Choutri, Salah Eddine; Djehiche, Boualem; Mazhar, Othmane

Article

Empirical validation of novel non-asymptotic bounds on the least squares estimator for LTI systems with applications in economics

Asian Journal of Economics and Banking (AJEB)

Provided in Cooperation with:
Ho Chi Minh University of Banking (HUB), Ho Chi Minh City

Suggested Citation: Choutri, Salah Eddine; Djehiche, Boualem; Mazhar, Othmane (2025) : Empirical validation of novel non-asymptotic bounds on the least squares estimator for LTI systems with applications in economics, Asian Journal of Economics and Banking (AJEB), ISSN 2633-7991, Emerald, Leeds, Vol. 9, Iss. 3, pp. 491-509, https://doi.org/10.1108/AJEB-12-2024-0147

This Version is available at:
https://hdl.handle.net/10419/334156

Asian Journal of Economics and Banking

491

Empirical validation of novel non-asymptotic bounds on the least squares estimator for LTI systems with applications in economics

Salah Eddine Choutri
Research Center for Interacting Urban Networks (CITIES) and Research Center for Stability, Instability and Turbulence (SITE), New York University Abu Dhabi, Abu Dhabi, United Arab Emirates
Boualem Djehiche
Department of Mathematics, School of Engineering Sciences, KTH Royal Institute of Technology, Stockholm, Sweden, and
Othmane Mazhar
Laboratoire de Probabilités, Statistique et Modélisation, Université Paris Cité, Paris, France and Sorbonne Université, Paris, France

Received 26 December 2024
Revised 20 January 2025
Accepted 3 February 2025

Abstract

Purpose – This study aims to investigate the performance of new theoretical non-asymptotic bounds of the least squares estimator for linear time-invariant (LTI) state-space models, building on prior foundational research.

Design/methodology/approach – Using both simulated and real-world datasets, the analysis examines the influence of noise characteristics, controllability and noise conditioning on estimation accuracy. Furthermore, the study validates the derived bounds on real datasets through, among others, cross-validation and bootstrap techniques.

Findings – The results confirm the validity of the theoretical non-asymptotic error bounds across various system configurations. Key insights highlight the role of system stability, dimensionality and noise properties in shaping estimation performance.

Research limitations/implications – While this study focuses on specific configurations and datasets, expanding the scope to include alternative estimation techniques and broader real-world systems could yield additional insights into the practical applicability of these bounds.

Practical implications – The findings provide actionable insights and practical guidelines for improving econometric modeling, financial time-series forecasting, adaptive systems control and environmental modeling, enabling improved system identification and parameter estimation techniques in these domains.

Originality/value – By using real datasets and advanced validation techniques, this study bridges the gap between theoretical bounds and practical applications, hopefully opening a novel perspective on least squares estimation performance in LTI systems.

Keywords Time series, Air pollution, Non-asymptotic bounds

Paper type Research paper

JEL Classification — B23 History of Economic Thought: Quantitative and Mathematical, C01 Econometrics, C02 Mathematical Methods, C15 Statistical Simulation Methods: General

This work was supported in part by the NYUAD Center for Interacting Urban Networks (CITIES), funded by Tamkeen under the NYUAD Research Institute Award CG001 and in part by the NYUAD Research Center on Stability, Instability and Turbulence (SITE), funded by Tamkeen under the NYUAD Research Institute Award CG002.

Asian Journal of Economics and Banking
Vol. 9 No. 3, 2025
pp. 491-509
Emerald Publishing Limited
e-ISSN: 2633-7991
p-ISSN: 2615-9821
DOI 10.1108/AJEB-12-2024-0147

1. Introduction

The estimation of Linear time-variant (LTI) systems is a fundamental problem in statistics, with applications spanning diverse fields such as system identification, finance, biology, and various engineering and control domains (Ljung, 1986).

The performance of least squares estimators is particularly relevant in economics and finance, where accurate modeling of time-series data and state-space systems is essential for tasks such as financial forecasting, portfolio optimization, and macroeconomic policy modeling. This study bridges theoretical bounds and practical applications in these fields.

Historically, research on this topic has predominantly focused on the asymptotic convergence properties of specific estimation techniques, such as ordinary least squares (Ljung, 1976a), prediction error methods (Ljung, 1976b), and maximum likelihood. However, the finite-time behavior of estimation schemes has received comparatively less attention.

Recently, analyzing the finite-time properties of these estimation methods for LTI systems became a crucial area of interest because of new use cases and challenges in machine learning. For instance, these error rates dictate the performance of model-based reinforcement learning algorithms that rely on a linear quadratic regulator (Arora *et al*., 2018). These studies primarily investigate the finite-time properties of classical estimation methods, including ordinary least squares, for both observable (Simchowitz *et al*., 2018) and unobservable states (Oymak and Ozay, 2019; Djehiche and Mazhar, 2022). The findings often leverage advanced tools from statistical learning theory, random matrix theory, and self-normalizing processes.

In addition, some recent works have provided information-theoretic perspectives, offering lower bounds on the sample complexity required to achieve a specified estimation accuracy (Jedra and Proutière, 2019; Djehiche and Mazhar, 2021; Djehiche *et al*., 2025). These lower bounds quantify the minimal number of observations needed to estimate the system's parameters.

Among the most comprehensive contributions to this field is (Djehiche *et al*., 2025), which establishes both upper and lower bounds for the least squares estimator in LTI systems. These bounds are nearly tight, differing only by an unknown positive constant and parameters reflecting the estimation task's complexity, such as system stability, observability, and noise intensity.

The current work builds on these theoretical foundations and aims to validate the derived error bounds in practical settings using both simulated and real-world data. Specifically, this study has three primary objectives:

(1) Validate theoretical error bounds for least squares estimation across various system configurations.

(2) Investigate the influence of factors such as system stability, noise properties, and system dimensionality on estimation performance.

(3) Demonstrate the practical applicability of the results through experiments on synthetic and real-world datasets.

Unlike simulated data, where the system matrices are generated with controlled properties, real-world datasets present additional challenges such as unknown dynamics, and data variability. Preprocessing steps, including normalization and ensuring compatibility with the LTI modeling framework, were applied to prepare the datasets for analysis. The two datasets are

- **Beijing Multi-Site Air-Quality Dataset:** A dataset from the UCI Machine Learning Repository containing hourly air pollutant measurements for evaluating model robustness in environmental monitoring.
- **Dynamical System Multivariate Time Series Forecast Dataset:** A synthetic dataset from Kaggle designed to mimic real-world dynamical systems for testing estimation

methods. This dataset mirrors the dynamics observed in financial markets, where multivariate time series are used to model interactions between economic indicators or asset prices.

Validation on these datasets involves estimating the system matrices and comparing the observed estimation errors with the theoretical bounds derived from the controllability and noise structure. This real-world analysis complements the simulated experiments, highlighting the bounds' applicability across diverse scenarios. The chosen datasets and simulation setups are representative of challenges faced in financial econometrics and macroeconomic modeling, where noise, stability, and high-dimensional data are common obstacles.

The structure of the paper is as follows. After this introduction and a brief outline of commonly used notations, we introduce the LTI system estimation problem in Section 2. This is followed by a review of the classical least squares solution, the standard assumptions for deriving error estimates, and the key results from (Djehiche *et al.*, 2025) that provide the primary error bounds, discussed in Section 1.

Section 3 outlines the methodology for validating these theoretical findings. It details the randomized generation process for LTI system parameters used in numerical experiments, the simulation procedure for generating system trajectories, the estimation method for LTI parameters, the validation tests conducted on simulated data, and the datasets employed to assess the practicality and applicability of the theoretical bounds in real-world conditions.

Section 4 presents an analysis of the simulation results, highlighting the degree to which the theoretical upper and lower error bounds predict the least squares estimator's performance and examining the complexity parameters' influence on estimation errors. Section 5 demonstrates the practical effectiveness of the least squares estimator in real-world data and evaluates the accuracy of the theoretical bounds using two datasets: one for the task of time series forecasting and another for the environmental task of air quality estimation in Beijing.

Finally, Section 6 discusses the key insights from this empirical study, and concludes the paper.

1.1 Frequently used notation

Throughout this work, we use the following notation:

- We denote by $W := \mathrm{diag}(w_1, \ldots, w_d)$ a matrix in $\mathcal{M}_{d\times d}(\mathbb{C})$ whose only nonzero entries are on the diagonal, with $W_{ii} = w_i$. A real matrix $W \in \mathcal{M}_{d\times d}(\mathbb{R})$ admits a singular value decomposition (SVD) $W = U\Sigma V$, where $U, V \in O(d)$ (the group of orthogonal matrices) and $\Sigma = \mathrm{diag}(s_1(W), \ldots, s_d(W))$ contains the singular values, satisfying $0 \leq s_1(W) \leq s_2(W) \leq \cdots \leq s_d(W)$. We define $s_{\max}(W) = s_d(W)$, $s_{\min}(W) = s_1(W)$, and the condition number $cond(W) = \frac{s_{\max}(W)}{s_{\min}(W)}$.
- A matrix $W \in \mathcal{M}_{d\times d}(\mathbb{R})$ is diagonalizable if it can be expressed as $W = S\Gamma S^{-1}$, where S is an invertible matrix (the change of basis matrix) and $\Gamma = \mathrm{diag}(\lambda_1, \ldots, \lambda_d)$ is the diagonal matrix of eigenvalues. Examples of diagonalizable matrices include symmetric and positive definite matrices. For a positive definite matrix, the decomposition $W = S\Gamma S^{-1}$ is referred to as its spectral decomposition, where all eigenvalues λ_i are positive. We denote the largest eigenvalue of a diagonalizable matrix W (in magnitude) by $\lambda_{\max}(W)$ and the smallest by $\lambda_{\min}(W)$, and recall the inequalities $s_{\max}(W) \geq |\lambda_{\max}(W)| \geq |\lambda_{\min}(W)| \geq s_{\min}(W)$.
- We denote by $(\Omega, \mathcal{F})$ the underlying measurable space. A probability measure defined on it is denoted by $\mathbb{P}$, and $\mathbb{E}$ represents the corresponding expectation operator.
- All norms are indicated with a subscript identifying the corresponding normed space, except for the absolute value $|\cdot|$. The p-norms of a vector $x \in \mathbb{R}^n$ are given by

$$|x|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \quad \text{for } 1 \le p < \infty, \qquad |x|_\infty = \max_{i\in[[1,n]]} |x_i|.$$

The p-Schatten norm of a matrix W in the Schatten space S_p is defined as

$$|W|_{S_p} = \left(\mathrm{tr}\left(W^* W\right)^{p/2}\right)^{1/p}, \quad \text{for } 1 \le p < \infty, \qquad |W|_{S_\infty} = \max_{x\in\mathbb{S}_2^{n-1}} |Wx|_2,$$

where $\mathbb{S}_2^{n-1} = \{x \in \mathbb{R}^n \mid |x|_2 \le 1\}$ denotes the Euclidean unit sphere in $\mathbb{R}^n$. We also have $|W|_{S_\infty} = s_{\max}(W)$.

The spaces $L_p(\Omega, \mathcal{F}, \mathbb{P}) = \{x : \Omega \to \mathbb{R} \text{ measurable}, \mathbb{E}(|x|^p) < \infty\}$ of random variables are equipped with the L_p-norms:

$$|x|_{L_p} = (\mathbb{E}(|x|^p))^{1/p}, \quad \text{for } 1 \le p < \infty, \qquad |x|_{L_\infty} = \underset{\omega\in\Omega}{ess\,sup}|x(\omega)|,$$

where *ess* sup denotes the essential supremum with respect to the measure $\mathbb{P}$.

- Throughout this work, C denotes positive constants whose exact values are immaterial to the results. The notation $x \lesssim y$ (resp. $x \gtrsim y$) means a positive constant C exists such as $x \le Cy$ (resp. $x \ge Cy$). The minimum (maximum) of two real numbers x and y is denoted as $\min(x, y) = x \wedge y$ ($\max(x, y) = x \vee y$).
- The notation $x \simeq y$ indicates that x and y are comparable up to constant factors, meaning there exist positive constants c and C such that $cy \le x \le Cy$.

2. The theoretical non-asymptotic bounds

We consider the LTI state-space model:

$$x_{i+1} = Ax_i + B\varepsilon_i, \quad i \in [[1, N]], \tag{1}$$

where $A \in \mathcal{M}_{d\times d}(\mathbb{R})$ is a parameter matrix and $B\varepsilon_i$ represents Gaussian noise. Here, $B \in \mathcal{M}_{d\times r}(\mathbb{R})$ is an unknown, rectangular, full-rank matrix with $r \le d$, and the noise sequence $(\varepsilon_i)_i$ is *i.i.d.* multivariate normal $\mathcal{N}(0, I_r)$. For simplicity, we assume $x_0 = 0$. Expanding (1), we derive the following expression for the observed covariates for $i \in [[1, N]]$:

$$x_i = \sum_{k=0}^{i-1} A^{i-1-k} B\varepsilon_k. \tag{2}$$

In the following definition, we introduce the concept of a controllable system.

Definition 2.1. *The system* $[A, B]$ *with* $A \in \mathcal{M}_{d\times d}(\mathbb{C})$ *and* $B \in \mathcal{M}_{d\times r}(\mathbb{C})$ *is said to be controllable if there exists an integer* $\overline{d}$ *such that the columns of the* $\overline{d}$*-order controllability matrix*

$$C(A, B) = \left[B\ AB\ A^2B\ \ldots\ A^{\overline{d}}B\right]$$

span $\mathbb{R}^d$. *We also define the controllability order* $\check{d}$ *as the smallest such* $\overline{d}$.

The least-squares estimator for the parameter matrix A is given by:

$$\widehat{A} := \underset{A \in \mathcal{M}_{d\times d}(\mathbb{R})}{\operatorname{argmin}} \sum_{t=1}^{N} \|x_{t+1} - Ax_t\|_2^2 = \left(\sum_{t=1}^{N} x_{t+1}x_t^*\right)\left(\sum_{t=1}^{N} x_t x_t^*\right)^{-1}, \tag{3}$$

where M^* denotes the conjugate transpose of the matrix M, and $\|y\|_2^2 := \sum_{i=1}^{n} |y_i|^2$. In Lemma 2.1 of (Djehiche *et al.*, 2025) it is shown that the sample covariance matrix for the model (1) is invertible with probability one if the pair $[A, B]$ is controllable and $N \gtrsim d$.

The behavior of the least square estimator for estimating the LTI dynamics governed by $x_{t+1} = Ax_t + B\varepsilon_t$ is strongly influenced by the stability properties of the matrix A. Stability classification depends on the eigenvalues of A, with the following cases:

- *Stable systems*: $|\lambda_i| < 1$ for all eigenvalues λ_i of A.
- *Unstable systems*: $|\lambda_i| > 1$ for any eigenvalue λ_i of A.
- *Limit-stable systems*: $|\lambda_i| = 1$ for some eigenvalues λ_i of A.

Theoretical bounds on the least-squares estimator $\widehat{A}$ offer valuable insights into its estimation accuracy. These bounds are also influenced by the dimension of the matrix A and the controllability properties of the pair $[A, B]$. Specifically, for the least-squares estimator $\widehat{A}$ of the auto-regression matrix A, the following results hold for the estimation loss $\mathcal{L}_p(\widehat{A}, A)$ for $p \in [1, \infty]$ and the mean square estimation risk $\mathcal{E}_2(\widehat{A}, A)$ defined as follows:

$$\mathcal{L}_p(\widehat{A}, A) = |\widehat{A} - A|_{S_p} \quad \text{for} \quad p \in [1, \infty], \quad \mathcal{E}_2(\widehat{A}, A) = \left(\mathbb{E}\left(|\widehat{A} - A|_{S_2}^2\right)\right)^{1/2}.$$

Below, we summarize the bounds for the estimation loss and mean square risk, depending on the stability of the matrix A, while highlighting the contribution of the complexity parameters of the pair $[A, B]$. By complexity parameters, we refer to the dimension of the matrix A, the controllability order $\check{d}$ for the system $[A, B]$, the noise properties (i.e. the strength $|B|_{S_\infty}$ and the conditioning of B), and the controllability properties (i.e. the locations of the singular values and the condition number of the controllability $C(A, B)$). The following upper bounds are derived from Theorems 2.4 and 2.5 of (Djehiche *et al.*, 2025), while the lower bounds are based on Theorem 3.2 and Corollary 4.1 of the same work:

2.1 Stable matrix

If A consists solely of a stable part, then for $\delta \in (0, 1)$ and sample size N satisfying:

$$N \gtrsim \frac{\check{d}}{1 - s_{\min}(A)^{2\check{d}}} \vee \frac{(d \vee \log(1/\delta))|B|_{S_\infty}^2}{\left(1 - |A|_{S_\infty}\right)^2 s_{\min}^2(C(A, B))}, \tag{4}$$

with probability at least $1 - \delta$, the least squares estimator $\widehat{A}$ satisfies for the estimation loss:

$$\mathcal{L}_p(\widehat{A}, A) \lesssim \frac{|B|_{S_\infty}}{s_{\min}(S)} \frac{d^{1/2+1/p} + d^{1/p}\log^{1/2}\left(\frac{1}{\delta}\right)}{N^{1/2} s_{\min}(C(A, B))}. \tag{5}$$

the following lower bound for the mean square estimation risk holds:

$$\mathcal{E}_2(\widehat{A},A)\gtrsim\frac{|B|_{S_2}}{s_{\max}(S)}\,\frac{d^{1/2}\left(1-s_{\max}^{2\check{d}}(A)\right)^{1/2}}{N^{1/2}s_{\max}(C(A,B))}. \tag{6}$$

2.2 Unstable matrix

If A consists solely of an unstable part, then for $\delta\in(0,1)$ and N satisfying:

$$N\gtrsim\frac{\check{d}}{1-s_{\min}\left(A^{-1}\right)^{2\check{d}}}\vee\frac{\log(N/\check{d}\,)}{2\check{d}\log(s_{\min}(A))}\vee\frac{(d\vee\log(1/\delta))|B|_{S_\infty}^2}{\left(1-|A|_{S_\infty}\right)^2 s_{\min}^2\left(C\left(A^{-1},B\right)\right)}, \tag{7}$$

with probability at least $1-\delta$, the least square estimator satisfies the following bound on the estimation loss:

$$\mathcal{L}_p(\widehat{A},A)\lesssim\frac{|B|_{S_\infty}}{s_{\min}(S)}\,\frac{d^{1/2+1/p}+d^{1/p}\log^{1/2}\left(\frac{1}{\delta}\right)}{s_{\min}^{N-3\check{d}}(A)s_{\min}(C(A,B))}. \tag{8}$$

and the following lower bound for the mean square estimation risk is given by:

$$\mathcal{E}_2(\widehat{A},A)\gtrsim\frac{|B|_{S_2}}{s_{\max}(S)}\,\frac{d^{1/2}\left(s_{\max}^{2\check{d}}(A)-1\right)}{s_{\max}^{N+3\check{d}}(A)s_{\max}(C(A,B))}. \tag{9}$$

2.3 Limit-stable matrix

If A consists of a limit-stable part, then for $\delta\in(0,1-e^{-5})$ and

$$N\gtrsim cond^2(C(A,B))\left(1\vee\frac{1}{s_{\min}^2(C(A,B))}\right)d\log\left(\frac{Ncond(C(A,B))|B|_{S_\infty}}{\delta}\right), \tag{10}$$

the least squares estimator satisfies the following bound on the estimation loss:

$$\mathcal{L}_p(\widehat{A},A)\lesssim\frac{d^{1/2}|B|_{S_\infty}}{s_{\min}(S)}\,\frac{d^{1/p}(d+\log\zeta_1+\log\log N+\log(1/\delta))^{1/2}\zeta_2}{Ns_{\min}(C(A,B))}, \tag{11}$$

where

$$\zeta_1=\frac{\check{d}|B|_{S_\infty}}{s_1(C(A,B))},\quad \zeta_2=cond^2(C(A,B))\log^{1/2}\left(cond(C(A,B))\frac{N}{\delta}\right). \tag{12}$$

Furthermore, the following lower bound holds for the mean square estimation risk:

$$\mathcal{E}_2(\widehat{A},A)\gtrsim\frac{|B|_{S_2}}{s_{\max}(S)}\,\frac{\check{d}^{1/2}}{N\left(d^{1/2}\vee\frac{\log\zeta_1}{d^{1/2}}\right)\zeta_1 s_{\max}(C(A,B))}. \tag{13}$$

Empirical validation of these theoretical bounds is explored under varying conditions of B and $C(A, B)$.

3. Methodology

This section outlines the methodology for validating theoretical bounds in linear regression for LTI dynamic systems. The methodology includes the systematic generation of the data from a dynamic systems across stability regimes, evaluation of noise and matrix conditioning effects, and validation using simulated and real-world data.

3.1 Data generation across stability regimes

LTI Dynamic systems are modeled using the state evolution equation:

$$x_{t+1} = Ax_t + B\epsilon_t, \tag{14}$$

where A is the system matrix, B is the control matrix, and ϵ_t is Gaussian noise with zero mean and covariance matrix I_r. To evaluate theoretical bounds under diverse conditions, the eigenvalues of A are constrained to three stability regimes, we avoid mixing eigenvalues from different regimes for the sake of clarity of the results:

- **Stable Systems:** Eigenvalues λ_i of A are sampled uniformly from $U(0.1, 0.9)$, ensuring $|\lambda_i| < 1$. The matrix B is randomly generated and normalized to have a Schatten 2-norm of 1. Well-conditioned B matrices have condition numbers below 10, while ill-conditioned B matrices have condition numbers between 40 and 10,000.
- **Unstable Systems:** Eigenvalues λ_i of A are sampled uniformly from $U(1.05, 1.2)$, ensuring $|\lambda_i| > 1$. The matrix B is normalized similarly to the stable case. Well-conditioned B matrices have condition numbers below 10, and ill-conditioned B matrices have condition numbers between 200 and 500.
- **Limit-Stable Systems:** Eigenvalues lie on the unit circle ($|\lambda_i| = 1$), either constructed as $\lambda_i = e^{i\theta}$, with θ uniformly spaced between 0 and 2π, or rescaled from an initial A matrix. Well-conditioned B matrices have condition numbers below 10, and ill-conditioned B matrices have condition numbers between 10 and 1,000.

The controllability matrix $C(A, B)$ is constructed iteratively to ensure the (A, B) pair satisfies the controllability criterion. The rank and condition number of $C(A, B)$ are verified, and if the pair does not meet the required conditions, the matrices A and B are regenerated or adjusted through scaling, perturbation, or rank modification.

3.2 Simulation framework

For each (A, B) pair, system trajectories are simulated over $N = 1{,}000$ time steps. The state evolution is governed by the equation $x_{t+1} = Ax_t + B\epsilon_t$, with ϵ_t as Gaussian noise scaled by σ. Two noise levels ($\sigma = 0.1, 1.0$) are used to evaluate the impact of noise on estimation accuracy.

Using the least squares algorithm, multiple estimates $\widehat{A}$ are computed for each simulated trajectory. Two key metrics are evaluated: (1) the Mean Squared Error (MSE) risk $\mathcal{E}_2(\widehat{A}, A)$, computed as the square Frobenius norm (Schatten 2-norm) of $A - \widehat{A}$, averaged over 100 trials, and (2) the estimation error $\mathcal{L}_2(\widehat{A}, A)$, computed as the Schatten 2-norm of $A - \widehat{A}$ for a single trial.

3.3 Validation tests

Two primary tests are conducted to validate the theoretical bounds:

- **Estimation Error Decay Rate and Dimension:** This test evaluates the error decay rates for different stability regimes:
  - **Stable Systems:** the predicted decay rate is

$$|\widehat{A} - A|_{S_2} \simeq \frac{d^{1/2}|B|_{S_\infty}}{s_{\min}(S)} \frac{d^{1/2}}{N^{1/2} s_{\min}(C(A_s, B_s))}. \tag{15}$$

– **Unstable Systems:** the theory predicts an exponential decay rate given as

$$|\widehat{A} - A|_{S_2} \simeq \frac{d^{1/2}|B|_{S_\infty}}{s_{\min}(S)} \frac{d^{1/2}}{s^N_{\min}(A) s_{\min}(C(A_u, B_u))}. \tag{16}$$

– **Limit-Stable Systems:** the theory gives the same decay rate up to a dimension mismatch indeed the upper bound rate is

$$|\widehat{A} - A|_{S_2} \lesssim \frac{d^{1/2}|B|_{S_\infty}}{s_{\min}(S)} \frac{d}{N s_{\min}(C(A_{\text{ls}}, B_{\text{ls}}))},$$

The rates above disregard contributions from minor terms and are given for the upper bounds. The rates for the lower bounds are similar with minor changes where instead of taking $s_{\min}$ and $d^{1/2}|B|_{S_\infty}$ you would get $s_{\max}$ and $|B|_{S_2}$, except for limit stable systems where a factor d is missing from the lower bound.

- **Noise and Conditioning Impact:** This test investigates the sensitivity of estimation error to noise characteristics and matrix conditioning, the tests are conducted by:
  - Varying the noise variance σ^2 with a noise matrix σB of fixed B and varying the condition number cond(B) for a fixed noise variance σ^2.
  - Varying the conditioning of the controllability matrix $C(A, B)$ between regimes of well-conditioned and poorly-conditioned.

3.4 Validation with real data

To assess the practical applicability of the theoretical bounds, validation was performed using real-world datasets. Unlike simulated data, where the system matrices A and B are generated with controlled properties, real-world datasets provide only the observed time series $(x_i)_i$. Therefore, the characteristics of A and B must be estimated before proceeding to the estimation and the analysis. Preprocessing steps include normalization, noise filtering, and ensuring compatibility with the modeling framework. The following datasets were used for validation:

3.4.1 Dynamical System Multivariate Time Series Forecast Dataset. This dataset, available on Kaggle [Data], contains multivariate time series data generated from a dynamical system. It is well-suited for testing control theory models, including Linear Time-Invariant (LTI) systems, as it captures complex time-dependent relationships in a controlled environment. The dataset provides a structured setting to validate theoretical bounds under conditions that mimic dynamic system behavior.

3.4.2 Beijing Multi-Site Air-Quality Dataset. This dataset, hosted by the UCI Machine Learning Repository [Data], tracks hourly measurements of air pollutants across multiple locations in Beijing. Here, we focus on the area of Aotizhongxin. The dynamics of air quality, influenced by time-varying processes such as weather conditions and emissions, can be approximated using LTI models. This dataset is particularly valuable for evaluating the robustness of theoretical bounds when applied to noisy, real-world processes.

Validation on these datasets involves estimating the system matrices A and B and comparing the observed estimation errors with the theoretical bounds derived from the controllability and noise structure. This real-world analysis complements the simulated experiments, highlighting the bounds' applicability across diverse scenarios.

4. Results and analysis

4.1 Error decay rates

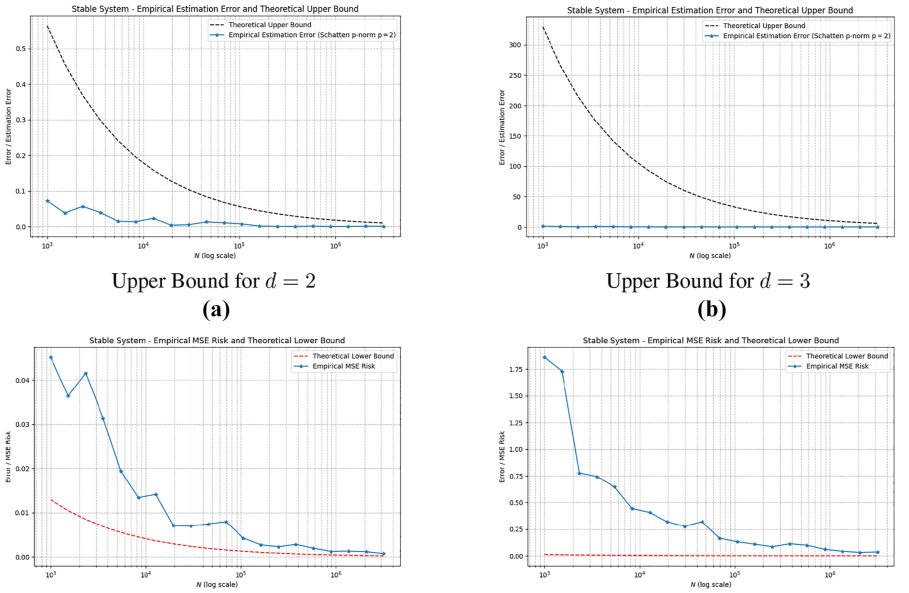
This section evaluates the error and risk decay rates for different system configurations depending on the stability of the matrix *A* and validates theoretical results that we summarized previously.

4.1.1 Stable systems. We conduct a series of simulations on stable systems with varying dimensions, tracking the estimation error and MSE risk. These results are then compared to the theoretical upper and lower bounds, which are plotted for reference.

- **Theoretical Prediction**: According to the previously mentioned theoretical results in (15), the upper and lower estimation risk both agree in their dependence on the sample size *N* and the dimension *d*.
- **Empirical Results**: This first series of simulation represented in Figure 1 show that indeed the estimated error stays between the predicted upper bound on the estimation loss and the predicted lower bound for the mean square estimation risk.

4.1.2 Unstable systems. We conduct a second series of simulation with unstable systems of varying dimensions while tracking the estimation error. Since the estimation error is theorized to decrease exponentially fast in the sample size with the largest eigenvalue, the experiment is made for relatively small samples and the simulated systems do not have eigenvalues deep in the instability region.

- **Theoretical Prediction**: According to the theoretical results the upper estimation loss and the lower estimation risk are both of the same order of dependence on the dimension *d*, and decrease exponentially fast with the sample size. The order for the upper bound



Upper Bound for $d = 2$ **(a)**

Upper Bound for $d = 3$ **(b)**

Lower Bound for $d = 2$ **(c)**

Lower Bound for $d = 3$ **(d)**

Note(s): d = 2 on the left, d = 3 on the right

Source(s): Figure by authors

Figure 1. Comparison of upper and lower bounds for stable system

depends on the smallest singular value of A and for the lower bound on the largest singular value of A. Up to this mismatch the order of both is roughly as in (16).

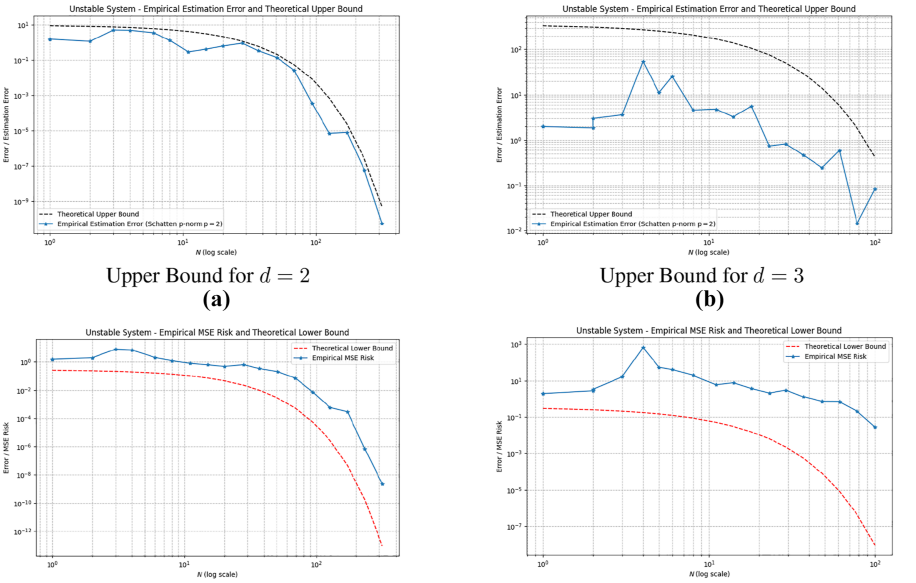
- **Empirical Results**: The simulations in Figure 2 also confirm the theoretical error rates since the estimated error stays between the predicted upper bound and lower bound. The rate in N is indeed exponential and depends on the spectral properties of the matrix A. This exponential behavior makes the estimation task rather easy, something to be contrasted with the control task since controlling an unstable system is in general an onerous task as the controller has to bring the system into stability first.

4.1.3 Limit stable systems. We conduct a third series of simulations with limit stable systems of varying dimensions while tracking the estimation error. We recall that there is a mismatch in the dependence on the dimension of the matrix A in the theory between upper and lower bound estimation error estimates, plus an improvement in the dependence on the sample size N over stable systems.

- **Theoretical Prediction**: According to the theoretical results the upper estimation loss and the lower estimation risk are both of the same order of dependence on the sample size N up to some lower order logarithmic term, while they differ by a factor of d in the dimension. Indeed the least square estimator satisfies the following bound on the estimation loss:

$$\mathcal{L}_2(\widehat{A}, A) \lesssim \frac{|B|_{S_2}}{s_{\min}(S)} \frac{d^{1/2}(d + \log\log N)^{1/2}}{N s_{\min}(C(A,B))}, \tag{17}$$

and the following lower bound for the mean square estimation risk:



Upper Bound for $d = 2$ **(a)**

Upper Bound for $d = 3$ **(b)**

Lower Bound for $d = 2$ **(c)**

Lower Bound for $d = 3$ **(d)**

Note(s): d = 2 on the left, d = 3 on the right

Source(s): Figure by authors

Figure 2. Comparison of upper and lower bounds for unstable system

$$\mathcal{E}_2(\widehat{A}, A) \gtrsim \frac{|B|_{S_2}}{s_{\max}(S)} \frac{1}{N s_{\max}(C(A,B))}. \tag{18}$$

- **Empirical Results**: While the simulations in Figure 3 confirms the theoretical error rates dependence on the sample size *N*, the dependence on the dimension is not as clear although the empirical results are closer to the upper bounds which suggest that the convergence rate for the error estimate is *d*/*N*.

4.2 Noise sensitivity, conditioning and controllability impact

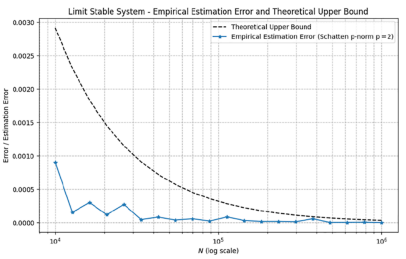
This section presents an empirical study to validate the theoretical upper and lower bounds on estimation accuracy. We evaluate the effects of noise level, structure, and conditioning of *B*, as well as the controllability of the system matrix *C*(*A*, *B*). For all experiments, the Schatten norm $|B|_{S_2} = 1$ is maintained to focus on the influence of noise conditioning and system controllability.

In this study:

(1) The number of observations $N = 1{,}000$ (Stable case), $N = 100$ (Unstable case), $N = 1{,}000$ (Limit stable case);

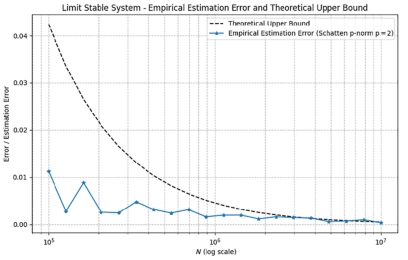
(2) For a square matrix *B*, a 2×2 dimension is used;

(3) For a rectangular matrix *B*, a 2×3 dimension is used and



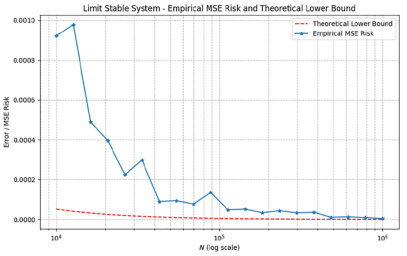
Upper Bound for $d = 2$

(a)



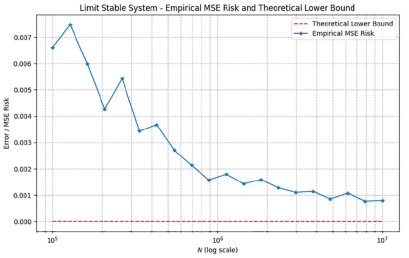
Upper Bound for $d = 3$

(b)



Lower Bound for $d = 2$

(c)



Lower Bound for $d = 3$

(d)

Note(s): d = 2 on the left, d = 3 on the right

Source(s): Figure by authors

Figure 3. Comparison of upper and lower bounds for limit stable system

(4) For the tables related to controllability and noise conditioning impact on estimation accuracy, the noise level is set to $\sigma = 1$.

Tables 1, 3, and 5 show results where B structure, noise level σ, and conditioning cond(B) are varied. Estimation error, mean squared estimation risk (MSE risk), and theoretical bounds are reported to measure the impact of noise structure on estimation accuracy.

Tables 2, 4, and 6 present results for systems with varying controllability $C(A, B)$ conditions, which are either well-conditioned or poorly-conditioned, in conjunction with different structures of B and its condition number cond(B). The results demonstrate how controllability affects estimation error under these varying conditions.

Table 1. Noise characteristics impact on estimation accuracy and bounds

Dim. B	cond(B)	Noise level	MSE risk	Est. err	Upper bnd	Lower bnd
(2, 2)	2.13	0.1	0.060	0.038	0.295	0.040
(2, 3)	1.83	0.1	0.059	0.046	0.254	0.041
(2, 2)	6149.42	0.1	13.927	9.969	90.886	0.037
(2, 3)	7251.18	0.1	1.163	1.151	11.993	0.036
(2, 2)	2.13	1.0	0.053	0.053	0.295	0.040
(2, 3)	1.83	1.0	0.063	0.081	0.254	0.041
(2, 2)	6149.42	1.0	3.164	3.134	24.977	0.037
(2, 3)	7251.18	1.0	1.149	1.134	13.436	0.036

Source(s): Table by authors

Table 2. Controllability and noise conditioning impact on estimation accuracy and bounds

Dim. B	Rank(B)	Cond(B)	Cond(C(A, B))	MSE risk	Est. error	Upper bnd	Lower bnd
(2, 2)	2	1.18	1.12	0.045	0.019	0.178	0.02734
(2, 2)	2	1.87	1483.43	0.057	0.035	0.233	0.03374
(2, 2)	2	235.50	8.64	0.162	0.078	2.550	0.04250
(2, 2)	2	1396.32	1386.36	6.548	3.780	64.674	0.04262
(2, 3)	2	2.06	2.08	0.076	0.059	0.259	0.03299
(2, 3)	1	1.36	1.35	0.045	0.035	0.202	0.02769
(2, 2)	1	62.32	6.00	0.147	0.041	1.429	0.03511
(2, 2)	1	10052.75	1183.64	1.070	0.382	18.200	0.04262

Source(s): Table by authors

Table 3. Noise characteristics impact on estimation accuracy and bounds

Dim. B	Cond(B)	Noise level	MSE risk	Est. err	Upper bnd	Lower bnd
(2, 2)	2.13	0.1	0.342	0.348	0.029	5.28×10^{-4}
(2, 3)	5.26	0.1	0.402	0.335	0.072	4.85×10^{-4}
(2, 2)	7170.76	0.1	0.191	0.054	23.617	4.77×10^{-4}
(2, 3)	6521.42	0.1	0.535	0.215	21.717	4.78×10^{-4}
(2, 2)	2.13	1.0	0.205	0.316	0.029	5.28×10^{-4}
(2, 3)	5.26	1.0	0.509	0.078	0.072	4.85×10^{-4}
(2, 2)	7170.76	1.0	0.755	2.203	23.617	4.77×10^{-4}
(2, 3)	6521.42	1.0	1.111	0.029	21.717	4.78×10^{-4}

Source(s): Table by authors

Table 4. Controllability and noise conditioning impact on estimation accuracy and bounds

Dim. B	Rank(B)	Cond(B)	Cond(C(A, B))	MSE risk	Est. error	Upper bnd	Lower bnd
(2, 2)	2	1.18	1.11	0.0766	0.0028	0.0243	2.42×10^{-5}
(2, 2)	2	2.34	131.82	0.0078	0.0169	0.0580	3.23×10^{-5}
(2, 2)	2	33.87	8.27	0.1739	0.0036	123.2029	4.87×10^{-6}
(2, 2)	2	25.80	109.71	0.0034	0.0032	25.8779	1.01×10^{-5}
(2, 3)	2	2.40	3.02	0.0061	0.0127	0.0600	3.27×10^{-5}
(2, 3)	1	4.98	8.27	0.0152	0.0032	0.1876	4.10×10^{-5}
(2, 3)	1	2.68	143.65	0.0123	0.0030	0.0702	3.43×10^{-5}
(2, 2)	1	38.77	8.66	0.0634	0.0044	297.3403	3.15×10^{-6}
(2, 2)	1	40.13	124.09	0.2197	0.0021	404.2661	2.75×10^{-6}

Source(s): Table by authors

Table 5. Noise characteristics impact on estimation accuracy and bounds

Dim. B	Cond(B)	Noise level	MSE	Est. err	Upper bnd	Lower bnd
(2, 2)	2.13	0.1	0.005	0.002	0.041	5.52×10^{-4}
(2, 3)	1.83	0.1	0.003	0.001	0.088	5.70×10^{-4}
(2, 2)	5072.54	0.1	0.039	0.024	52.397	5.00×10^{-4}
(2, 3)	6767.88	0.1	0.011	0.002	1.166	5.00×10^{-4}
(2, 2)	2.13	1.0	0.009	0.007	0.041	5.52×10^{-4}
(2, 3)	1.83	1.0	0.005	0.004	0.088	5.70×10^{-4}
(2, 2)	5072.54	1.0	0.047	0.014	52.397	5.00×10^{-4}
(2, 3)	6767.88	1.0	0.013	0.007	1.166	5.00×10^{-4}

Source(s): Table by authors

Table 6. Controllability and noise conditioning impact on estimation accuracy and bounds

Dim. B	Rank(B)	Cond(B)	Cond(C(A, B))	MSE risk	Est. error	Upper bnd	Lower bnd
(2, 2)	2	2.13	2.13	0.005	0.009	0.041	5.36×10^{-4}
(2, 2)	2	3.66	787.28	0.006	0.011	0.405	3.83×10^{-4}
(2, 2)	2	574.63	574.63	0.009	0.003	0.030	9.32×10^{-6}
(2, 2)	2	12.76	9.00	0.022	0.032	25.607	1.80×10^{-4}
(2, 3)	2	1.95	1.95	0.008	0.006	0.101	5.70×10^{-4}
(2, 3)	2	1.30	143.91	0.004	0.004	0.028	7.84×10^{-4}
(2, 3)	1	1.18	1.18	0.008	0.008	0.016	8.50×10^{-4}
(2, 2)	1	20.82	20.82	0.067	0.083	42.183	1.15×10^{-4}

Source(s): Table by authors

4.2.1 Stable case. The theoretical bounds in the table are consistent with the observed errors, particularly for well-conditioned cases, where the bounds are tight. Moreover, the estimation error and the MSE stay close to the lower bound even though the upper bound diverges.

The results show that for stable systems, the error rate stays close to the lower bound regardless of the condition number and the dimension of the matrix B. This stability highlights the robustness of the estimator and validates the theoretical predictions. The result confirms also the observation that the error is invariant to the noise level and is only more sensitive to the noise matrix conditioning.

Indeed, if we take $B = \sigma I_d$, we obtain, for the least square estimator,

$$\mathbb{E}\left(|\widehat{A}-A|_{S_2}^2\right)=\mathbb{E}\left(\left|\left(\sum_{i=1}^{N}x_i x_{i-1}^*\right)\left(\sum_{i=1}^{N}x_{i-1}x_{i-1}^*\right)^{-1}-A\right|_{S_2}^2\right)$$

$$=\mathbb{E}\left(\left|\left(\sum_{i=1}^{N}(Ax_{i-1}+\sigma\varepsilon_{i-1})x_{i-1}^*\right)\left(\sum_{i=1}^{N}x_{i-1}x_{i-1}^*\right)^{-1}-A\right|_{S_2}^2\right) \tag{19}$$

$$=\mathbb{E}\left(\left|\sum_{i=1}^{N}\varepsilon_{i-1}\frac{x_{i-1}^*}{\sigma}\left(\sum_{i=1}^{N}\frac{x_{i-1}}{\sigma}\frac{x_{i-1}^*}{\sigma}\right)^{-1}\right|_{S_2}^2\right)$$

which is independent of the noise variance since $\frac{x_i}{\sigma}=\sum_{k=0}^{i-1}A^{i-1-k}\varepsilon_k$.

Table 2 validates the influence of the controllability matrix $C(A, B)$ on the derived theoretical bounds. For well-conditioned controllability matrices ($\text{cond}(C(A, B)) < 10$), the bounds are tight and align well with the observed MSE risks and estimation errors. However, as $\text{cond}(C(A, B))$ increases, the bounds diverge, particularly for poorly conditioned cases, reflecting the amplified estimation uncertainty predicted by the theory.

The rank and dimension of B have minimal impact on the accuracy as long as $C(A, B)$ is well-conditioned, confirming that the bounds depend primarily on the characteristics of $C(A, B)$. These results emphasize the importance of maintaining well-conditioned controllability matrices to achieve reliable error bounds and accurate estimations.

4.2.2 Unstable case. The theoretical bounds for the unstable case are also consistent with the observed errors and are only affected by the matrices B and $C(A, B)$ conditioning. Nonetheless, we observe that for the well-conditioned case, the experimental results fall outside the theoretical upper and lower bounds both in Tables 3 and 4. This is because the estimation error decreases exponentially fast for unstable matrices, while the theoretical bounds hold for sample sizes that do not allow capturing this phenomenon in the experiment.

Indeed, the theoretical bounds hold for

$$N\gtrsim cond^2(C(A,B))\left(1\vee\frac{1}{s_{\min}^2(C(A,B))}\right)d\log\left(\frac{Ncond(C(A,B))|B|_{S_\infty}}{\delta}\right), \tag{20}$$

which gives theoretical upper and lower bounds that are well below the precision limits. Instead, we report experimental results for $N = 100$ to keep the error rates reasonable. When, doing experiments with ill-conditioned matrices B and $C(A, B)$ the error rate does not decrease too fast, so the MSE and estimation error both stay between the theoretical upper and lower bounds.

4.2.3 Limit stable case. Similar to the stable case, the theoretical bounds for the limit stable case align well with the observed errors for well-conditioned systems, where the bounds remain tight both for well-conditioned or ill-conditioned matrices B and $C(A, B)$. These results are reported in Tables 5 and 6.

5. Real-world data validation

We apply the least squares estimator to two real-world datasets to demonstrate its practical applicability.

5.1 Multivariate time series forecasting dataset

This synthetic dataset simulates multivariate time series data from dynamical systems with stable configuration. The dataset consists of multiple variables evolving over time, capturing interactions between components of the system.

5.1.1 Estimation results and analysis. This section presents the results of the system estimation and analysis. The system matrices A and B were estimated from a large dataset after preprocessing, which included differencing for stationarity. Cross-validation and bootstrap techniques were used to evaluate the robustness and confidence in the estimation. The key results are summarized in Table 7. The estimated A-matrix captures the system dynamics, while the B-matrix represents the effect of white noise disturbances. Eigenvalue analysis of A indicates that the system is stable since all eigenvalues lie within the unit circle. Additionally, the mean cross-validation error and a 99% confidence interval for the estimation error were calculated to assess the accuracy of the estimated matrices. A theoretical upper bound for the estimation error was also derived based on the controllability properties of the system.

5.1.2 Bootstrap analysis and visualization. The accuracy and robustness of the estimated matrices were further validated using bootstrap analysis. Figure 4 shows the distribution of bootstrap estimation errors along with the mean cross-validation error and the 99% confidence interval bounds. The histogram indicates that the errors are tightly clustered, demonstrating the reliability of the estimation.

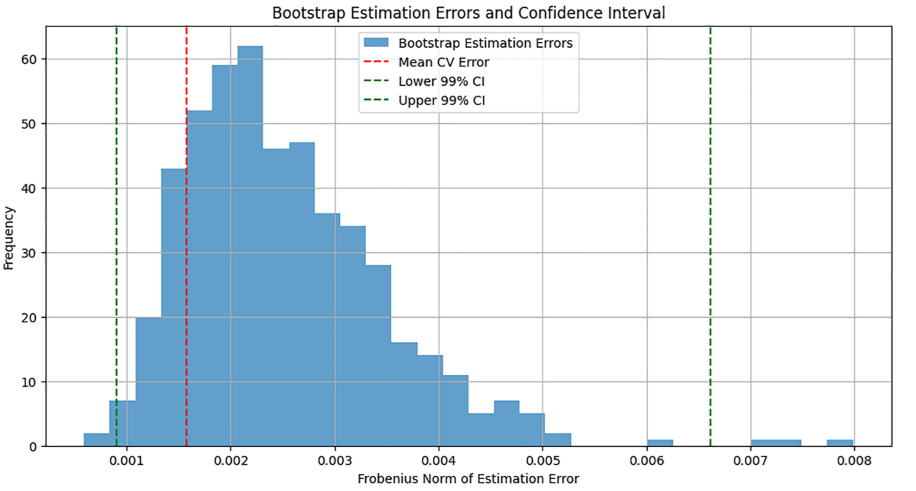
5.1.3 Discussion.

- **Stability Analysis:** The eigenvalues of the estimated A-matrix lie within the unit circle, confirming the stability of the identified system dynamics. This stability is critical for ensuring the validity of the theoretical bounds and the applicability of their rates in practice. In this case, we would expect a decay rate of $(d/N)^{1/2}$ and adjust the sample size accordingly.
- **Cross-Validation and Bootstrap Analysis:** The mean cross-validation error is remarkably low, with a value of 0.00158, and the 99% confidence interval spans a

Table 7. Summary of estimation results

Parameter	Value
Estimated A Matrix	$\begin{bmatrix} -2.23\times10^{-6} & 4.16\times10^{-3} & 4.45\times10^{-4} & 1.80\times10^{-4} \\ -4.77\times10^{-3} & 7.30\times10^{-3} & 7.45\times10^{-2} & -1.87\times10^{-2} \\ 1.35\times10^{-3} & -3.04\times10^{-3} & 8.99\times10^{-1} & 1.43\times10^{-2} \\ -1.03\times10^{-4} & 1.40\times10^{-3} & -2.93\times10^{-2} & 9.81\times10^{-1} \end{bmatrix}$
Estimated B Matrix	$\begin{bmatrix} -1.18\times10^{-3} & -5.87\times10^{-3} & 1.60\times10^{-3} \\ -5.13\times10^{-3} & -3.76\times10^{-3} & 2.52\times10^{-3} \\ 1.64\times10^{-3} & -2.80\times10^{-5} & -2.37\times10^{-3} \\ 5.66\times10^{-4} & -9.85\times10^{-4} & -1.00\times10^{-4} \end{bmatrix}$
Eigenvalues of A	$0.00378991 + 0.00243652i$, $0.00378991–0.00243652i$, 0.90487607, 0.97503101
System Stability	Stable (all eigenvalues within the unit circle)
Mean Cross-Validation Estimation Error	0.0062 (Frobenius norm)
99% Confidence Interval for Estimation Error	Lower Bound: 0.0049, Upper Bound: 0.0161
Theoretical Upper Bound for Estimation Error	0.352
Theoretical Lower Bound for Estimation Risk	0.0047
Empirical Estimation risk	0.6773

Source(s): Table by authors



Source(s): Figure by authors

Figure 4. Distribution of bootstrap estimation errors with confidence interval

narrow range from 0.0009 to 0.0066. Bootstrap analysis further corroborates the reliability of the estimation, demonstrating that errors are tightly clustered and well within the confidence bounds.

- **Conservative Theoretical Upper Bound:** The derived theoretical upper bound of 0.43631 for the estimation error is higher than the observed errors. This discrepancy highlights the conservativeness of the theoretical bound, providing assurance that practical applications are likely to perform even better than theoretically anticipated.
- **Estimation Effectiveness:** The low estimation errors observed across various analyses underscore the effectiveness of the estimation process. These results emphasize the method's ability to capture system dynamics accurately, ensuring reliable forecasting performance for multivariate time series data.

5.2 Beijing Multi-Site Air-Quality Dataset

The dataset, titled *"Beijing Multi-Site Air-Quality Data"*, includes hourly measurements of air pollutants such as PM2.5, PM10, SO_2, NO_2, CO, and O_3 from 12 nationally controlled monitoring sites in Beijing. In this study, we use the data related the area of Aotizhongxin in Beijin. The data spans from March 1, 2013, to February 28, 2017, and exhibits both seasonal patterns and long-term trends influenced by meteorological and anthropogenic factors.

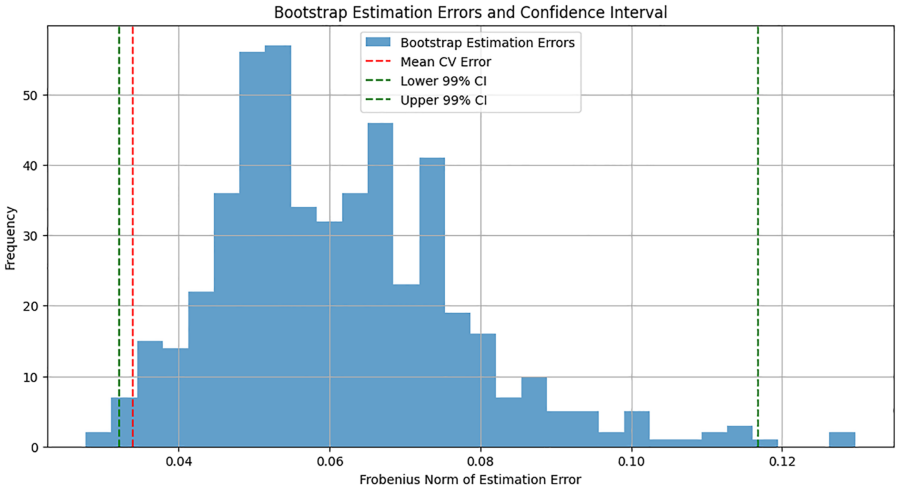
5.2.1 Estimation results and analysis. As in the previous subsection, we present the results of estimating the system matrices A and B from the dataset. The key results are summarized in Table 8. Preprocessing steps included differencing for stationarity and Z-score normalization. The same methods were used to evaluate the robustness and accuracy of the estimation.

5.2.2 Bootstrap analysis and visualization. Bootstrap analysis was conducted to validate the robustness of the estimation process. Figure 5 displays the distribution of bootstrap estimation errors, highlighting that the errors are tightly clustered around the mean, indicating reliable estimation results.

Table 8. Summary of estimation results for Beijing air pollution dataset

Parameter	Value
Estimated A Matrix	$\begin{bmatrix} 0.0824 & -0.0681 & -0.0355 & -0.0599 & -0.0263 \\ -0.0266 & 0.1785 & -0.0452 & -0.0328 & -0.0217 \\ 0.0103 & -0.0118 & 0.0994 & -0.0372 & 0.0618 \\ -0.0137 & -0.0356 & -0.0749 & 0.2092 & -0.0255 \\ 0.0021 & -0.0034 & -0.0367 & 0.0165 & 0.0208 \end{bmatrix}$
Estimated B Matrix	$\begin{bmatrix} 0.0019 & 0.0053 & -0.0068 & -0.0049 \\ -0.0041 & 0.0138 & 0.0002 & -0.0008 \\ 0.0016 & -0.0011 & -0.0016 & -0.0107 \\ 0.0047 & -0.0020 & 0.0002 & -0.0134 \\ -0.0061 & 0.0034 & 0.0077 & 0.0022 \end{bmatrix}$
Eigenvalues of A	$0.0434, 0.0565 \pm 0.0339i, 0.1969, 0.2369$
System Stability	Stable (all eigenvalues within the unit circle)
Mean Cross-Validation Estimation Error	0.0339 (Frobenius norm)
99% Confidence Interval for Estimation Error	Lower Bound: 0.0321, Upper Bound: 0.1168
Theoretical Upper Bound for Estimation Error	2.2927
Theoretical Lower Bound for MSE Risk	0.0857
Empirical MSE Risk	0.2131
Controllability Matrix Rank	5 (Full Rank)
Condition Number of Controllability Matrix	11.92
Smallest Singular Value of Controllability Matrix	0.0017

Source(s): Table by authors



Source(s): Figure by authors

Figure 5. Distribution of bootstrap estimation errors for Beijing dataset

5.2.3 Discussion.

- **System Stability:** The estimated system is confirmed to be stable, as all eigenvalues of the *A*-matrix lie within the unit circle. This stability analysis is critical for ensuring reliable and predictable system dynamics estimates. It allows us to adjust N the sample size to achieve the desired accuracy using the theoretical upper bounds for guidance.
- **Low Estimation Errors:** Both cross-validation and bootstrap analyses demonstrate low estimation errors, with the mean Frobenius norm of the error recorded at 0.0339. These results indicate the high accuracy of the parameter estimation process.
- **Conservative Theoretical Bound:** The theoretical upper bound for the estimation error, calculated as 2.2927, is consistent with the observed results. However, this bound is conservative when compared to the empirical risk measures, providing an additional margin of confidence in practical applications.
- **Controllability Metrics:** The system exhibits strong controllability characteristics, supported by a full-rank controllability matrix and acceptable singular values. These metrics validate the system's well-conditioned nature and enhance the reliability of the estimated parameters.

6. Discussion and conclusion

This study highlights the utility of non-asymptotic bounds in improving econometric modeling and financial system identification. By understanding the finite-time behavior of least squares estimators, practitioners in finance and economics can better handle noise and dimensionality challenges, enabling more robust forecasting and decision-making models. The following key insights are drawn from the analysis:

Theoretical Implications: The results confirm that error decay rates align closely with theoretical predictions under varying stability regimes. For stable systems, the error decreases proportionally to $\mathcal{O}(d^{1/2}/N^{1/2})$, consistent with derived upper and lower bounds. In unstable systems, the exponential decay rate observed is governed by the spectral properties of the system matrix, highlighting the interplay between eigenvalues and sample size. Limit-stable systems exhibit an improved convergence rate, albeit with a mismatch in dimensional dependence between theoretical bounds. These findings reinforce the robustness of theoretical frameworks in capturing the finite-time behavior of least squares estimators across different configurations.

Practical Implications: Noise characteristics and system dimensionality significantly impact estimation accuracy in practical scenarios. Sensitivity tests revealed that estimation errors increase with higher noise levels and poorly conditioned matrices. Real-world validation on the Beijing Air Quality dataset demonstrated that empirical estimation errors are consistently below conservative theoretical upper bounds, confirming the practical utility of these bounds. Moreover, system stability, verified through eigenvalue analysis, emerged as a critical factor for ensuring reliable estimation accuracy in dynamic environments since it allows us to adjust the sample size.

In conclusion, this work bridges theoretical and empirical analyses, demonstrating the efficacy of least squares estimators in diverse scenarios. While confirming theoretical bounds, the study also highlights practical uses of the theoretical upper and lower bounds for controlling the accuracy by estimating the system complexity parameters and adjusting the sample size accordingly. By bridging theoretical insights and practical validations, this study provides valuable tools for modeling economic systems, such as analyzing the stability of macroeconomic policies, and financial systems, like assessing the impact of shocks on asset prices. Future work can extend this research by deriving similar bounds to high dimensional nonlinear systems and applying them to more complex data observed in global financial markets and economic interactions.

References

Arora, S., Hazan, E., Lee, H., Singh, K., Zhang, C. and Zhang, Y. (2018), "Towards provable control for unknown linear dynamical systems", available at: https://openreview.net/forum?id=BygpQlbA-

Djehiche, B. and Mazhar, O. (2021), "Non-asymptotic estimation lower bounds for LTI state space models with Cramér-Rao and van Trees", *arXiv preprint*, arXiv:2109.08582, available at: https://arxiv.org/abs/2109.08582

Djehiche, B. and Mazhar, O. (2022), "Efficient learning of hidden state LTI state space models of unknown order", *arXiv preprint*, arXiv:2202.01625, available at: https://arxiv.org/abs/2202.01625

Djehiche, B. and Mazhar, O. (2025), "An optimal non-asymptotic journey through system identification with least squares", in Thach, N.N., Trung, N.D., Ha, D.T. and Kreinovich, V. (Eds), *Artificial Intelligence and Machine Learning for Econometrics: Applications, Regulation (and Related Topics)*, Springer, Studies in Systems, Decision, and Control series.

Jedra, Y. and Proutière, A. (2019), "Sample complexity lower bounds for linear system identification", *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics (AISTATS 2019)*, pp. 2397-2405.

Ljung, L. (1976a), "Consistency of the least-squares identification method", *IEEE Transactions on Automatic Control*, Vol. 21 No. 5, pp. 779-781, doi: 10.1109/tac.1976.1101344.

Ljung, L. (1976b), "On the consistency of prediction error identification methods", in Mehra, R.K. and Lainiotis, D.G. (Eds), *Mathematics in Science and Engineering*, Elsevier, Vol. 126, pp. 121-164, doi: 10.1016/s0076-5392(08)60871-1.

Ljung, L. (1986), *System Identification: Theory for the User*, Prentice-Hall, 0138816409.

Oymak, S. and Ozay, N. (2019), "Non-asymptotic identification of LTI systems from a single trajectory", *IEEE Transactions on Automatic Control*, Vol. 64 No. 6, pp. 2345-2355.

Simchowitz, M., Mania, H., Tu, S., Jordan, M.I. and Recht, B. (2018), "Learning without mixing: towards a sharp analysis of linear system identification", *Proceedings of the 31st Conference on Learning Theory (COLT 2018)*, pp. 1298-1320.

Corresponding author
Boualem Djehiche can be contacted at: boualem@kth.se