u) \text{ is invertible.}$$

When $\mathbf{S}_Z(u) \succ 0$, the cross-moment representation (11) further gives

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

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

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

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

$$\mathbf{K} := (\mathcal{X} \mathcal{X}^*)^{-1} \mathcal{X} \dot{\mathcal{X}}^* = \left(\int_0^T x(t) x(t)^* dt \right)^{-1} \left(\int_0^T x(t) \dot{x}(t)^* dt \right) \in \mathbb{C}^{n \times n}.$$

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

$$\text{rank}(\mathcal{P}(\lambda)) < n \implies \lambda \in \sigma(\mathbf{K}).$$

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

Proof. See Appendix A.3. □

Corollary 2.4 (Finite checking for the data-driven test). *Assume $\text{rank}(\mathcal{X}) = n$ and $\mathbf{S}_Z(u) \succ 0$, and let \mathbf{K} be as in Theorem 2.3. Define the candidate set $\Lambda := \sigma(\mathbf{K})$, hence $|\Lambda| \leq n$. If*

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

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

Proof. See Appendix A.4. □

Lemma 2.5 (A quantitative lower bound for $\sigma_{\min}(\mathbf{H}_\lambda(T))$ via \mathbf{K}). *Assume $\text{rank}(\mathcal{X}) = n$, $\mathbf{S}_Z(u) \succ 0$, and let \mathbf{K} be as in Theorem 2.3. Then for every $\lambda \in \mathbb{C}$,*

$$\sigma_{\min}(\mathbf{H}_\lambda(T)) \geq \sqrt{\sigma_{\min}(\mathbf{S}_Z(u))} \frac{\sigma_{\min}(\mathbf{K} - \bar{\lambda} \mathbf{I})}{\|(\mathcal{X} \mathcal{X}^*)^{-1} \mathcal{X}\|}.$$

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

Proof. See Appendix A.5. □

Consequently, for rank certification it is enough to check the finitely many candidates $\Lambda = \sigma(\mathbf{K})$.

2.5. Best conditioning of control inputs. The data-driven estimate (12) and the resulting singular-value margin are controlled—up to the model-dependent factor $\sigma_{\min}(\mathbf{P}_\lambda)$ —by the smallest singular value of the stacked Gramian $\mathbf{S}_Z(u)$. Since $\mathbf{S}_Z(u)$ 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 $\mathbf{S}_Z(u)$ in (11) is numerically stable. Under an L^2 energy budget, the best possible conditioning corresponds to spreading the energy isotropically across the m input channels.

Remark 2.6. The matrix $\mathbf{S}_U(u)$ is a principal submatrix of $\mathbf{S}_Z(u)$, hence

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

Thus, even though $\sigma_{\min}(\mathbf{S}_Z(u))$ 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 $\sigma_{\min}(\mathbf{S}_Z(u))$ without model knowledge.

Proposition 2.7 (Best conditioning under $\|u\|_{L^2} \leq 1$). Define the input Gramian

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

If $\|u\|_{L^2(0,T)}^2 \leq 1$, then

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

Moreover, there exist u with $\|u\|_{L^2(0,T)} \leq 1$ such that $\mathbf{S}_U(u) = \frac{1}{m}\mathbf{I}_m$. In particular, the upper bound is tight.

Proof. See Appendix A.6. □

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

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

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

Proposition 2.8 (Best conditioning under $\|u\|_{H^1} \leq 1$). Let $u \in H^1(0, T; \mathbb{R}^m)$ satisfy $\|u\|_{H^1(0,T)} \leq 1$, and define $\mathbf{S}_U(u) := \int_0^T u(t)u(t)^\top dt$. Then

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

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

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

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

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

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

Proof. See Appendix A.7. □

Remark 2.9. As $T \rightarrow \infty$ the derivative penalty vanishes and $\alpha \rightarrow 1/m$, recovering the L^2 -budget optimum in Proposition 2.7. For short horizons (or large m), the optimal H^1 -budget design suppresses high-frequency components and yields a smaller isotropic eigenvalue α .

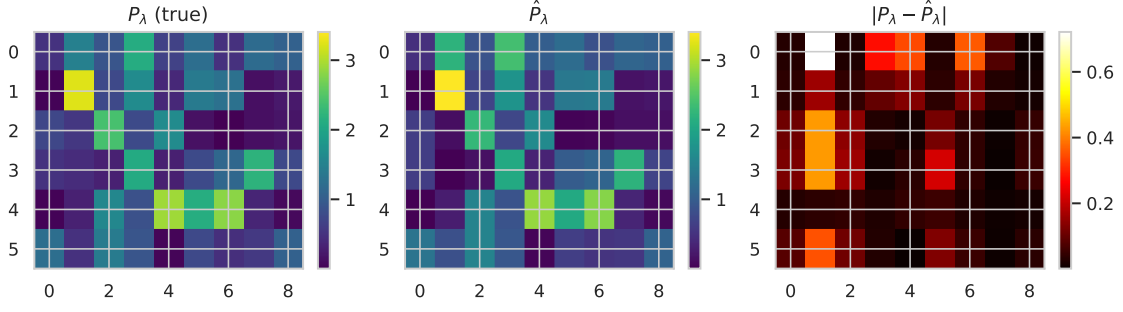


FIGURE 1. Element-wise comparison of the true Hautus matrix \mathbf{P}_λ (left) and its time-domain estimate $\hat{\mathbf{P}}_\lambda$ (center), with absolute difference (right). The small difference magnitudes confirm the accuracy of the cross-moment estimator.

3. 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

$$dx(t) = (\mathbf{A}x(t) + \mathbf{B}u(t)) dt + \beta dW(t),$$

where $\mathbf{A} \in \mathbb{R}^{n \times n}$ is the drift matrix, $\mathbf{B} \in \mathbb{R}^{n \times m}$ the input matrix, $\beta \in \mathbb{R}^{n \times q}$ 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

$$u_i(t) = \sin(\omega_i t), \quad i = 1, \dots, m,$$

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

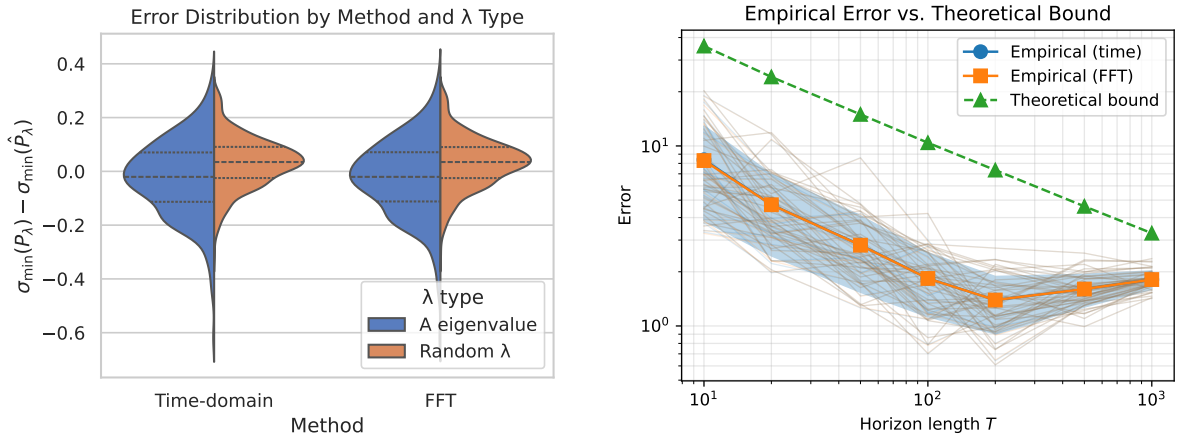
3.1. Validation of the Data-Driven Hautus Test. Our first experiment validates the data-driven Hautus test by comparing the estimated matrix $\hat{\mathbf{P}}_\lambda$ with the true Hautus matrix $\mathbf{P}_\lambda = [\mathbf{A} - \lambda \mathbf{I}, \mathbf{B}]$. We simulate a system with $n = 6$ states, $m = 3$ inputs, and $q = 2$ noise dimensions over a horizon $T = 100$ with time step $dt = 0.05$. 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 (12).

Figure 1 shows the element-wise comparison between the true and estimated matrices for a representative $\lambda = 0.3 + 1.3i$. Both methods yield accurate estimates, with estimation errors $\|\hat{\mathbf{P}}_\lambda - \mathbf{P}_\lambda\|_2$ 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 $\sigma_{\min}(\mathbf{P}_\lambda) - \sigma_{\min}(\hat{\mathbf{P}}_\lambda)$, showing near-zero bias and comparable variance for time-domain and FFT estimators.

3.2. Error Convergence Rate. Proposition 2.2 predicts that the estimation error $\|\hat{\mathbf{P}}_\lambda - \mathbf{P}_\lambda\|_2$ 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



(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).

standard deviation across trials. We also overlay the theoretical bound from Proposition 2.2 (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.

3.3. 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 2.4 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

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

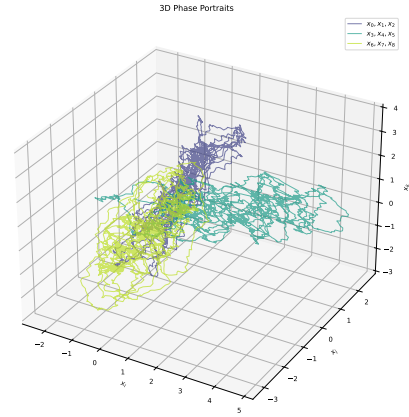
We simulate a controllable system with $n = 6$, $m = 3$, $q = 2$ over a horizon $T = 100$. The candidate eigenvalues $\sigma(\mathbf{K})$ are computed from the trajectory data and compared with the true eigenvalues $\sigma(\mathbf{A})$. For each candidate λ_i , we compute the estimated minimum singular value $\sigma_{\min}(\hat{\mathbf{P}}_{\lambda_i})$ 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 $\sigma_{\min}(\hat{\mathbf{P}}_\lambda) > 0$ for all candidates, the system is correctly identified as controllable.

3.4. Comparison of Time-Domain and FFT Methods. We compare two computational approaches for evaluating the cross-moments: (i) the *time-domain method*, which computes $\mathbf{H}_\lambda(T)$ via direct numerical integration, and (ii) the *FFT method*, which leverages the Parseval identity with frequency-domain representations as in (17).

For a system with $n = 8$ states, $m = 4$ inputs, $q = 2$ noise dimensions, and horizon $T = 100$ with time step $dt = 0.02$ (yielding 5001 sample points), we test both methods

ω_{\max}	$\sigma_{\min}(\hat{\mathbf{P}}_\lambda)$	$\ \hat{\mathbf{P}}_\lambda - \mathbf{P}_\lambda\ _2$
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 ω_{\max} on FFT method accuracy for $\lambda = 0.3 + 0.8i$.

(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.

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 ω_{\max} to truncate the integral. Figure 3a shows the effect of ω_{\max} on the estimation error for $\lambda = 0.3 + 0.8i$. 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 offers modest speedups for long trajectories due to the $\mathcal{O}(N \log N)$ complexity of the FFT versus $\mathcal{O}(N)$ for direct integration, though both are fast in practice.

3.5. 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 $\mathbf{S}_Z(u)$ is well-conditioned, which is crucial for accurate estimation of the Hautus matrix and reliable controllability certification.

REFERENCES

- [1] M.K. Camlibel et al. “The Shortest Experiment for Linear System Identification”. In: *Systems & Control Letters* 197 (Mar. 2025), p. 106045. ISSN: 01676911. DOI: [10.1016/j.sysconle.2025.106045](https://doi.org/10.1016/j.sysconle.2025.106045). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0167691125000271> (visited on 01/20/2026).
- [2] 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.

- [3] Dennis Gramlich et al. *Fast Identification and Stabilization of Unknown Linear Systems*. arXiv.org. Aug. 22, 2022. URL: <https://arxiv.org/abs/2208.10392v3> (visited on 12/11/2025).
- [4] 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](https://doi.org/10.1109/CDC45484.2021.9683709). URL: <https://ieeexplore.ieee.org/document/9683709> (visited on 12/15/2025).
- [5] 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](https://doi.org/10.1137/0301010). URL: <http://epubs.siam.org/doi/10.1137/0301010> (visited on 12/08/2025).
- [6] 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](https://doi.org/10.1109/LCSYS.2020.3003770). URL: <https://ieeexplore.ieee.org/document/9121339/> (visited on 12/11/2025).
- [7] 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](https://doi.org/10.1016/j.sysconle.2023.105603). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0167691123001500> (visited on 01/20/2026).
- [8] 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](https://doi.org/10.1109/LCSYS.2022.3205550). URL: <https://ieeexplore.ieee.org/document/9882335/> (visited on 01/20/2026).
- [9] 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](https://doi.org/10.1109/LCSYS.2025.3576276). arXiv: [2503.12489](https://arxiv.org/abs/2503.12489) [math]. URL: <http://arxiv.org/abs/2503.12489> (visited on 12/08/2025).
- [10] 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).
- [11] Henk J. Van Waarde. “Beyond Persistent Excitation: Online Experiment Design for Data-Driven Modeling and Control”. In: *IEEE Control Systems Letters* 6 (2022), pp. 319–324. ISSN: 2475-1456. DOI: [10.1109/LCSYS.2021.3073860](https://doi.org/10.1109/LCSYS.2021.3073860). URL: <https://ieeexplore.ieee.org/document/9406124/> (visited on 01/20/2026).
- [12] 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](https://doi.org/10.1109/TAC.2020.2966717). URL: <https://ieeexplore.ieee.org/document/8960476/> (visited on 12/11/2025).
- [13] Roman Vershynin. “An Introduction with Applications in Data Science”. In: ().
- [14] 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](https://doi.org/10.48550/arXiv.2002.01023). arXiv: [2002.01023](https://arxiv.org/abs/2002.01023) [math]. URL: <http://arxiv.org/abs/2002.01023> (visited on 12/08/2025). Pre-published.
- [15] 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.](https://doi.org/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 (7), this gives $v^* \mathbf{H}_\lambda(u) = v^* \mathbf{P}_\lambda \mathbf{S}_Z(u) = 0$, hence $\text{rank}(\mathbf{H}_\lambda(u)) < n$. This proves the contrapositive of the first implication in Theorem 2.1.

Conversely, assume (\mathbf{A}, \mathbf{B}) is controllable and $\mathbf{S}_Z(u)$ is invertible. Then \mathbf{P}_λ has full row rank for every $\lambda \in \mathbb{C}$ and (7) implies

$$\text{rank}(\mathbf{H}_\lambda(u)) = \text{rank}(\mathbf{P}_\lambda \mathbf{S}_Z(u)) = \text{rank}(\mathbf{P}_\lambda) = n \quad \forall \lambda \in \mathbb{C},$$

which proves the converse. \square

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

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

$$\mathbb{E}[\mathbf{M}(T)_{i,:}^\top \mathbf{M}(T)_{i,:} \mid z] = \int_0^T z(t) z(t)^\top dt = \mathbf{S}_Z(u), \quad i = 1, \dots, q.$$

Therefore, conditional on z , we have the distributional identity $\mathbf{M}(T) \stackrel{d}{=} \mathbf{G} \mathbf{S}_Z(u)^{1/2}$ where $\mathbf{G} \in \mathbb{R}^{q \times (n+m)}$ has i.i.d. $\mathcal{N}(0, 1)$ entries. Using (13),

$$\hat{\mathbf{P}}_\lambda(T) - \mathbf{P}_\lambda \stackrel{d}{=} \beta \mathbf{G} \mathbf{S}_Z(u)^{-1/2},$$

conditional on z , hence

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

The standard Gaussian matrix bound $\mathbb{P}(\|\mathbf{G}\|_2 \geq \sqrt{q} + \sqrt{n+m} + t) \leq e^{-t^2/2}$ (valid for all $t \geq 0$; see, e.g., [13]) yields (16) by choosing $t = \sqrt{2 \log(1/\delta)}$ and using $\sigma_{\min}(\mathbf{S}_Z(u)) = T \sigma_{\min}(\bar{\mathbf{S}}_Z(T))$. \square

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

Proof. Fix $\lambda \in \mathbb{C}$ and assume $\text{rank}(\mathcal{P}(\lambda)) < n$. 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. $\dot{\mathcal{X}}^* w = \bar{\lambda} \mathcal{X}^* w$. Left-multiply by $(\mathcal{X} \mathcal{X}^*)^{-1} \mathcal{X}$ to obtain

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

so $\bar{\lambda} \in \sigma(\mathbf{K})$. Since x is real-valued, the moments defining \mathbf{K} are real and hence $\sigma(\mathbf{K})$ is closed under complex conjugation, which yields $\lambda \in \sigma(\mathbf{K})$. \square

A.4. Proof of Corollary 2.4.

Proof. Suppose there existed $\lambda_0 \in \mathbb{C}$ with $\text{rank}(\mathbf{H}_{\lambda_0}(T)) < n$. Since $\mathbf{S}_Z(u) \succ 0$, the matrix

$$\mathbf{G}_{\lambda_0}(u) = \mathbf{H}_{\lambda_0}(T) \mathbf{S}_Z(u)^{-1} \mathbf{H}_{\lambda_0}(T)^*$$

is singular, hence (19) implies $\text{rank}(\mathcal{P}(\lambda_0)) < n$. Theorem 2.3 then yields $\lambda_0 \in \sigma(\mathbf{K}) = \Lambda$, contradicting the hypothesis. \square

A.5. 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

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

Moreover,

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

hence

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

Using (18), $\|(\dot{\mathcal{X}}^* - \bar{\lambda} \mathcal{X}^*)w\|_{L^2(0, T)}^2 = w^* \mathbf{G}_\lambda(u) w$, so

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

Since $\mathbf{S}_Z(u) \succ 0$, (11) gives

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

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

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

i.e.,

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

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

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

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

A.6. Proof of the L^2 -budget isotropic design.

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

$$\begin{aligned} \text{tr}(\mathbf{S}_U(u)) &= \int_0^T \text{tr}(u(t)u(t)^\top) dt \\ &= \int_0^T \|u(t)\|_2^2 dt \\ &= \|u\|_{L^2(0, T)}^2 \leq 1, \end{aligned}$$

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 $\mathbf{S}_U(u) = \frac{1}{m} \mathbf{I}_m$. \square

A.7. Proof of the H^1 -budget isotropic design.

Proof. Expand u in the orthonormal Neumann basis $\{\psi_k\}_{k \geq 0}$ as $u(t) = \sum_{k \geq 0} \psi_k(t) a_k$ with coefficients $a_k \in \mathbb{R}^m$, so that

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

Let $W := \text{diag}(w_0, w_1, \dots)$ with $w_k := 1 + \left(\frac{k\pi}{T}\right)^2$, and define $A := [a_0 \ a_1 \ \dots]$, so that $\mathbf{S}_U(u) = AA^\top$ and $\|u\|_{H^1(0,T)}^2 = \text{tr}(AWA^\top)$. Write $A = \mathbf{S}_U(u)^{1/2} R$ where R has orthonormal rows ($RR^\top = \mathbf{I}_m$), and set $\Pi := R^\top R$, a rank- m orthogonal projector. Then

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

Since W is diagonal with nondecreasing entries, the minimum of $\text{tr}(W \Pi)$ over rank- m projectors Π equals $\sum_{k=0}^{m-1} w_k$, attained by projecting onto $\text{span}\{\psi_0, \dots, \psi_{m-1}\}$. Using $\|u\|_{H^1(0,T)}^2 \leq 1$ gives the stated upper bound on $\lambda_{\min}(\mathbf{S}_U(u))$. The construction with the first m Neumann modes and orthogonal Q yields $\mathbf{S}_U(u) = \alpha \mathbf{I}_m$ and saturates $\|u\|_{H^1(0,T)}^2 = 1$. \square