Error analysis on the initial state reconstruction problem

Rocio Diaz Martin¹ · Ivan Medri¹ · Juliana Osorio²

Received: 19 September 2022 / Accepted: 3 May 2023 / Published online: 17 May 2023
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2023

Abstract
In this paper, we propose a method to estimate the initial state of a linear dynamical system from noisy observation based on Kalman Filters and Optimal Smoothing techniques. The method allows the user to have estimations in real-time, that is, to have a new estimation for each new observation. Moreover, at each step, the covariance matrix of the error is found, and is the one that minimizes the error variance after applying a linear filter. We analyze the stability of the method when there is no noise in the dynamics. This hypothesis is the one usually seen in dynamical sampling articles, and differs from the classical framework in control theory in which the covariance of the system noise plays a fundamental role. In particular, we prove that the dynamic of the state estimator error is always Lyapunov stable. Also, necessary and sufficient conditions are given to guarantee asymptotic stability for the error dynamics of a linear time-invariant dynamical system, which is itself a linear time-variant system.

Keywords  Control theory · Dynamical sampling · Kalman filter · Observability · Stability

Mathematics Subject Classification  93B07 · 93C05 · 93C10 · 93D05 · 93D20 · 94A12 · 94A20

Communicated by Mark Iwen.

Rocio Diaz Martin, Ivan Medri and Juliana Osorio contributed equally to this work.

Ivan Medri
cocarojasjorgeluis@gmail.com
Rocio Diaz Martin
rocio.p.diaz.martin@vanderbilt.edu
Juliana Osorio
juli.osorio@gmail.com

1 Department of Mathematics, Vanderbilt University, 1326 Stevenson Center Ln, Nashville, TN 37212, USA
2 Departamento de Matematica, FCEyN, Universidad de Buenos Aires, Ciudad Universitaria, Buenos Aires 1428, Argentina
1 Introduction

The aim of this paper is to recover the initial state of a given dynamical system (without noise in the dynamic) from noisy observations and principally to analyze the asymptotic stability of the method.

Consider the discrete-time linear dynamical system

\[
\begin{align*}
    x(k + 1) &= A_{k+1}x(k), & x(0) &= x_0 \\
    y(k) &= H_k x(k) + v_k, & v_k &\sim N(0, R_k)
\end{align*}
\]  

(1)

where, for each \( k \in \mathbb{N} \cup \{0\} \), \( x(k) \in \mathbb{C}^d \) are the states of the system, \( y(k) \in \mathbb{C}^m \) are the observations, and \( A_k \) and \( H_k \) are matrices in \( \mathbb{C}^{d \times d} \) and \( \mathbb{C}^{m \times d} \), respectively, called the dynamic operator and the observation or sampling operator at time \( k \). In applications, it is commonly assumed \( m < d \). If the dynamic and sampling operators are constant (\( A_k = A \), and \( H_k = H \)), we say that the system is linear time-invariant (LTI), otherwise it is linear time-variant (LTV). The sequence \( \{v_k\} \) is a random process with a known distribution. Typically, the random vectors \( v_k \) are assumed to be normal with zero means, such that \( \mathbb{E}(v_j v_k^*) = R_k \delta_{j,k} \), where the matrices \( R_k \) are symmetric positive definite.

In this paper, we deal with the observability problem for the dynamical system (1). That is, we will assume that the initial state \( x_0 \) is unknown and the goal is to recover or estimate it from the set of measurements \( \{y(k)\}_k \). In the field of Sampling Theory, this problem is known as the Dynamical Sampling Problem [1–5].

Some methods for approaching this problem, proposed in Control Theory, are known as Kalman filter or Optimal Smoothing methods (see, e.g., [6–10]). They can be deduced by making a sequence of estimates that come from solving a recursive weighted least square problem, and are applied for systems that include both, noise in the observations and in the dynamic. A priori, the development of Kalman filters does not center on estimating the initial state by knowing samples in later times, but is based on reconstructing the final state through the knowledge of previous observations. That is, the focus is on filtering an actual state of the system or on predicting future states from past noisy observations. Nonetheless, both problems, the one of recovering the initial state and the one of recovering the final state, can be related by introducing auxiliary variables that remain static over time, thus obtaining that the final state of the augmented system always contains the initial state. This alternate form of the Kalman filter is known under the name of Optimal Smoothing (see, e.g., [9, Chapter 9]). However, in presence of noise, this leads to the reconstruction problem of the initial state being less robust than that of the final state, and then a specific error control is required. Colloquially speaking, a new observation at step \( n \) is a fresh observation of the state at time \( n \) but it is a dated and corrupted (by the evolution of the noise) observation of the initial state that happened a long time ago.

Notably, in the recent work [11], the author reviewed the classical stability analysis of Kalman filter methods to include dynamical systems where the dynamics are unperturbed, as the one considered in (1). Not having noise in the dynamic but only in
the measurements is a common scenario in physical-world applications where information about the models is precise. This framework is also the one usually considered in recent dynamical sampling articles [2, 4, 12–15].

It was only in the aforementioned work [11] that the stability analysis was considered for noiseless dynamical systems and final state reconstruction. In classical texts, the asymptotic stability of a filter is only guaranteed under the presence of noise. See, for e.g., [7, Ch. 7, Sec. 6, Th. 7.4] where complete controllability is assumed, which requires noise in the state evolution and moreover with a positive definite covariance matrix. These conditions are usually not satisfied in the adaptations of Kalman filters to optimal smoothing. Yet, a proper analysis of this is never considered in the literature. This is one of the main reasons why we try to solve the proposed problem and give an explicit analysis of the stability of the method. Moreover, for initial state reconstruction, the noiseless approach helps to understand how the quality of the measurements decay based solely on the system dynamics. This is a phenomenon that should be considered separately from how corrupted a measurement is due to noise in the dynamic. It is the combination of both issues, dynamic evolution and measurements, that will determine how good an observation is. Indeed, whereas one part determines whether new measurements still conserve a good amount of information about the initial state, the other part speaks about how big is the noise content of the data.

For LTI dynamical systems, when the observation operator $H$ is presented as

$$Hx = \{\langle x, g \rangle \}_{g \in G},$$

for some set of vectors $G \subseteq \mathbb{C}^d$ (of cardinality less than or equal to $d$), the observability problem is known as the dynamical sampling problem. The name resembles the fact that the collection of the observations can be viewed as time-space samples of the form

$$\{y(k)\}_{0 \leq k \leq L-1} = \{\langle x(k), g \rangle \}_{g \in G, 0 \leq k \leq L-1} = \{\langle A^k x_0, g \rangle \}_{g \in G, 0 \leq k \leq L-1}. \quad (2)$$

The dynamical sampling problem has been widely studied (see, e.g., [12, 16] and the references therein). For an explicit formulation of the connections between Dynamical Sampling and Control Theory we mention the recent article [4].

We emphasize that in the framework of dynamical sampling, no perturbation in the dynamic operator is considered. The papers [13, 16] developed algorithms for solving the dynamical sampling problem where the space measurements are inexact. The techniques behind these papers lie on least squares methods and denoising using Cadzow algorithms [17]. Furthermore, the authors treat the case where the dynamic operator $A$ is unknown. However, several assumptions on the dynamic operator are made. In particular, the case of a circulant matrix $A$ was analyzed in great detail. With this extra assumption, Fourier analysis techniques are suitable.

Our objective is mainly to analyze the stability of the dynamic of the errors resulting from an initial state estimation method for a non-corrupted discrete-time dynamical system based on noisy observations.

We organize the paper as follows: In Sect. 2 we introduce the notation used throughout the paper. Then, we define what we denote as a uniformly observable dynamical...
system by adapting the notion of uniformly completely observable dynamical system given in [7, 11]. Since our goal is to reconstruct the initial state (not the final state), we translate the requirements on the so-called information matrix (cf. [7]) to the observability matrix (given in (6)). Apart from that, we include some preliminary results from stability theory of dynamical systems. The main contributions of this paper start in Sect. 3. We describe a method to recover the initial state of a system. This approach is based on Kalman methods (or equivalently, based on recursive least squares algorithms). Therefore, the ideas to set up the algorithm are not new in the area of control theory. However, we particularly seek to provide a new scheme to those given on the side of dynamical sampling theory (see Remark 1) by introducing techniques from other fields. Finally, in Sect. 4 we deal with the problem of stability of the method. For general LTV dynamical systems we obtain Lyapunov stability for the dynamic of the estimation error (Theorem 5), whereas for LTI systems we prove asymptotic stability under an extra hypothesis on the eigenvalues of the evolution matrix $A$ (Theorem 8). Numerical examples are presented at the end of this section.

2 Preliminaries

2.1 Notation and assumptions

For the dynamical system (1) the initial state $x_0$ is a deterministic but unknown vector that we want to estimate. For ease in the exposition, we use the following convention:

$$A(k, k) := \mathbb{I}, \quad A(k + 1, k) := A_{k+1},$$

where $\mathbb{I}$ denotes the identity matrix. For all $k \in \mathbb{N} \cup \{0\}$ the matrices $A_k$ are assumed to be invertible, so the dynamic is time-reversible. If $j < k$

$$A(k, j) = \prod_{\ell=0}^{k-j-1} A_{k-\ell} = A_k A_{k-1} \cdots A_{j+2} A_{j+1}$$

and

$$A(j, k) = A^{-1}(k, j).$$

The matrix $A(k, j)$ is called the state transition matrix of the system and with this notation we can go from the state $x(j)$ at time step $j$ to the state $x(k)$ at time step $k$ via $x(k) = A(k, j)x(j)$. Then, we have

$$\begin{cases}
x(k) = A(k, 0)x_0 \\
y(k) = H_k A(k, 0)x_0 + v_k
\end{cases} \quad (3)$$

Also, we denote

$$\tilde{H}_k := H_k A(k, 0) \quad (4)$$

which is the operator that observes the exact evolution of the initial state $x_0$ at time $k$. 

\copyright Birkhäuser
For the case of a linear time-invariant systems (LTI), we have \( A_k = A \) and \( H_k = H \), therefore the transition matrices and \( \tilde{H}_k \) take a simpler form:
\[
A(k, 0) = A^k, \quad \tilde{H}_k = HA^k.
\]

Finally, regarding the noise covariances \( R_k \), we assume that they are all positive definite \( m \times m \) matrices with a strictly positive lower bound \( \sigma^2 I \), i.e,
\[
R_k \geq \sigma^2 I \quad \forall k \in \mathbb{N} \cup \{0\}.
\]

### 2.2 Observability

One may say that to observe the system (1) means to recover the initial data \( x_0 \) by knowing only the output function \( k \mapsto y(k) \). With more mathematical rigor, notice that observations of the dynamical system (1) are given by
\[
\begin{bmatrix}
y(0) \\
y(1) \\
\vdots \\
y(L-1)
\end{bmatrix} = \mathcal{A}_L x_0 + \begin{bmatrix}
v_0 \\
v_1 \\
\vdots \\
v_{L-1}
\end{bmatrix} \tag{5}
\]
where \( \mathcal{A}_L \) is the \( mL \times d \) block matrix
\[
\mathcal{A}_L := \begin{bmatrix}
H_0 \\
H_1 A(1, 0) \\
\vdots \\
H_{L-1} A(L - 1, 0)
\end{bmatrix} = \begin{bmatrix}
\tilde{H}_0 \\
\tilde{H}_1 \\
\vdots \\
\tilde{H}_{L-1}
\end{bmatrix}.
\]

Let \( R_L \) be the \( mL \times mL \) diagonal block matrix
\[
R_L := \begin{bmatrix}
R_0 & & \\
& R_1 & \\
& & \ddots \\
& & & R_{L-1}
\end{bmatrix}.
\]

The observability problem consists of solving a weighted least squares problem (with weight given by \( R^{-1}_L \)) which reduces to see if \( \mathcal{A}_L^T R^{-1}_L \mathcal{A}_L \) is positive definite. This is equivalent to
\[
\mathcal{O}(L, 0) := \sum_{j=0}^{L-1} A(j, 0)^* H_j^* R^{-1}_j H_j A(j, 0) > 0. \tag{6}
\]

The matrix \( \mathcal{O} \) is called the observability matrix. We mention that in the noise-free case, \( \mathcal{O}(L, 0) = \sum_{j=0}^{L-1} A(j, 0)^* H_j^* H_j A(j, 0) \).
Definition 1 The dynamical system (1) is said to be uniformly observable if there exist \( L \in \mathbb{N} \) and \( \rho > 0 \) such that

\[
\mathcal{O}(k + L, k) = \sum_{j=k}^{k+L-1} A(j, k)^* H_j^* R_j^{-1} H_j A(j, k) \geq \rho I \quad \text{for all } k \geq 0.
\]

For the case of an LTI dynamical systems, uniform observability is equivalent to observability, that is, there exists \( L > 0 \) such that the observability matrix

\[
\mathcal{O}(L, 0) = \sum_{j=0}^{L-1} (A^j)^* H^* R_j^{-1} H A^j
\]

is strictly positive definite.

Remark 1 Without noise, to solve the dynamical sampling problem is equivalent to have \( \mathcal{A}_L \) of full-rank. Indeed, the frame operator \( \mathcal{A}_L^* \mathcal{A}_L \) is the observability matrix \( \mathcal{O}(L, 0) \) associated to the LTI dynamical system.

We point out that in [16], assuming that for some fixed number \( L \geq 1 \) the matrix \( \mathcal{A}_N \) has full rank for every \( N \geq L \), and considering noisy measurements as in (5) with random noise having \( \mathbb{E}(v_j) = 0 \) and the same covariance matrix \( \sigma^2 I \) for all \( j \in \mathbb{N} \cup \{0\} \) (for some \( \sigma > 0 \)), the authors propose the \( N \)-th estimation \( \tilde{x}_N \) of the initial state as

\[
\tilde{x}_N = \arg\min_{x \in \mathbb{C}^d} \sum_{j=0}^{N-1} \|H A^j x - y(j)\|_2^2 \quad (7)
\]

which is a classical linear least squares problem. The method used in [16] to solve (7) is based on an update of a proper QR decomposition which requires the new observation and the \( R \) factor from the previous step. Then the authors also considered other situations, such as the problem where the dynamic operator \( A \) is unknown, and also they worked with the case when \( A \) is of convolution type, and applied Cadzow-like techniques, that is beyond the scope of this work. In this paper, we will tackle the problem differently (allowing time-variant dynamical systems), and for updating the \( N \)-th estimation we will only require the previous estimation \( (N - 1) \) with its corresponding covariance matrix, and the last observation \( y(N - 1) \).

2.3 Lyapunov stability

The algorithm proposed in this work to estimate the initial state of a given observable system will produce a new dynamical system for the estimation errors. Since our goal is to prove the stability for this time-varying dynamical system, we recall different notions of stability and characterizations for asymptotic stability in the case of linear time-varying systems. For a clarifying exposition on the subject see [18, 19].
Definition 2  Given the discrete dynamical system

$$x(k + 1) = \Psi_{k+1}x(k), \quad x(k_0) = x_0, \quad k \geq k_0$$  \hspace{1cm} (8)$$

with $\Psi_k$ a linear operator for all $k$, we say that the zero solution (the only equilibrium point) of (8) is:

(i) **Lyapunov stable**, if for every $k_0 \in \mathbb{N}$ there exists $\delta = \delta(\epsilon, k_0) > 0$ such that $\|x_0\| < \delta$ implies that $\|x(k)\| < \epsilon$ for all $k \geq k_0$.

(ii) **Uniformly stable**, if for every $\epsilon > 0$ there exists $\delta = \delta(\epsilon) > 0$, independent on $k_0$ such that $\|x_0\| < \delta$ implies that $\|x(k)\| < \epsilon$ for all $k \geq k_0$.

(iii) **Asymptotically stable**, if it is Lyapunov stable and for every $k_0 \in \mathbb{N}$ there exists $\delta = \delta(k_0) > 0$ such that $\|x_0\| < \delta$ implies $\lim_{k \to \infty} x(k) = 0$.

(iv) **Uniformly asymptotically stable**, if it is uniformly stable and there exists $\delta > 0$, independent of $k_0$, such that $\lim_{k \to \infty} \|x(k)\| = 0$ uniformly in $k_0$ for all $\|x_0\| < \delta$.

(v) **Globally uniformly asymptotically stable**, if for every $x_0$, $\lim_{k \to \infty} \|x(k)\| = 0$.

Proposition 1  (cf. [18]) Let $\Psi(k, k_0)$ denote the state transition matrix for (8) from time $k_0$ to time $k$, then the following are equivalent:

1. The equilibrium point of system (8) is uniformly asymptotically stable.
2. $\lim_{k \to \infty} \|\Psi(k, k_0)\| = 0$ uniformly in $k_0$.
3. $\|\Psi(k, k_0)\| \leq \alpha e^{-\lambda(k-k_0)} \forall k \geq k_0$ for some positive constants $\alpha, \lambda$.
4. The equilibrium point of the system (8) is globally uniformly asymptotically stable.

When item 3. above holds, the zero solution of (8) is called **exponentially stable**. Thus, Proposition 1 shows that in the cases of linear systems asymptotic stability and exponential stability are equivalent.

3 Kalman-based method for initial state recovering

In this section we develop an algorithm to recover the initial state of a general linear dynamical system in the presence of noisy observations with known distributions as in (1). The idea of this method is to estimate the initial state $x_0$ in the following way: We start from an initial guess denoted by $\hat{x}_0$ and for each time step we improve the estimation by correcting the previous one with a so-called **innovation term**. In this way, for the next estimation we take into account the last available observation of the system. More explicitly, the estimator is of the form

$$\hat{x}_k = \hat{x}_{k-1} + K_k \left( y(k - 1) - \tilde{H}_{k-1}\hat{x}_{k-1} \right) \quad k = 1, 2, \ldots$$

$$\hat{x}_0 = \text{initial guess}$$  \hspace{1cm} (9)$$

where $\hat{x}_k$ denotes the $k$-th estimation of the initial state $x_0$, and $\tilde{H}_{k-1}$ is given in (4). The matrices $K_k$ are to be defined according to some optimal criterion.
Note that the \( k \)-th estimation of the initial state \( \tilde{x}_k \) takes into account all the first \( k \) noisy observations \( \{y(n)\}_{n=0}^{k-1} \). In fact, the matrices \( \tilde{H}_{k-1} \) can be thought of as observation matrices of the initial state \( x_0 \).

Let \( e_k := \tilde{x}_k - x_0 \) be the error given by the \( k \)-th estimation of the initial state \( x_0 \). Its covariance matrix is given by

\[
P_k := \text{Cov}(e_k, e_k) = \mathbb{E} \left[ (\tilde{x}_k - x_0)(\tilde{x}_k - x_0)^* \right] - \mathbb{E} \left[ (\tilde{x}_k - x_0) \right] \mathbb{E} \left[ (\tilde{x}_k - x_0) \right]^*.
\]  

For each step \( k \), the matrix \( K_k \) is chosen such that it minimizes the following energy functional

\[
J(\tilde{x}_k) = \frac{1}{2} \text{Tr}(P_k).
\]

The trace of the covariance is called the variance of the error estimation. In this sense, the defined estimator is the linear minimum variance estimator for the initial state (cf. [7, Example 7.4]) among those which take into account the information given by \( \{y(n)\} \) given in (3).

The next theorem provides the formulas needed to write the algorithm to compute the estimation of the initial state of our system.

**Theorem 2** Given the dynamical system (1), the solution to the minimization problem given above is

\[
K_k = P_{k-1} \tilde{H}_{k-1}^* \left( \tilde{H}_{k-1} P_{k-1} \tilde{H}_{k-1}^* + R_{k-1} \right)^{-1} (11)
\]

and the update of the covariance error matrix is given by

\[
P_k = P_{k-1} - K_k \tilde{H}_{k-1} P_{k-1} (12)
\]

with \( P_0 \) the initial guess for the covariance error.

**Proof** From equations (1) and (9) we have a recursive formula for the error

\[
e_k = \tilde{x}_k - x_0 = (I - K_k \tilde{H}_{k-1})\tilde{x}_{k-1} + K_k y(k-1) - x_0
\]

\[
= (I - K_k \tilde{H}_{k-1})\tilde{x}_{k-1} + K_k (\tilde{H}_{k-1}x_0 + v_{k-1}) - x_0
\]

\[
= (I - K_k \tilde{H}_{k-1})e_{k-1} + K_k v_{k-1}. (13)
\]

Therefore, the covariance matrix of the error at the \( k \)-th estimation is

\[
P_k = \mathbb{E} \left[ ((I - K_k \tilde{H}_{k-1})e_{k-1} + K_k v_{k-1})((I - K_k \tilde{H}_{k-1})e_{k-1} + K_k v_{k-1})^* \right] - \mathbb{E} \left[ (I - K_k \tilde{H}_{k-1})e_{k-1} + K_k v_{k-1} \right] \mathbb{E} \left[ (I - K_k \tilde{H}_{k-1})e_{k-1} + K_k v_{k-1} \right]^*
\]

\[
= (I - K_k \tilde{H}_{k-1}) \mathbb{E} (e_{k-1}^* e_{k-1}^*) (I - K_k \tilde{H}_{k-1})^* + K_k \mathbb{E} (e_{k-1}^* v_{k-1}^*) K_k^*
\]

\[
+ (I - K_k \tilde{H}_{k-1}) \mathbb{E} (v_{k-1}^* e_{k-1}^*) K_k^* + K_k \mathbb{E} (v_{k-1}^* v_{k-1}) (I - K_k \tilde{H}_{k-1})^*
\]

\[
- (I - K_k \tilde{H}_{k-1}) \mathbb{E} (e_{k-1}^* v_{k-1}^*) (I - K_k \tilde{H}_{k-1})^* - (I - K_k \tilde{H}_{k-1}) \mathbb{E} (e_{k-1}^* e_{k-1}^*) (I - K_k \tilde{H}_{k-1})^*
\]

\[
= (I - K_k \tilde{H}_{k-1}) P_{k-1} (I - K_k \tilde{H}_{k-1})^* + K_k R_{k-1} K_k^* (14)
\]
where we have used that $\mathbb{E}(v_{k-1}e_{k-1}^*) = 0$ since $v_{k-1}$ and $e_{k-1}$ rise from different times. In fact, $v_{k-1}$ is the noise of the $(k-1)$-th observation $y(k-1)$ and $e_{k-1}$ is the error of the $(k-1)$-th estimation of the initial state which takes into account the first $(k-1)$ observations starting from 0, that is $\{y(0), y(1), \ldots, y(k-2)\}$.

Then, the linear functional that we have to minimize is given by

$$J(K_k) = \frac{1}{2} \text{Tr}(P_k)$$

$$= \frac{1}{2} \text{Tr} \left( (I - K_k \tilde{H}_k) P_{k-1} (I - K_k \tilde{H}_k)^* + K_k R_{k-1} K_k^* \right)$$

$$= \frac{1}{2} \text{Tr} (P_{k-1}) - \text{Tr} \left( K_k \tilde{H}_k P_{k-1} \right)$$

$$+ \frac{1}{2} \text{Tr} \left( K_k (\tilde{H}_k P_{k-1} \tilde{H}_k^* + R_{k-1}) K_k^* \right).$$

Imposing that

$$\frac{\partial J}{\partial K_k} = -(\tilde{H}_k P_{k-1})^* + K_k (\tilde{H}_k P_{k-1} \tilde{H}_k^* + R_{k-1}) = 0,$$

we obtain $K_k$ in (11).

Finally, replacing (11) in (14) we have

$$P_k = P_{k-1} + K_k (\tilde{H}_k P_{k-1} \tilde{H}_k^* + R_{k-1}) K_k^* - K_k \tilde{H}_k P_{k-1} - P_{k-1} \tilde{H}_k K_k$$

$$= P_{k-1} + P_{k-1} \tilde{H}_k^* K_k - K_k \tilde{H}_k P_{k-1} - P_{k-1} \tilde{H}_k K_k$$

$$= (I - K_k \tilde{H}_k) P_{k-1}.\quad (15)$$

\[\square\]

**Algorithm 1** Kalman-based algorithm to recover $x_0$

**Require:** $A \in \mathbb{R}^{d \times d}$, $H \in \mathbb{R}^{m \times d}$ the dynamic and observation operators, $R_k$ the covariances for the noise.

**Ensure:** $\hat{x}_0$ the initial state estimation after $T_f$ steps.

**Initialization:** $\hat{x}_0$, $P_0$ initial guesses for the initial state and covariance error matrix.

$\tilde{H}_0 = H$

for $k = 1$ to $T_f$ do

$K_k \leftarrow P_{k-1} \tilde{H}_k^* (\tilde{H}_k P_{k-1} \tilde{H}_k^* + R_{k-1})^{-1}$

$\hat{x}_k \leftarrow \hat{x}_{k-1} + K_k (y_k - \tilde{H}_k \hat{x}_{k-1})$

$P_k \leftarrow P_{k-1} - K_k \tilde{H}_k P_{k-1}$

$\tilde{H}_k \leftarrow \tilde{H}_k A$

end for

return $\hat{x}_{T_f}$
4 On the stability of the state estimation error dynamics

As can be deduced from (13) and (15), the dynamics for $\mathbb{E}(e_k)$ and for the matrices $P_k$ are the same. Indeed, since the noise has zero mean,

$$\mathbb{E}(e_k) = \Psi_k \mathbb{E}(e_{k-1}) + K_k \mathbb{E}(v_{k-1}) = \Psi_k \mathbb{E}(e_{k-1}) \quad \text{and} \quad P_k = \Psi_k P_{k-1}$$

where the transitions are given by the time-varying matrix

$$\Psi_k = I - K_k \tilde{H}_{k-1}$$ \hspace{1cm} (16)

with $K_k$ computed according to (11) in Theorem 2. In this section we look for the asymptotic stability of a linear time-varying system of the form

$$z(k) = \Psi_k z(k-1), \quad z(0) = z_0,$$ \hspace{1cm} (17)

with $\Psi_k$ as in (16), which corresponds to the dynamics of the expectation of the state dynamic error and of the covariance matrices of the proposed estimation method. The next proposition summarizes the properties of the transition matrices.

**Proposition 3** The matrices $\Psi_k$ satisfy the following:

(i) $\Psi_k = P_{k}P_{k-1}^{-1}$ for all $k \geq 1$.

(ii) $\Psi(k, j) = P_{k}P_{j}^{-1}$ for all $k \geq j$. Therefore, $\Psi(j, k) = P_{j}P_{k}^{-1}$.

**Proof** Item i) follows by (15). To prove ii) we simply write the transition matrix and use that $\Psi(k, i - 1) = \Psi_i$:

$$\Psi(k, j) = \Psi(k, k - 1)\Psi(k - 1, k - 2)\Psi(k - 2, k - 3) \cdots \Psi(j + 2, j + 1)\Psi(j + 1, j)$$

$$= P_{k}P_{k-1}^{-1}P_{k-1}P_{k-2}^{-1} \cdots P_{j+2}P_{j+1}^{-1}P_{j+1}P_{j}^{-1} = P_{k}P_{j}^{-1}.$$

Finally, by item i) $\Psi_k$ is invertible as it is the product of two non-singular matrices, therefore $\Psi_k^{-1} = P_{k-1}P_{k}^{-1}$. Then,

$$\Psi(j, k) = \Psi(k, j)^{-1} = P_{j}P_{k}^{-1}.$$ \hspace{1cm} $\square$

**Remark 2** It can be shown that

$$P_k \leq O(k, 0)^{-1}.$$ \hspace{1cm} (18)

Indeed, by considering the weighted least squares estimation $\tilde{x}_k$ given in (7), and using (5) we have

$$x_0 - \tilde{x}_k = x_0 - O(k, 0)^{-1} \sum_{j=0}^{k-1} \tilde{H}_j^* R_j y(j) = O(k, 0)^{-1} \sum_{j=0}^{k-1} \tilde{H}_j^* R_j v_j.$$
Thus, from $E(v_jv_j^*) = R_k\delta_{j,k}$ we obtain

$$E[(x_0 - \tilde{x}_k)(x_0 - \tilde{x}_k)^*] = \mathcal{O}(k, 0)^{-1}.$$  

Therefore, since $P_k$ is the covariance matrix of the error given by the linear estimator for the initial state that has minimum variance among all the estimators that consider the observations $\{y(n)\}$ given in (3), we have that $P_k$ is bounded from above by $\mathcal{O}(k, 0)^{-1}$, as stated in (18). Hence, if $\|\mathcal{O}(k, 0)^{-1}\| \to 0$ as $k \to \infty$, then $P_k$ approaches 0. Compare this with the proof of Lemma 7.1 in [7]. In what follows we will obtain a more precise relation between $P_k$ and $\mathcal{O}(k, 0)$. After that, in the next section, we will analyze necessary and sufficient conditions on the dynamic to guarantee that $\|P_k\| \to 0$ as $k$ increases, as well as stability of the method.

**Proposition 4** If the covariance matrix $P_0$ of the initial guess is strictly positive, then $P_k$ is upper bounded for all $k \geq 0$. Moreover,

$$P_0 \geq P_1 \geq P_2 \geq \cdots \geq P_k \geq P_{k+1} \geq \cdots$$  

(19)

and their inverses satisfy the relation

$$P_k^{-1} = P_0^{-1} + \mathcal{O}(k, 0).$$  

(20)

**Proof** Replacing (11) in (12) we have

$$P_k = P_{k-1} - P_{k-1} \tilde{H}_{k-1}^* (\tilde{H}_{k-1} P_{k-1} \tilde{H}_{k-1}^* + R_{k-1})^{-1} \tilde{H}_{k-1} P_{k-1}.$$

(21)

Take the matrix inverses of both sides of (21) while applying the matrix equality

$$(G + V^* W V)^{-1} = G^{-1} - G^{-1} V^* \left(W^{-1} + V G^{-1} V^* \right)^{-1} V G^{-1}$$

with

$$\begin{aligned}
G &= P_{k-1} \\
V &= \tilde{H}_{k-1} P_{k-1} \\
W &= - (\tilde{H}_{k-1} P_{k-1} \tilde{H}_{k-1}^* + R_{k-1})^{-1}
\end{aligned}$$

we obtain a recursion for $P_k^{-1}$:

$$P_k^{-1} = P_{k-1}^{-1} + \tilde{H}_{k-1}^* R_{k-1}^{-1} \tilde{H}_{k-1} \geq P_{k-1}^{-1}.$$  

From that recursion, we obtain (19). Moreover, it can be seen that

$$P_0^{-1} > 0 \quad \text{(by hypothesis)}$$

$$P_1^{-1} = P_0^{-1} + H^* R_0^{-1} H$$
and recursively we obtain

\[ P_k^{-1} = P_0^{-1} + \sum_{j=0}^{k-1} \tilde{H}_j^* R_j^{-1} \tilde{H}_j, \]

that is,

\[ P_k^{-1} = P_0^{-1} + O(k, 0) \geq P_0^{-1} \quad \text{for all } k \geq 0. \]

Finally, since \( P_0 \) is symmetric positive definite, we have that \( P_0 \leq \rho_0 I \) with \( \rho_0 = \| P_0 \|. \) Therefore,

\[ P_k^{-1} \geq \frac{1}{\rho_0} I \quad \text{and} \quad P_k \leq \rho_0 I. \]

\[ \square \]

### 4.1 Lyapunov stability

Lyapunov’s direct method \cite{19} is one of the most popular ways to show all kinds of stability for a given dynamical system. In our case, we will prove Lyapunov stability for the estimation error dynamics considering as a candidate for Lyapunov function \( V(k, z(k)) := z(k)^* P_k^{-1} z(k) \) with \( P_k \) the error covariance matrices described in the previous proposition.

**Theorem 5** Let \( \Psi_k \) as in (16). If the initial covariance matrix \( P_0 \) is strictly positive then the zero solution of the system

\[
\begin{align*}
z(k) &= \Psi_k z(k - 1) \\
z(0) &= z_0
\end{align*}
\]

is Lyapunov stable.

**Proof** To prove the stability it is enough to give a function \( V(k, z(k)) \) satisfying the following (see \cite[Theorem 13.11]{19}):

\[
\begin{cases}
V(k, 0) = 0 \\
V(k, z(k)) \geq \alpha(\| z(k) \|) \\
\Delta V(k, z) := V(k, z(k)) - V(k - 1, z(k - 1)) \leq 0
\end{cases}
\]

for some strictly increasing function \( \alpha : [0, \infty) \to [0, \infty) \).

Consider

\[ V(k, z(k)) = z(k)^* P_k^{-1} z(k) \quad \text{(22)} \]

as our Lyapunov function. By Proposition 4 we have

\[ V(k, z(k)) \geq \frac{1}{\rho_0} \| z(k) \|^2 \quad \forall k \geq 0. \]
Now,

\[ V(k, z(k)) = z(k)^* P_{k-1}^{-1} z(k) \]

\[ = z(k - 1)^* (\mathbb{I} - K_k \tilde{H}_{k-1})* P_{k-1}^{-1} (\mathbb{I} - K_k \tilde{H}_{k-1}) z(k - 1) \quad \text{(by (17))} \]

\[ = z(k - 1)^* (\mathbb{I} - K_k \tilde{H}_{k-1})* P_{k-1}^{-1} P_k P_{k-1}^{-1} z(k - 1) \quad \text{(by (15))} \]

\[ = V(k - 1, z(k - 1)) - z(k - 1)^* \tilde{H}_{k-1}^* K_k^* P_{k-1}^{-1} z(k - 1) \]

\[ = V(k - 1, z(k - 1)) - z(k - 1)^* \tilde{H}_{k-1}^* (\tilde{H}_{k-1} P_{k-1} \tilde{H}_{k-1}^* + R_{k-1})^{-1} \tilde{H}_{k-1} z(k - 1) \quad \text{(by (11)).} \]

Hence,

\[ \Delta V(k, z) = -z(k - 1)^* \tilde{H}_{k-1}^* \Sigma_{k-1}^{-1} \tilde{H}_{k-1} z(k - 1) \leq 0 \]

where

\[ \Sigma_{k-1} = \tilde{H}_{k-1} P_{k-1} \tilde{H}_{k-1}^* + R_{k-1}. \]

\[ \square \]

**Remark 3** Note that with the Lyapunov function given in (22) one cannot prove the asymptotic stability for the zero solution, one of the reasons being that the matrices \( \tilde{H}_{k}^* \Sigma_{k}^{-1} \tilde{H}_{k} \) are only positive semi-definite.

### 4.2 Asymptotic stability for the error dynamics of an LTI dynamical systems

From now on, we will consider the system (1) to be LTI. Before we state our main result about asymptotic stability, we introduce some notation and enunciate two important results concerning eigenvalues and singular values of a given matrix.

Given any \( B \in \mathbb{C}^{d \times d} \), let

\[ \lambda_{\text{max}}(B) = \max\{|\lambda|: \lambda \text{ eigenvalue of } B\}, \]

\[ \lambda_{\text{min}}(B) = \min\{|\lambda|: \lambda \text{ eigenvalue of } B\}. \]

Also, \( s_i(B) \) denotes the \( i \)-th singular value and \( s_{\text{min}}(B), s_{\text{max}}(B) \) denote the smallest and largest singular values, respectively. For \( B \in \mathbb{C}^{d \times d} \) Hermitian denote \( \{\lambda_i(B)\} \) the set of eigenvalues of \( B \) ordered by

\[ \lambda_1(B) \geq \lambda_2(B) \geq \cdots \geq \lambda_d(B). \]

If \( C \in \mathbb{C}^{d \times d} \) is also Hermitian then by Weyl’s Theorem [20] we have

\[ \lambda_i(B) + \lambda_d(C) \leq \lambda_i(B + C) \leq \lambda_i(B) + \lambda_1(C). \quad (23) \]

Finally, we cite a Gelfand type result for the asymptotic behavior of the singular values of the powers of a matrix.
Theorem 6 (cf. [21]) Let $A \in \mathbb{C}^{d \times d}$ be a matrix with eigenvalues $\lambda_k(A)$, $k = 1, \ldots, d$ ordered by non-increasing absolute values $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_d|$. For any $n \in \mathbb{N}$, denote $s_i(A^n)$ the singular values of $A^n$ also ordered non-increasingly $s_1(A^n) \geq s_2(A^n) \geq \cdots \geq s_d(A^n)$. Then

$$\lim_{n \to \infty} (s_i(A^n))^{1/n} = |\lambda_i|.$$ 

The following lemma describes the asymptotic behavior of the smallest eigenvalue of the observability matrix. More explicitly, we show that the non-decreasing sequence \{\min(O(k, 0))\}_{k \in \mathbb{N}} either goes to infinity or it is bounded, depending on the locus of the eigenvalues of the dynamic operator $A$.

Lemma 7 Let the system (1) be observable and with dynamics given by $A$. Then the following holds:

1. If $\min(A) > 1$, then

$$\lim_{k \to \infty} \min(O(k, 0)) = \infty.$$ 

2. If $\min(A) < 1$, then the sequence \{\min(O(k, 0))\}_{k \in \mathbb{N}} converges to a positive constant.

Proof (Part 1.) Let $L \in \mathbb{N}$ and $\rho > 0$ given by the observability condition as in Definition 1. We will show that the subsequence \{\min(O(nL, 0))\}_{n \in \mathbb{N}} goes to infinity. We have that

$$O(nL, 0) = \sum_{j=0}^{n-1} (A^*)^{jL} O((j + 1)L, jL) A^{jL} \geq \rho \sum_{j=0}^{n-1} (A^*)^{jL} A^{jL}$$

for all $n \in \mathbb{N}$.

Let us denote

$$O_{A,L}(n, 0) := \sum_{j=0}^{n-1} (A^*)^{jL} A^{jL}.$$ 

Therefore,

$$\min(O(nL, 0)) \geq \rho \min(O_{A,L}(n, 0)).$$

Then,

$$\min(O(nL, 0)) \geq \rho \sum_{j=0}^{n-1} \min((A^*)^{jL} A^{jL})$$

$$\geq \rho \min((A^*)^{(n-1)L} A^{(n-1)L}) = \rho s_{\min}^2(A^{(n-1)L})$$

where in the first inequality we have used Weyl’s inequality (23). From Theorem 6, we can take $\epsilon$ sufficiently small, with $\min(A) - \epsilon > 1$, such that for $n$ large enough,

$$s_{\min}^2(A^{(n-1)L}) > (\min(A) - \epsilon)^{2(n-1)}.$$
and we get the result.

Furthermore, for \( k \) sufficiently large we can write \( k = nL + m \) for \( m < L \). Then,

\[
\lambda_{\min}(O(k, 0)) \geq \lambda_{\min}(O(nL, 0)) \geq \rho e^{\beta k}
\]

with \( \beta \) some positive constant depending on \( L \) and \( \lambda_{\min}(A) \). This exponential order will be useful for our main result.

**Part 2.** Since \( O(k, 0) \) is symmetric and positive definite, for \( k \geq L \) we now that

\[
0 < \lambda_{\min}(O(k, 0)) = \min_{\|x\|=1} \langle O(k, 0)x, x \rangle
\]

Let \( v \) be a (normalized) eigenvector of \( A \) corresponding to the eigenvalue of minimum absolute value. Then,

\[
\lambda_{\min}(O(k, 0)) \leq \sum_{j=0}^{k-1} \langle (A^*)^j H^* R^j H A^j v, v \rangle \leq \frac{\|H\|^2}{\sigma^2} \sum_{j=0}^{k-1} \lambda_{\min}(A)^{2j} < \infty
\]

where \( \sigma^2 \) is the positive lower bound for the noise covariances given in Sect. 2.1. \( \square \)

**Remark 4** For the first part, notice that in the case of a normal operator \( A \) the result is rather straightforward and it also includes the case \( \lambda_{\min}(A) = 1 \). In fact, given the eigen-decomposition of \( A = U \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_d)U^* \) we have

\[
O_{A,L}(n, 0) = U \begin{bmatrix}
\sum_{j=0}^{n-1} |\lambda_1|^{2jL} & 0 & \ldots & 0 \\
0 & \sum_{j=0}^{n-1} |\lambda_2|^{2jL} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \sum_{j=0}^{n-1} |\lambda_d|^{2jL}
\end{bmatrix} U^*
\]

**Theorem 8** Consider the observable dynamical system (1). Let \( \Psi_k \) as in (16). If \( \lambda_{\min}(A) > 1 \), then the equilibrium point of

\[
\begin{cases}
\dot{z}(k) = \Psi_k z(k - 1) \\
z(0) = z_0
\end{cases}
\]

is uniformly asymptotically stable. On the other hand, if \( \lambda_{\min}(A) < 1 \) then the equilibrium point is not asymptotically stable.

**Proof** In order to prove the result let us observe the following. By Proposition 3 and equation (20) we have

\[
\Psi(k, k_0) = P_k P_{k_0}^{-1} \quad \forall k \geq k_0
\]

with

\[
P_n^{-1} = P_0^{-1} + O(n, 0) \quad \text{for every } n \in \mathbb{N}.
\]
Thus, by Proposition 1 it is enough to show that there exist $\alpha, \beta > 0$ such that
$$
\|\Psi(k, k_0)\| \leq \|P_k\| \|P_{k_0}^{-1}\| \leq \alpha e^{-\beta(k-k_0)} \quad \forall k \geq k_0 \geq 0.
$$

Since the matrices $P_k$ are symmetric positive definite, by Weyl’s Theorem mentioned before we have the handy inequality for $\|P_k\| = \lambda_{\text{max}}(P_k)$:
$$
\frac{1}{\lambda_{\text{max}}(P_0^{-1}) + \lambda_{\text{min}}(O(k, 0))} \leq \lambda_{\text{max}}(P_k) \leq \frac{1}{\lambda_{\text{min}}(P_0^{-1}) + \lambda_{\text{min}}(O(k, 0))}. \quad (25)
$$

For a fixed $k_0 \in \mathbb{N}$,
$$
\|P_{k_0}^{-1}\| \leq \|P_0^{-1}\| + \|O(k_0, 0)\| \leq \|P_0^{-1}\| + \frac{\|H^*H\|}{\sigma^2} \sum_{j=0}^{k_0} s_{\text{max}}(A)^{2j} \leq Ce^{\beta'k_0}
$$
with $C, \beta'$ positive constants depending on $\|H^*H\|, \|P_0^{-1}\|, \sigma$ and $s_{\text{max}}(A)$.

Now, from the right hand inequality in (25) and the exponential bound (24) obtained in Lemma 7 we get
$$
\|P_k\| \leq \frac{1}{\rho} e^{-\beta k}.
$$

Therefore,
$$
\|\Psi(k, k_0)\| \leq \alpha e^{-\beta(k-k_0)} \quad \forall k \geq k_0 \geq 0,
$$
with $\alpha, \beta$ positive constants depending on $L, \rho, \sigma^2, \lambda_{\text{min}}(A), s_{\text{max}}(A), \|H^*H\|$ and $\|P_0^{-1}\|$.

For the second assertion of the theorem observe that, due to the left hand inequality (25) and Lemma 7,
$$
\lim_{k \to \infty} \|P_k\| \neq 0.
$$

But, as it was stated before in (12), the matrices $P_k$ follow the dynamic $P_k = \Psi_k P_{k-1}$ with initial condition at time $k = 0$ given by $P_0$. This in turn implies, by Proposition 1, that $\lim_{k \to \infty} \|\Psi(k, 0)\| \neq 0$ which is equivalent to say that the equilibrium point is not asymptotically stable.

**Remark 5** 1. If $\lambda_{\text{max}}(A) < 1$, it can be easily seen that the equilibrium point of the system (17) is uniformly stable.

2. If the dynamics operator $A$ is normal, then $\lambda_{\text{min}}(A) \geq 1$ is a necessary and sufficient condition for the uniform asymptotic stability of the estimation error dynamics.

The next result enlightens us on the importance of the previous theorem. First, we mention a very agreeable property of a given estimator.

An estimator is said to be *consistent* if, roughly speaking, has the property that as the number of observations (data) increases indefinitely, the resulting sequence
of estimates converges in probability to the quantity to be estimated. A sufficient criterion for an estimator to be consistent is that its mean squared error converges to 0. The estimator is said to be asymptotically unbiased if its mean converges to the true quantity to be estimated, i.e., \( \lim_{k \to \infty} \| E(e_k) \| = 0 \). For a full exposition on the subject see [22].

Theorem 8 asserts that for an LTI observable dynamical system with dynamic operator \( A \) having all its eigenvalues outside the unit disk, the zero solution of the time varying system (17) is asymptotically stable. As it was already mentioned at the beginning of Sect. 4, this dynamical system is in fact the dynamical system of the expected value of the error estimation \( E(e_k) \). Then, under the aforementioned hypotheses we have that

\[
\lim_{k \to \infty} \| E(e_k) \| = \lim_{k \to \infty} \| P_k \| = 0.
\]

Recall that the matrices \( P_k \) given in (10) satisfy

\[
P_k = E(e_k e_k^*) - E(e_k)E(e_k)^*.
\]

Then, by taking traces in this expression we get

\[
\text{Tr} P_k = E(\| e_k \|^2) - \| E(e_k) \|^2
\]  

(26)

where we have used the identities

\[
\text{Tr}(E(e_k e_k^*)) = E(e_k^* e_k) = E(\| e_k \|^2)
\]

\[
\text{Tr}(E(e_k)E(e_k)^*) = E(e_k)^* E(e_k) = \| E(e_k) \|^2.
\]

The first term in the right hand side of (26) is called the mean squared error for the estimator \( \hat{x}_k \). In view of the above discussion, we have the following result.

**Corollary 9**  For the observable dynamical system (1) with \( \lambda_{\min}(A) > 1 \), Algorithm 1 gives an asymptotically unbiased and consistent estimator.

### 4.3 Numerical examples

In this section we give two numerical examples to depict the stability properties discussed above. In Example 1 we consider a dynamics satisfying the hypotheses of Theorem 8. In Example 2 we have a dynamical system for which the error dynamics will be only Lyapunov stable.

#### 4.3.1 Example 1

Consider the LTI system

\[
\begin{aligned}
x(k+1) &= Ax(k) \\
y(k) &= Hx(k) + v_k
\end{aligned}
\]
(a) True and estimated coordinates of $x_0$.

(b) Variance and MSE.

(c) Trajectory of the four eigenvalues of the matrices $P_k$.

**Fig. 1** Error analysis for Example 1

with

$$A = \begin{bmatrix} 1.99 & -0.32 & 0 & 0.07 \\ 0.43 & 1.17 & 0.02 & 0 \\ 0.13 & -0.09 & 1.52 & -0.13 \\ 0.28 & -0.14 & 0.03 & 1.22 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

and $v_k \sim N(0, \sigma I)$ for $\sigma = 0.01$. The true initial state is $x_0 = [0.2, 0.4, 0.5, 0.3]$.

Consider initial conditions $\tilde{x}_0 = [0.376, 0.502, 0.421, 0.366]$, initial error covariance $P_0 = 10^{-2} I$.

Figure 1a illustrates the plots of original signal (true signal) and its estimations after 5, 10, 40 time steps. Figure 1b shows the evolution of the mean squared error $E(e_k^2)$ for 40 time steps. Figure 1c shows the evolution for the four eigenvalues of the covariance matrix $P_k$ at every instant $k$ during 40 time steps.

**4.3.2 Example 2**

Consider the LTI system

$$\begin{align*}
x(k+1) &= Ax(k) \\
y(k) &= Hx(k) + v_k
\end{align*}$$
(a) True and estimated coordinates of $x_0$.

(b) Variance and MSE.

**Fig. 2** Error analysis for Example 2

with

$$A = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}, \quad H = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

and $v_k \sim N(0, \sigma I)$ for $\sigma = 0.001$. The true initial state is $x_0 = [0.83053274, 0.35472554]$. The initialization for the algorithm is $\hat{x}_0 = [0.9065169, 0.1988222]$, and initial error covariance $P_0 = 10^{-2}I$.

Figure 2a illustrates the plots of original signal (true signal) and its estimations after 2, 5, 20 time steps. Figure 2b shows the evolution of the mean squared error $E(e_k^2)$ for 40 time steps.

**Funding** The research for this article has mainly been carried out while R. Díaz Martín was holding a postdoctoral fellowship from CONICET, and then finished as a Postdoctoral Research Scholar at Vanderbilt University. Besides, I. Medri was holding postdoctoral fellowships from Vanderbilt University.

**Declarations**

**Conflict of interest**: On behalf of all authors, the corresponding author states that there is no conflict of interest.

**References**

1. Aldroubi, A., Krishtal, I.: Krylov subspace methods in dynamical sampling. Sampl. Theory Signal Image Process. **15**, 9–20 (2016). https://doi.org/10.1007/BF03549595
2. Aldroubi, A., Cabrelli, C., Cakmak, A.F., Molter, U., Petrosyan, A.: Iterative actions of normal operators. J. Funct. Anal. **272**(3), 1121–1146 (2017). https://doi.org/10.1016/j.jfa.2016.10.027
3. Cheng, J., Tang, S.: Estimate the spectrum of affine dynamical systems from partial observations of a single trajectory data. Inverse Prob. **38**(1), 015004 (2021)
4. Díaz Martín, R., Medri, I., Molter, U.: Dynamical sampling: a view from control theory. In: Excursions in Harmonic Analysis, Volume 6, pp. 269–295. Birkhäuser, Cham (2021). https://doi.org/10.1007/978-3-030-69637-5_15
5. Tang, S.: System identification in dynamical sampling. Adv. Comput. Math. **43**, 555–580 (2017). https://doi.org/10.1007/s10444-016-9497-5
6. Bryson, A.E., Frazier, M.: Smoothing for linear and nonlinear dynamic systems. Proc. Optimum Systems Synthesis Conf. Wright-Patterson Air Force Base, Ohio. Technical Documentary Report ASD-TDR-63-119, 353–364 (1963)

7. Jazwinski, A.H.: Stochastic Processes and Filtering Theory. Courier Corporation, New York (2007)

8. Willman, W.: On the linear smoothing problem. IEEE Trans. Autom. Control 14(1), 116–117 (1969). https://doi.org/10.1109/TAC.1969.1099108

9. Simon, D.: Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches. John Wiley & Sons, New Jersey (2006). https://doi.org/10.1002/0470045345

10. Meditch, J.S.: Orthogonal projection and discrete optimal linear smoothing. SIAM J. Control 5(1), 74–89 (1967). https://doi.org/10.1137/0305006

11. Zhang, Q.: On stability of the kalman filter for discrete time output error systems. Syst. Control Lett. 107, 84–91 (2017). https://doi.org/10.1016/j.sysconle.2017.07.011

12. Cabrelli, C., Molter, U., Paternostro, V., Philipp, F.: Dynamical sampling on finite index sets. J. d’analyse Math. 140(2), 637–667 (2020). https://doi.org/10.1007/s11854-020-0099-2

13. Aldroubi, A., Davis, J., Krishtal, I.: Dynamical sampling: Time-space trade-off. Appl. Comput. Harmon. Anal. 34(3), 495–503 (2013). https://doi.org/10.1016/j.acha.2012.09.002

14. Cabrelli, C., Molter, U., Suárez, D.: Multi-orbital frames through model spaces. Compl. Anal. Oper. Theory 15(1), 1–22 (2021). https://doi.org/10.1007/s11785-020-01063-6

15. Martín, R.D., Medri, I., Molter, U.: Continuous and discrete dynamical sampling. J. Math. Anal. Appl. 499(2), 125060 (2021). https://doi.org/10.1016/j.jmaa.2021.125060

16. Aldroubi, A., Huang, L., Krishtal, I., Lederman, R.: Dynamical sampling with additive random noise. STSIP 17, 153–182 (2018). https://doi.org/10.1007/BF03549662

17. Cadzow, J.A.: Signal enhancement-a composite property mapping algorithm. IEEE Trans. Acoust. Speech Signal Process. 36(1), 49–62 (1988). https://doi.org/10.1109/29.14988

18. Bof, N., Carli, R., Schenato, L.: Lyapunov theory for discrete time systems. arXiv preprint arXiv:1809.05289 (2018)

19. Chellaboina, V., Haddad, W.M.: Nonlinear Dynamical Systems and Control: A Lyapunov-based Approach. Princeton University Press, New Jersey (2008). https://doi.org/10.2307/j.ctvcm4hws

20. Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, New York (2012)

21. Yamamoto, T.: On the extreme values of the roots of matrices. J. Math. Soc. Japan 19(2), 173–178 (1967). https://doi.org/10.2969/jmsj/01920173

22. Casella, G., Berger, R.L.: Statistical Inference. Cengage Learning, Pacific Grove (2021)

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.