# Filtering discrete-time systems with multiplicative noise in $L_2$ spaces with applications

Filippo Cacace, Massimiliano d'Angelo, Vittorio De Iuliis, and Alfredo Germani

***Abstract*—In this work we adopt a novel formulation of the distributed parameters recursive filter for discrete-time systems evolving in $L_2$ spaces to widen the class of systems that can be processed by a state estimation algorithm. Starting from a rigorous definition of Kronecker algebra on $L_2$ spaces that involves both elements and bounded operators of $L_2$, we provide a computationally efficient solution in the case of linear systems with multiplicative noises. We illustrate the potential application of the approach by developing a case-study concerning the conceptual design of a distributed thermo-couple in the presence of the Nyquist–Johnson noise.**

**Keywords:** Distributed parameters systems; Kalman Filtering; Discrete-time systems

## I. Introduction

The optimal filtering problem for finite-dimensional linear systems affected by Gaussian noises was solved in the seminal work by Kalman [1] (discrete-time). The following decades witnessed an astonishing development of several approaches to the filtering problem of broader classes of dynamical systems [2], [3]. Distributed parameters or infinite-dimensional systems attracted much attention and found a early solution in the infinite-dimensional Kalman-Bucy filter proposed by Falb [4], that includes distributed parameter systems with bounded integral operators but it is not suitable for the more interesting case of partial differential operators. Most of the subsequent efforts on filtering infinite-dimensional systems that were developed in the following decades have been focusing on the case of continuous-time systems (see [5], [6], [7] and the references therein). We point out the recent work [8] which provide a high gain nonlinear observer for a class of deterministic quasi-linear hyperbolic systems with an application to an epidemic model. The same problem has been tackled by [9] with an application to flow control. Finally, another interesting approach is the one of [10], where a Luenberger-type boundary observer via backstepping argument is employed for a class of time-varying linear hyperbolic PIDEs. We notice that for the latter works, the state vector evolve in $\mathbb{R}$.

Instead, only a few recent works take into consideration the case of sampled-data measurements [11], [12], [13] that is indeed of great relevance in modern digital applications.

Filippo Cacace f.cacace@unicampus.it is with Università Campus Bio-medico di Roma, Italy.

Massimiliano d'Angelo massimiliano.dangelo@unimib.it is with the Department of Biotechnology and Biosciences (BtBs), University of Milano-Bicocca, Italy.

Vittorio De Iuliis vittorio.deiuliis and Alfredo Germani alfredo.germani@univaq.it are with the Dipartimento di Ingegneria e Scienze dell'Informazione e Matematica, Università degli Studi dell'Aquila, Italy.

This work considers linear distributed parameter stochastic systems with discrete-time dynamics generated by an integral bounded operator and discrete-time measurements. The aforementioned setting is well suited to the case when the integral operator is the semigroup generated by an underlying continuous-time dynamics and allows to filter continuous-discrete infinite-dimensional systems, *i.e.* systems with a continuous-time dynamics represented through a partial differential equation and sampled measurements.

This paper extends our previous work [14] to take into account nonlinear noise terms in the spirit of the Extended Kalman Filter. Although in this case the filter equations do not have a stationary form, we show that an efficient implementation of the filter is possible. In order to emphasize the potential impact of the proposed method we develop in detail the design of a distributed thermo-couple to estimate the temperature profile of a heated bar from discrete-time scalar measurements of the voltage difference at the endpoints when a current flows through it. The problem fits well in our framework due to the presence of the Nyquist–Johnson nonlinear noise term in the measurement equation ([15], [16]).

The work is structured as follows. Section II provides the problem formulation and presents some mathematical background on the Kronecker algebra in $L_2$ spaces. Section III presents the estimation theory on Hilbert spaces and introduces the concept of optimal linear estimate through conditional expectations and orthogonal projections, providing a discussion and some insights to address the non-Gaussian estimation problem. Section IV describes the algorithm amenable to cope with nonlinear noise terms. Finally, in Section V we illustrate the method on the problem of estimating the temperature profile of a heated bar from measurements of the voltage difference at the endpoints.

*Notation:* We denote $\mathcal{H}^n = L_2(\Theta;\mathbb{R}^n)$ the Hilbert spaces of square integrable functions with respect to the Lebesgue measure from a domain $\Theta$ to $\mathbb{R}^n$. The notation $\mathcal{H}^i$ will also be used for the Hilbert space of functions in $\mathbb{R}^{n_i}$, $n_i$ positive integer. Thus, if $v \in \mathcal{H}^n$, $v(\xi) \in \mathbb{R}^n$ for every $\xi \in \Theta$. $\|v\|^2_{\mathcal{H}} = \int_\Theta \|v(\xi)\|^2 d\xi < \infty$ is the squared norm of $v$, where $\|v(\xi)\|$ is the standard Euclidean norm. Given a bounded linear operator $A : \mathcal{H}^1 \to \mathcal{H}^2$, we indicate with $\kappa^A : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^{n_2 \times n_1}$ the kernel of $A$. Given a probability space $(\Omega, \mathcal{F}, P)$, $\mathcal{W}^{\mathcal{F}}_{\mathcal{H}} = L_2(\Omega;\mathcal{F};\mathcal{H};P)$ is the Hilbert space of square integrable $\mathcal{F}$-measurable functions with values in $\mathcal{H}$. $\mathbb{E}[X] \in \mathcal{H}$ denotes the expected value of $X \in \mathcal{W}^{\mathcal{F}}_{H}$, $\mathbb{E}[X] = \int_\Omega X(\omega)\,\mathrm{d}P$. For $X, Y \in \mathcal{W}^{\mathcal{F}}_{\mathcal{H}}$, $[X,Y]_{\mathcal{W}^{\mathcal{F}}_{\mathcal{H}}} = \int_\Omega [X(\omega), Y(\omega)]_{\mathcal{H}}\,\mathrm{d}P = \mathbb{E}[X,Y]_{\mathcal{H}}$. $I_{\mathcal{H}}$ is the identity operator in $\mathcal{H}$ and $I_n$ the identity matrix in $\mathbb{R}^n$. For

$M \in \mathbb{R}^{m\times n}$, $\mathrm{st}(M) \in \mathbb{R}^{mn}$ denotes the vertical stack of $M$, and, for $v \in \mathbb{R}^{mn}$, $\mathrm{st}_m^{-1}(v) \in \mathbb{R}^{m\times n}$ is the inverse operation. Given two vectors or matrices $A$ and $B$ we denote with $A\otimes B$ their Kronecker product and with $A^{[2]} = A\otimes A$ the Kronecker square. Two well known properties used in the paper, that and hold for matrices or vectors of suitable dimensions, are

$$(AB)\otimes(CD) = (A\otimes C)(B\otimes D) \tag{1}$$
$$(A\otimes B)\,\mathrm{st}(C) = \mathrm{st}(BCA^\top). \tag{2}$$

## II. Problem Formulation and Preliminaries

On a probability space $(\Omega, \mathcal{F}, P)$, we address the state estimation problem for infinite-dimensional discrete-time systems, described by:

$$x_{k+1}(\omega) = Ax_k(\omega) + f(x_k(\omega))n_k(\omega) \tag{3}$$
$$y_k(\omega) = Cx_k(\omega) + g(x_k(\omega))n_k(\omega), \tag{4}$$

where the state $x_k \in \mathcal{W}^{\mathcal{F}}_{\mathcal{H}^n}$, $k \geq 0$, that is, $x_k(\omega) \in \mathcal{H}^n$ and $x_k(.)(\xi) \in \mathcal{W}^{\mathcal{F}}_{\mathbb{R}^n}$. $\Theta$ is a compact subset of $\mathbb{R}^m$. The measurement variable $y_k$ is in the Hilbert space $\mathcal{W}^{\mathcal{F}}_{\mathcal{H}^q}$. The stochastic sequence $n_k \in \mathcal{W}^{\mathcal{F}}_{\mathbb{R}^p}$ is zero-mean and white with covariance matrix $\mathbb{E}[n_k n_k^\top] = I_p$. $A : \mathcal{H}^n \to \mathcal{H}^n$ and $C : \mathcal{H}^n \longrightarrow \mathcal{H}^q$ are bounded linear operators such that

$$(A\varphi)(\xi) = \int_\Theta \kappa^A(\xi,\zeta)\varphi(\zeta)\mathrm{d}\zeta, \quad \text{with } \varphi \in \mathcal{H}^n,\ \xi \in \Theta$$
$$(C\varphi)(\xi) = \int_\Theta \kappa^C(\xi,\zeta)\varphi(\zeta)\mathrm{d}\zeta, \quad \text{with } \varphi \in \mathcal{H}^n,\, \xi \in \Theta,$$

where the kernels $\kappa^A : \mathbb{R}^m\times\mathbb{R}^m \to \mathbb{R}^{n\times n}$, $\kappa^C : \mathbb{R}^m\times\mathbb{R}^m \to \mathbb{R}^{q\times n}$ satisfy

$$\int_{\Theta\times\Theta} \|\kappa^A(\xi,\zeta)\|^2 \mathrm{d}\xi\mathrm{d}\zeta < \infty$$
$$\int_{\Theta\times\Theta} \|\kappa^C(\xi,\zeta)\|^2 \mathrm{d}\xi\mathrm{d}\zeta < \infty.$$

Moreover, $f : \mathcal{H}^n \to \mathcal{H}^{n\times p}$, $\mathcal{H}^{n\times p} = L_2(\Theta; R^{n\times p})$ and $g : \mathcal{H}^n \to \mathcal{H}^{q\times p}$ are known bounded operators. The state noise and the output noise are uncorrelated, namely for any $z, w \in \mathcal{W}^{\mathcal{F}}_{\mathcal{H}^n}$, we have $f(z(.)(\xi))g^\top(w(.)(\zeta)) = 0$, $\forall \xi,\zeta \in \Theta$.

The aim of the optimal filtering problem is to find an estimate of the state $x_k$, optimal with respect to some criterion, given the available measurement $y_0, \ldots, y_k$. In this paper, we provide an estimation algorithm by using the formalism of the recent paper [14] with an approximation in the spirit of the well-known Extended Kalman Filter (EKF). We shall clarify this point later in Section IV.

### A. Preliminary definitions and results on $L_2$ spaces

In this section we introduce a few definitions and results on the Kronecker algebra in $L_2$ spaces that will be instrumental in the remainder of the work.

*Definition 1:* Given $\mathcal{H}^i, \mathcal{H}^j$ with bases $\{e_i\}, \{e_j\}$ let $x \in \mathcal{H}^i$ and $y \in \mathcal{H}^j$. We define $x \boxtimes y$ the element in $\mathcal{H}^{i\otimes j}$ such that

$$(x\boxtimes y)(\xi,\zeta) = x(\xi)\otimes y(\zeta), \tag{5}$$

where $x(\xi)\otimes y(\zeta)$ is the standard Kronecker product and

$$\mathcal{H}^{i\otimes j} = \Big\{ z = \sum_{i,j} a_{ij}(e_i\otimes e_j) \in L_2(\Theta^2; \mathbb{R}^{n_i n_j}) \Big\} \tag{6}$$

such that $\sum_{i,j} a_{ij}^2 < \infty$. □

For $x \in \mathcal{W}^{\mathcal{F}}_{\mathcal{H}^i}$ and $y \in \mathcal{W}^{\mathcal{F}}_{\mathcal{H}^j}$, $x\boxtimes y$ is the element in $L_2(\Omega; \mathcal{F}; \mathcal{H}^{i\otimes j}; P)$ such that

$$(x\boxtimes y)(\omega,\xi,\zeta) = x(\omega,\xi)\otimes y(\omega,\zeta). \tag{7}$$

The inner product in $\mathcal{H}^{i\otimes j}$ enjoys special properties.

*Lemma 1:* If $x, y \in \mathcal{H}^{i\otimes j}$, $x = x_1\boxtimes x_2$, $y = y_1\boxtimes y_2$ then

$$[x,\, y]_{\mathcal{H}^{i\otimes j}} = [x_1,\, y_1]_{\mathcal{H}^i} \cdot [x_2,\, y_2]_{\mathcal{H}^j}. \tag{8}$$

Thus, $\|x\|^2_{\mathcal{H}^{1\otimes 2}} = \|x_1\|^2_{\mathcal{H}^1}\|x_2\|^2_{\mathcal{H}^2}$.

Lemma 1 implies that the norm of $x\boxtimes y$ is finite whenever $x$ and $y$ have finite norm. For $x \in \mathcal{H}_n$ we define $x^{\boxed{2}} \in \mathcal{H}^{n\otimes n}$

$$x^{\boxed{2}}(\xi,\zeta) = (x\boxtimes x)(\xi,\zeta) = x(\xi)\otimes x(\zeta). \tag{9}$$

The operation "$\boxtimes$" is bilinear and associative. In order to derive further properties we introduce an analogue tensor product among bounded linear operators, that we still denote by $\boxtimes$ (the appropriate product shall be clear from the context).

*Definition 2:* Let $M : \mathcal{H}^1 \to \mathcal{H}^2$ and $N : \mathcal{H}^3 \to \mathcal{H}^4$ be two bounded linear operators with kernels $\kappa^M : \mathbb{R}^m\times\mathbb{R}^m \to \mathbb{R}^{n_2\times n_1}$ and $\kappa^N : \mathbb{R}^m\times\mathbb{R}^m \to \mathbb{R}^{n_4\times n_3}$, respectively. Then, the operator $M\boxtimes N : \mathcal{H}^{1\otimes 3} \to \mathcal{H}^{2\otimes 4}$ is defined as

$$\forall z \in \mathcal{H}^{1\otimes 3}: \quad \left((M\boxtimes N)z\right)(\xi,\zeta) = \int_{\Theta\times\Theta} \left(\kappa^M(\xi,s)\otimes\kappa^N(\zeta,t)\right) z(s,t)\mathrm{d}s\,\mathrm{d}t. \tag{10}$$

*Lemma 2:* Let $x \in \mathcal{H}^1$, $y \in \mathcal{H}^3$ and $M : \mathcal{H}^1 \to \mathcal{H}^2$ and $N : \mathcal{H}^3 \to \mathcal{H}^4$ be two bounded linear operators with kernels $\kappa^M$ and $\kappa^N$ respectively. Then,

$$(Mx)\boxtimes(Ny) = (M\boxtimes N)(x\boxtimes y). \tag{11}$$

*Lemma 3:* Let $\mathcal{H}^i$ be Hilbert spaces for $i = 1,2,3$, $M : \mathcal{H}^2 \to \mathcal{H}^3$ and $N : \mathcal{H}^1 \to \mathcal{H}^2$ be two bounded linear operators with kernel $\kappa^M$ and $\kappa^N$, respectively. Then, the *composition* $M\circ N :\ \mathcal{H}^1 \to \mathcal{H}^3$ is given by

$$\big((M\circ N)x\big)(\xi) = \int_\Theta \kappa^{MN}(\xi,\eta)x(\eta)\mathrm{d}\eta \tag{12}$$

with

$$\kappa^{MN}(\xi,\eta) \triangleq \int_\Theta \kappa^M(\xi,s)\kappa^N(s,\eta)\mathrm{d}s \tag{13}$$

*Lemma 4:* Let $M : \mathcal{H}^2 \to \mathcal{H}^3$, $N : \mathcal{H}^1 \to \mathcal{H}^2$, $P : \mathcal{H}^5 \to \mathcal{H}^6$, and $Q : \mathcal{H}^4 \to \mathcal{H}^5$ be bounded linear operators with kernels $\kappa^M$, $\kappa^N$, $\kappa^P$, $\kappa^Q$ respectively. Then,

$$(M\circ N)\boxtimes(P\circ Q) = (M\boxtimes P)\circ(N\boxtimes Q). \tag{14}$$

We are interested to Kronecker product between linear operators when one of the operators is the identity. The following result descends immediately from the definitions above.

*Corollary 1:* Let $M : \mathcal{H}^1 \to \mathcal{H}^2$ and $I_{\mathcal{H}^2}$ be the identity operator in $\mathcal{H}^2$. Then, $\forall x \in \mathcal{H}^1$, $y \in \mathcal{H}^2$,

$$(M\boxtimes I_{\mathcal{H}^2})(x\boxtimes y) = (Mx)\boxtimes y, \tag{15}$$
$$(I_{\mathcal{H}^2}\boxtimes M)(y\boxtimes x) = y\boxtimes(Mx) \tag{16}$$

Finally, we have the following definition[1]

*Definition 3:* Let $M : \mathcal{H}^1 \to \mathcal{H}^2$ and $I_{\mathcal{H}^2}$ the identity operator in $\mathcal{H}^2$. Then, $\forall z \in \mathcal{H}^{1\otimes 2}$,

$$((M \boxtimes I_{\mathcal{H}^2})z)(\xi,\zeta) \triangleq \int_\Theta (\kappa^M(\xi,s)\otimes I_{n_2})z(s,\zeta)\,\mathrm{d}s \quad (17)$$

$$((I_{\mathcal{H}^2} \boxtimes M)z)(\xi,\zeta) \triangleq \int_\Theta (I_{n_2}\otimes \kappa^M(\zeta,s))z(\xi,s)\,\mathrm{d}s. \quad (18)$$

Definitions 1, 2 can obviously be used when one of the Hilbert spaces is replaced by the Euclidean space $\mathbb{R}^n$. This is useful to express particular cases of (3)–(4), for example when the available output is finite-dimensional.

## III. Estimation theory on Hilbert spaces

### A. Covariance vectors, uncorrelated vectors and projections

We apply the framework introduced above to define the covariance vectors of the noise terms. In order to provide an intuitive analogy with the linear time-invariant finite-dimensional case where $F \in \mathbb{R}^{n\times p}$ we recall that, when $\mathbb{E}[n_k] = 0$, the covariance of the state noise, $\mathbb{E}[(Fn_k)(Fn_n)^\top] = FF^\top$, can be represented in an equivalent way as $\mathbb{E}[(Fn_k)^{[2]}] = F^{[2]}\cdot \mathrm{st}(I_p) = \mathrm{st}(FF^\top)$. In the infinite-dimensional case with multiplicative noise, let $F_k(\omega) := f(x_k(\omega)) \in \mathcal{H}^{n\times p}$. Clearly, $F_k \in \mathcal{W}^{\mathcal{F}}_{\mathcal{H}^{n\times p}}$ and $F_k(\omega)n_k(\omega) \in \mathcal{H}^n$. Thus, $\mathbb{E}[(F_kn_k)^{\boxed{2}}] \in \mathcal{H}^{n\otimes n}$. From Lemma 2 we obtain

$$\mathbb{E}[(F_kn_k)^{\boxed{2}}] = \mathbb{E}[(F_k\boxtimes F_k)n_k^{[2]}] = \bar{F}_k^{\boxed{2}}\,\mathrm{st}(I_p), \quad (19)$$

where $\bar{F}_k^{\boxed{2}} = \mathbb{E}[(F_k\boxtimes F_k)]$ . The elements $\bar{F}_k^{\boxed{2}}\,\mathrm{st}(I_p) \in \mathcal{H}^{n\otimes n}$ and $\bar{G}_k^{\boxed{2}}\,\mathrm{st}(I_p) \in \mathcal{H}^{q\otimes q}$, with $\bar{G}_k^{\boxed{2}} = \mathbb{E}[(G_k \boxtimes G_k)]$, are (respectively) the state and measurement noise covariance vectors. For concision we denote them as $Q_k = \bar{F}_k^{\boxed{2}}\,\mathrm{st}(I_p)$ and $R_k = \bar{G}_k^{\boxed{2}}\,\mathrm{st}(I_p)$, with

$$Q_k(\xi,\zeta) = \mathbb{E}[(F_k(\omega)(\xi)\otimes F_k(\omega)(\zeta))]\,\mathrm{st}(I_p), \quad (20)$$
$$R_k(\xi,\zeta) = \mathbb{E}[(G_k(\omega)(\xi)\otimes G_k(\omega)(\zeta))]\,\mathrm{st}(I_p). \quad (21)$$

The operation "$\boxtimes$" can be used to define uncorrelation in $\mathcal{W}^{\mathcal{F}}_{\mathcal{H}}$. We recall another definition and a lemma of [14].

*Definition 4:* Given $x \in \mathcal{W}^{\mathcal{F}}_{\mathcal{H}^1}$ and $y \in \mathcal{W}^{\mathcal{F}}_{\mathcal{H}^2}$, we say that $x$ and $y$ are *uncorrelated* and write $x \perp y$ when

$$\mathbb{E}[x\boxtimes y] = 0. \quad (22)$$

Clearly, when $x\perp y$, $\mathbb{E}[x(\omega)(\xi)\otimes y(\omega)(\zeta)] = 0$, $\forall \xi,\zeta\in\Theta$. Notice that, in analogy with the finite-dimensional case, when $\mathcal{H}^1 = \mathcal{H}^2 = \mathcal{H}^n$ then $x\perp y \;\Rightarrow [x,y]_{\mathcal{W}^{\mathcal{F}}_{\mathcal{H}^n}} = 0$, but the converse is true only when $n = 1$.

*Lemma 5:* If $x\in \mathcal{W}^{\mathcal{F}}_{H_1}$, $y \in \mathcal{W}^{\mathcal{F}}_{H_2}$ and $x\perp y$, then $Mx \perp Ny$ for any bounded linear operators $M$, $N$.

We can use Definition 4 to express in an alternative way that state and measurement noise in (3)–(4) are uncorrelated.

*Corollary 2:* The noises $F_k(\omega)n_k(\omega)$, $G_k(\omega)n_k(\omega)$ are uncorrelated if and only if $\mathbb{E}[(F_k(\omega)\boxtimes G_k(\omega))]\,\mathrm{st}(I_p) = 0$.

*Proof.* The proof is readily obtained since

$$\begin{aligned}\mathbb{E}[(F_kn_k)\boxtimes(G_kn_k)] &= \mathbb{E}[(F_k\boxtimes G_k)]\mathbb{E}[n_k^{[2]}]\\ &= \mathbb{E}[(F_k\boxtimes G_k)]\,\mathrm{st}(I_p) = 0. \end{aligned}\quad (23)$$

In our case the structure of the functions $f$ and $g$ ensures that (23) always holds true. Finally, we recall another definition of [14] on the projection onto subspaces.

*Definition 5:* Given $x\in\mathcal{W}^{\mathcal{F}}_{\mathcal{H}}$ and a subspace $\mathcal{W}^{\mathcal{F}^1}_{\mathcal{H}}$ of $\mathcal{W}^{\mathcal{F}}_{\mathcal{H}}$, the projection of $x$ on $\mathcal{W}^{\mathcal{F}^1}_{\mathcal{H}}$, denoted $\Pi[x|\mathcal{W}^{\mathcal{F}^1}_{\mathcal{H}}]$, is the unique vector $z\in\mathcal{W}^{\mathcal{F}^1}_{\mathcal{H}}$ such that $x - z\perp w$, $\forall w\in \mathcal{W}^{\mathcal{F}^1}_{\mathcal{H}}$.

As remarked above, it follows that $[z, x{-}z]_{\mathcal{W}^{\mathcal{F}}_{H}} = 0$. Moreover, if $\forall y\in\mathcal{W}^{\mathcal{F}}_{\mathcal{H}}$, $x\perp y$ we write $x\perp\mathcal{W}^{\mathcal{F}}_{\mathcal{H}}$ and it follows from the uniqueness of the projection that $\Pi[x|\mathcal{W}^{\mathcal{F}}_{\mathcal{H}}] = 0$.

### B. Optimal linear estimate

The optimal estimate $\hat{x}_k$ of $x_k \in \mathcal{W}^{\mathcal{F}}_{\mathcal{H}^n}$ in the minimum variance sense is the conditional expectation with respect to the $\sigma$-algebra $\mathcal{F}^y_k$ generated by the output sequence $Y^k = \{y_0,\dots,y_k\}$, that is, $\hat{x}_k = \mathbb{E}[x_k|\mathcal{F}^y_k]$. The conditional expectation can be equivalently expressed as the *projection* of $x_k \in \mathcal{W}^{\mathcal{F}}_{\mathcal{H}^n}$ on the the subspace of the $\mathcal{F}^y_k$-measurable functions $\Omega\to\mathcal{H}^n$, that is, $\hat{x}_k = \Pi[x_k|\mathcal{W}^{\mathcal{F}^y_k}_{\mathcal{H}^n}]$.

In the hypothesis that the sequences $x_k$ and $y_k$ are jointly Gaussian, the projection of $x_k$ onto $\mathcal{W}^{\mathcal{F}^y_k}_{\mathcal{H}^n}$ is identical to the projection on the space $\mathcal{L}^y_k$ of the affine functions of $Y^k$ having value in $\mathcal{W}^{\mathcal{F}^y_k}_{\mathcal{H}^n}$.

$$\mathcal{L}^y_k = \{z\in\mathcal{W}^{\mathcal{F}^y_k}_{\mathcal{H}^n} :\; z = \sum_{i=0}^k K_iy_i + N,\; N\in\mathcal{H}^n\}, \quad (24)$$

where $K_i : \mathcal{H}^q\to\mathcal{H}^n$ are bounded linear operators. We can then write

$$\hat{x}_k = \Pi[x_k|\mathcal{L}^y_k]. \quad (25)$$

When the sequences $x_k$ or $y_k$ are not Gaussian, the conditional expectation $\mathbb{E}[x_k|\mathcal{F}^y_k]$ does not coincide with (25), that however provides the best linear unbiased estimator in the minimum variance sense. In our case, both the state and output processes are allowed to be non-Gaussian because of the nonlinear map $f$ and $g$. Moreover, even in the finite-dimensional case, *i.e.* $x_k\in\mathbb{R}^n$, in the case of non-Gaussian noises, the conditional expectation is the solution of an infinite dimensional problem ([17]) that can be solved by numerical approximate solutions with high computational burden. Thus, it is essential to provide feasible approximate solutions. For these reasons, we shall provide an approximation for the solution of (25) in the spirit of the EKF.

## IV. Algorithm for the filter, the approximated covariance vectors and the gain kernel

The discrete-time Kalman filter for linear systems on $L_2$ spaces is derived in [14]. Here we propose a sub-optimal filter for systems with nonlinear noise terms which enlarges consistently the class of systems that can be processed without increasing the complexity. We summarize the algorithm for the computation of the approximated covariance vectors and

[1](17)–(18) do not descend from Corollary 1. They can be derived by representing the kernel of $I_{\mathcal{H}}$ as a Dirac function. The latter is however not in $L_2$. To maintain a simple $L_2$ framework we use Definition 3.

the kernel of the distributed Kalman gain. We consider the classical recursive Kalman-like structure

$$\hat{x}_k(\omega) = A\hat{x}_{k-1}(\omega) + K_k\left(y_k(\omega) - (C \circ A)\hat{x}_{k-1}(\omega)\right), \quad (26)$$

where $\hat{x}_k$ is the state estimate and $K_k : \mathcal{H}^q \to \mathcal{H}^n$ is the Kalman gain operator at time $k$. If $\Sigma_0(\xi,\zeta) \triangleq \mathbb{E}\left[x_0(\xi) \otimes x_0(\zeta)\right]$ is known, $K_k$ and the approximated covariance vectors can be computed by iteratively solving for $k \geq 0$

$$P_0 = \Sigma_0 \quad (27)$$

$$P^p_{k+1} = A^{\boxed{2}} P_k + \hat{Q}_k \quad (28)$$

$$P^o_{k+1} = C^{\boxed{2}} P^p_{k+1} + \hat{R}_k \quad (29)$$

$$\left(K_{k+1} \boxtimes I_{\mathcal{H}^q}\right) P^o_{k+1} = \left(I_{\mathcal{H}^n} \boxtimes C\right) P^p_{k+1} \quad (30)$$

$$P_{k+1} = P^p_{k+1} - \left((K_{k+1} \circ C) \boxtimes I_{\mathcal{H}^n}\right) P^p_{k+1}. \quad (31)$$

In these integral equations $\hat{Q}_k = \hat{F}_k^{\boxed{2}}\,\mathrm{st}(I_p)$, $\hat{R}_k = \hat{G}_k^{\boxed{2}}\,\mathrm{st}(I_p)$, $\hat{F}_k := f(\hat{x}_k(\omega)) \in \mathcal{H}^{n\times p}$, $\hat{G}_k := f(\hat{x}_{k-1}(\omega)) \in \mathcal{H}^{q\times p}$

$$\hat{Q}_k(\xi,\zeta) = \left(\hat{F}_k(\omega)(\xi) \otimes \hat{F}_k(\omega)(\zeta)\right)\mathrm{st}(I_p), \quad (32)$$

$$\hat{R}_k(\xi,\zeta) = \left(\hat{G}_k(\omega)(\xi) \otimes \hat{G}_k(\omega)(\zeta)\right)\mathrm{st}(I_p),. \quad (33)$$

This corresponds to computing for $k \geq 0$, $\forall(\xi,\zeta) \in \Theta\times\Theta$,

$$P^p_{k+1}(\xi,\zeta) = \int_{\Theta\times\Theta} (\kappa^A(\xi,s) \otimes \kappa^A(\zeta,t)) P_k(s,t)\mathrm{d}s\,\mathrm{d}t + \hat{Q}_k(\xi,\zeta) \quad (34)$$

$$P^o_{k+1}(\xi,\zeta) = \int_{\Theta\times\Theta} (\kappa^C(\xi,s) \otimes \kappa^C(\zeta,t)) P^p_{k+1}(s,t)\mathrm{d}s\,\mathrm{d}t + \hat{R}_k(\xi,\zeta) \quad (35)$$

$$\int_\Theta (\kappa^K_{k+1}(\xi,s) \otimes I_q) P^o_{k+1}(s,\zeta)\,\mathrm{d}s = = \int_\Theta (I_n \otimes \kappa^C(\zeta,s)) P^p_{k+1}(\xi,s)\,\mathrm{d}s \quad (36)$$

$$P_{k+1}(\xi,\zeta) = P^p_{k+1}(\xi,\zeta) - \int_{\Theta^2} (\kappa^K_{k+1}(\xi,t)\kappa^C(t,s)) \otimes I_n) P^p_{k+1}(s,\zeta)\,\mathrm{d}t\,\mathrm{d}s \quad (37)$$

In the above recursion the difficult step is solving for $\kappa^K_{k+1}$ the integral equation (36) at each $k$. We notice that, by using Definition 3, (2) and the symmetry of $\mathrm{st}_n^{-1}\left(P^p_{k+1}\right)$, equation (36) can be written $\forall(\xi,\zeta) \in \Theta\times\Theta$:

$$\int_\Theta \kappa^K_{k+1}(\xi,s)\,\mathrm{st}_q^{-1}\left(P^o_{k+1}(s,\zeta)\right)\,\mathrm{d}s = \int_\Theta \mathrm{st}_n^{-1}\left(P^p_{k+1}(\xi,t)\right)\kappa^C(\zeta,t)^\top\,\mathrm{d}t. \quad (38)$$

*The case of finite-dimensional output*

A common case is that the output is finite dimensional (see also theoretical motivations in [7]). In this section we obtain the closed-form solutions of (34)–(37) for this case, where $y_k \in \mathcal{W}^{\mathcal{F}}_{\mathbb{R}^q}$, $G_k \in \mathbb{R}^{q\times p}$ and $C$ is a bounded linear operator,

$$y_k(\omega) = \int_\Theta \kappa^C(s) x_k(\omega,s)\,\mathrm{d}s + G_k n_k(\omega), \quad (39)$$

with $\kappa^C(s) \in \mathbb{R}^{q\times n}$. We also notice that the filter gain $K_k : \mathbb{R}^q \to \mathcal{H}^n$ is such that, for $v \in \mathbb{R}^q$ and $\xi \in \Theta$,

$$(Kv)(\xi) = \int_\Theta \kappa^K(\xi,s)\,\mathrm{d}s \cdot v = K(\xi)v \quad (40)$$

with $K(\xi) \in \mathbb{R}^{n\times q}$. Although equation (34) is the same as in the general case, equation (35) for the output innovation covariance becomes

$$P^o_{k+1} = \int_{\Theta\times\Theta} \left(\kappa^C(s) \otimes \kappa^C(t)\right) P^p_{k+1}(s,t)\mathrm{d}s\,\mathrm{d}t + \hat{R}_k. \quad (41)$$

Thus, by letting

$$\widetilde{P}^p_k(\xi,\zeta) \triangleq \mathrm{st}_n^{-1}\left(P^p_k(\xi,\zeta)\right) \quad (42)$$

we finally have, from (38), (41)

$$\mathrm{st}_q^{-1}\left(P^o_{k+1}\right) = \int_{\Theta\times\Theta} \kappa^C(t)\widetilde{P}^p_{k+1}(s,t)\kappa^C(s)^\top\,\mathrm{d}s\,\mathrm{d}t + \hat{G}_k\hat{G}_k^\top \quad (43)$$

$$K_{k+1}(\xi) = \left(\int_\Theta \widetilde{P}^p_{k+1}(\xi,s)\kappa^C(s)^\top\,\mathrm{d}s\right) \cdot \left(\int_{\Theta\times\Theta} \kappa^C(t)\widetilde{P}^p_{k+1}(s,t)\kappa^C(s)^\top\,\mathrm{d}s\,\mathrm{d}t + \hat{G}_k\hat{G}_k^\top\right)^{-1} \quad (44)$$

This closed-form formula replaces the integral equation (36). The last step is the computation of $P_{k+1}$ from $P^p_{k+1}$ and $K_{k+1}$. Equation (31) becomes

$$(((K_{k+1} \circ C) \boxtimes I_{\mathcal{H}^n}) P^p_{k+1})(\xi,\zeta) \quad (45)$$

$$= \int_\Theta ((K_{k+1}(\xi)\kappa^C(s)) \otimes I_n) P^p_{k+1}(s,\zeta)\,\mathrm{d}s = \mathrm{st}\left(\int_\Theta \widetilde{P}^p_{k+1}(s,\zeta)\kappa^C(s)^\top\,\mathrm{d}s \cdot K^\top_{k+1}(\xi)\right) \quad (46)$$

$$P_{k+1}(\xi,\zeta) = P^p_{k+1}(\xi,\zeta) \quad (47)$$

$$- \mathrm{st}\left(\int_\Theta \widetilde{P}^p_{k+1}(s,\zeta)\kappa^C(s)^\top\,\mathrm{d}s \cdot K^\top_{k+1}(\xi)\right). \quad (48)$$

Summarizing, in the case of finite-dimensional output the Kalman filter can be computed by iterating the computation of equations (34), (44), (48) together with (26). It can be mentioned that when in addition the kernel $\kappa^C$ is constant, i.e. it is a constant matrix of size $q \times n$, equations (44) and (48) can be further simplified.

Another important case of finite-dimensional output is that of measurements taken at discrete spatial points,

$$y_k(\omega) = \mathrm{col}_{i=1}^q(C_i x_k(\omega,\xi_i)), \quad C_i \in \mathbb{R}^{1\times n}. \quad (49)$$

The theory developed so far does not allow for discrete observations of this kind, as this would require $C$ to be unbounded. In practice, however, the problem is easily solved, as already proposed in the continuous-time case for example in [18], by considering the components of $y_k$ as a linear function of the weighted average value of the state over a small spatial neighborhood of the discrete point $\xi_i$. This amounts to a definition of the kernel $\kappa^C : \mathcal{H}^n \to \mathbb{R}^q$ that is amenable to be processed by the filter (26), (34), (44), (48).

## V. Application: Design of a Distributed Thermo-Couple in the Presence of Thermal Noise

We illustrate an application of the distributed Kalman Filter to the conceptual design of a distributed thermo-couple that solves the problem of estimating the temperature profile of a bar of length 1 from measurements of the voltage difference of a current flowing through it.

In order to provide an explicit computation of the underlying kernels, we consider the heat equation on the interval $[0,1]$ with Dirichlet boundary conditions. The functions $\phi_m(\xi) = \sqrt{2}\sin(m\pi\xi)$, with $m = 1, 2, \ldots$, constitute an orthonormal basis for $L_2([0,1])$. Given $x = x(t,\xi)$ we have

$$\frac{\partial x}{\partial t} = \frac{\partial^2 x}{\partial \xi^2} \tag{50}$$

with Dirichlet boundary conditions $x(t,0) = x(t,1) = 0$. The solution to (50) with the initial condition

$$x(0,\xi) = x_0(\xi) = \sum_{m=1}^{+\infty} [\phi_m, x_0]\phi_m(\xi) \tag{51}$$

is given by

$$x(t,\xi) = \sum_{m=1}^{+\infty} \mathrm{e}^{-t\lambda_m}[\phi_m, x_0]\phi_m(\xi), \tag{52}$$

where $\lambda_m = (m\pi)^2$ [19]. Thus, by the definition of the inner product, the solution (52) is

$$x(t,\xi) = \int_0^1 \sum_{m=1}^{+\infty} \mathrm{e}^{-t\lambda_m}\phi_m(\xi)\phi_m(\eta)x_0(\eta)\,\mathrm{d}\eta, \tag{53}$$

and by considering a sequence $\{t_k\}$ with discretization step $t_{k+1} - t_k = \Delta > 0$, we can write

$$x(t_{k+1},\xi) = \int_0^1 \sum_{m=1}^{+\infty} \mathrm{e}^{-\Delta\lambda_m}\phi_m(\xi)\phi_m(\eta)x(t_k,\eta)\,\mathrm{d}\eta. \tag{54}$$

Finally, by considering also the external noise term $Fn_k$, we can write the difference equation in the form of (3) as

$$x_{k+1}(\xi) = \int_0^1 \sum_{m=1}^{+\infty} \mathrm{e}^{-\Delta\lambda_m}\phi_m(\xi)\phi_m(\eta)x_k(\eta)\,\mathrm{d}\eta + F(\xi)n_k, \tag{55}$$

where we recognize the kernel of the underlying operator $A$

$$\kappa^A(\xi,\eta) = \sum_{m=1}^{+\infty} \mathrm{e}^{-\Delta\lambda_m}\phi_m(\xi)\phi_m(\eta). \tag{56}$$

We note that we consider the case in which $f$ does not depend on the state process, thus $f(x_k) = F : \mathbb{R}^p \to \mathcal{H}^n$.

A current source is applied to the bar and we measure the voltage difference at the endpoints. As a first approximation, the measurement $y(t)$ is given by $y(t) = I \cdot r(t)$, $I$ is the applied current intensity, and $r$ is the resistance of the bar that depends on the temperature through the equation

$$r(t) = \int_0^1 (\beta + \alpha x(t,\xi))\mathrm{d}\xi, \tag{57}$$

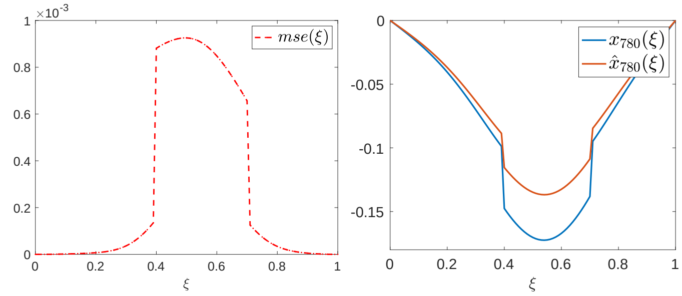


Fig. 1. Experimental average mean square error at each point (left). Temperature profile $x(\xi)$ of the bar at time $t \approx 7.8$ and its estimate $\hat{x}(\xi)$ (right).

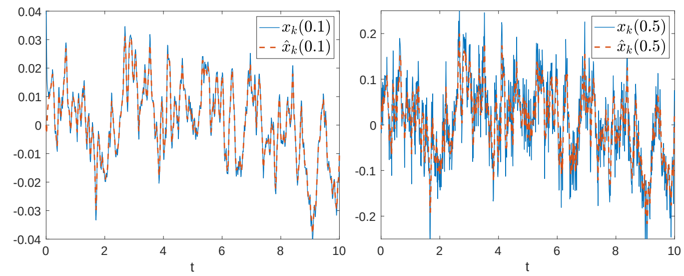


Fig. 2. Temperature $x_k(\xi)$ and its estimate $\hat{x}_k(\xi)$ as a function of time, for $\xi = 0.1$ (left) and $\xi = 0.5$ (right).

where $\alpha$ and $\beta$ are some non-negative coefficients depending on the bar. For simplicity, in what follows, we assume $\beta = 0$. By setting $c = I\alpha$ we have

$$y(t) = \int_0^1 c\,x(t,\xi)\mathrm{d}\xi. \tag{58}$$

In practice, measurements are acquired sampled and thus it is reasonable to consider the discrete-time model associated to (58). Also, we model the Johnson–Nyquist noise (or thermal noise), namely the electronic noise generated by the thermal agitation of the charge carriers inside the bar. In particular, by considering an infinitesimal portion of the bar $\mathrm{d}\ell$, the thermal noise, which depends on the temperature of the bar, is a voltage $\mathrm{d}V_T$ described by the following properties

$$\mathbb{E}[\mathrm{d}V_T(t,\xi)] = 0, \tag{59}$$

$$\mathbb{E}[(\mathrm{d}V_T(t,\xi))^2] = 4k_B x(t,\xi)\mathrm{d}r(t,\xi), \tag{60}$$

with $k_B$ the Boltzmann constant. Thus, being $\mathrm{d}r(t,\xi) = \alpha x(t,\xi)\mathrm{d}\xi$ from (57), we have

$$\mathbb{E}[V_T^2(t)] = \int_0^1 4\alpha K_B x^2(t,\xi)\mathrm{d}\xi = 4\alpha K_B \int_0^1 x^2(t,\xi)\mathrm{d}\xi. \tag{61}$$

Furthermore, with the previously defined sequence $\{t_k\}$, the discretization step $\Delta$, and the described Johnson–Nyquist noise term, the measurement equation becomes

$$y_k = \int_0^1 c\,x_k(\xi)\mathrm{d}\xi + G\left(\int_0^1 4\alpha K_B x_k^2(\xi)\mathrm{d}\xi\right)^{\frac{1}{2}} n_k \tag{62}$$

$$= Cx_k + G\left(\tilde{\kappa}\,Cx_k^2\right)^{\frac{1}{2}} n_k \tag{63}$$

where $\tilde{\kappa} = \frac{4\alpha K_B}{c}$, $G$ a constant matrix such that $FG^\top = 0$[2], and $C$ is the output map operator with kernel

$$\kappa^C(\xi, \eta) = \kappa^C = c. \quad (64)$$

We assume $n_k \in \mathbb{R}^2$ to be a Gaussian white noise vector, $F = (f(\xi), 0)$ and $G = (0, g)$ (s.t. $FG^\top = 0$). We suppose that the bar is affected by a random thermal noise for $\xi \in [0.4,\, 0.7]$. The state and measurement noise amplitudes are

$$f(\xi) = \begin{cases} 0.2 \cdot \xi(\xi - 1), & \xi \in [0.4,\, 0.7] \\ 0, & \xi \notin [0.4,\, 0.7] \end{cases}, \qquad g = 0.1 \quad (65)$$

The simulation results in this section refer to $N_s = 50$ simulations, with $t \in [0, \bar{t}]$, $\bar{t} = 10$, $\Delta = 1 \cdot 10^{-2}$ (i.e. $\bar{t}/\Delta = 1000$ time points). The measurement parameters are $I = 5$, $\alpha = 1, \beta = 0$, that yield $c = 5$. The initial condition is

$$x_0(\xi) = \; 0.1 \sin(2\pi\xi), \quad \xi \in [0,\, 1]. \quad (66)$$

The filter equation is given by (26). In the numerical computation we used the first 10 eigenfunctions $\phi_m$ (e.g., $m = 1, 2, \ldots, 10$), and a spatially discretized representation of the vectors $P$, $P^p$ with $N = 10^2$ uniformly distributed points in the interval $[0, 1]$. In order to test the numerical accuracy of the estimate, we show in Fig. 1 (left), the average mean square error (mse) at each $\xi \in [0, 1]$ across simulations. The mse is computed as

$$\text{mse}(\xi) = \frac{1}{N_s(N_t - \tau)} \sum_{i=1}^{N_s} \sum_{k=\tau}^{N_t} (x_k^{(i)}(\xi) - \hat{x}_k^{(i)}(\xi))^2, \quad (67)$$

where $N_t = \bar{t}/\Delta = 1000$ is the number of time samples, $x_k^i(\xi)$ and $\hat{x}_k^i(\xi)$ denote the true and estimated state of $i$-th simulation at time $t_k$. $\tau = 100$ has been chosen to exclude the transient. For illustrative purposes we plot in Fig. 1 (right) the true and estimated temperature profile at time $t \approx 7.8$ for a single noise realization, and in Fig. 2 the true and estimated temperature evolution at points $\xi = 0.1$ and $\xi = 0.5$ for the same noise realization.

## VI. Conclusions

This new formulation for filtering systems on Hilbert space provides an efficient tool for the state estimation of discrete-time infinite-dimensional linear stochastic systems with multiplicative noise. The formalism adopted here is promising for the design of polynomial estimators and regulators that could better cope with non-Gaussian noise terms ([20], [21]). Furthermore, other natural extensions of this method could be the consideration of an underlying packet dropping network with intermittent or delayed observations/control ([22], [23]). Finally, another pioneering direction could be the one of considering the distributed filtering problem in discrete-time ([24]) for infinite-dimensional systems.

[2]Note that it is enough to set $F = \text{col}(\tilde{F}, 0)$ and $G = \text{col}(0, \tilde{G})$ with the appropriate dimensions.

## References


[1] R. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME - Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[2] A. Jazwinski, *Stochastic processes and filtering theory*. Courier Corporation, 2007.

[3] H. Kushner, "Dynamical equations for optimal nonlinear filtering," *Journal of Differential Equations*, vol. 3, no. 2, pp. 179–190, 1967.

[4] P. Falb, "Infinite-dimensional filtering: The Kalman-Bucy filter in Hilbert space," *Information and Control*, vol. 11, pp. 102–137, 1967.

[5] A. Bensoussan, *Stochastic control by functional analysis methods*. Elsevier, 2011.

[6] R. Curtain, "Infinite-dimensional filtering," *SIAM Journal on Control*, vol. 13, no. 1, pp. 89–104, 1975.

[7] ——, "A survey of infinite-dimensional filtering," *Siam Review*, vol. 17, no. 3, pp. 395–411, 1975.

[8] C. Kitsos, G. Besancon, and C. Prieur, "High-gain observer design for a class of quasi-linear integro-differential hyperbolic systems-application to an epidemic model," *IEEE Transactions on Automatic Control*, 2021.

[9] F. Castillo, E. Witrant, C. Prieur, and L. Dugard, "Boundary observers for linear and quasi-linear hyperbolic systems with application to flow control," *Automatica*, vol. 49, no. 11, pp. 3180–3188, 2013.

[10] A. Deutschmann, L. Jadachowski, and A. Kugi, "Backstepping-based boundary observer for a class of time-varying linear hyperbolic PIDEs," *Automatica*, vol. 68, pp. 369–377, 2016.

[11] S. Sallberg, "Sampled-data Kalman filtering and multiple model adaptive estimation for infinite-dimensional continuous-time systems," *PhD dissertation AFIT/DS/ENG/07-08*, 2007.

[12] S. Sallberg, P. Maybeck, and M. Oxley, "Infinite-dimensional sampled-data Kalman filter," in *49th IEEE Conference on Decision and Control (CDC)*. IEEE, 2010, pp. 7363–7368.

[13] ——, "Infinite-dimensional sampled-data Kalman filtering and the stochastic heat equation," in *49th IEEE Conference on Decision and Control (CDC)*. IEEE, 2010, pp. 5062–5067.

[14] F. Cacace, M. d'Angelo, V. De Iuliis, and A. Germani, "Kalman filtering for linear discrete-time systems in $L_2$ spaces with applications," *Systems & Control Letters*, submitted.

[15] M. Piotto, A. Catania, A. Nannini, and P. Bruschi, "Thermal noise-boosting effects in hot-wire-based micro sensors," *Journal of Sensors*, vol. 2020, 2020.

[16] A. Vinante, R. Mezzena, G. A. Prodi, S. Vitale, M. Cerdonio, M. Bonaldi, and P. Falferi, "Thermal noise in a high qultracryogenic resonator," *Review of scientific instruments*, vol. 76, no. 7, p. 074501, 2005.

[17] M. Zakai, "On the optimal filtering of diffusion processes," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 11, no. 3, pp. 230–243, 1969.

[18] A. Bensoussan, "Optimization of sensors' location in a distributed filtering problem," in *Stability of stochastic dynamical systems*. Springer, 1972, pp. 62–84.

[19] R. Curtain and H. Zwart, *An introduction to infinite-dimensional linear systems theory*. Springer Science & Business Media, 2012, vol. 21.

[20] S. Battilotti, F. Cacace, M. d'Angelo, and A. Germani, "The polynomial approach to the LQ non-Gaussian regulator problem through output injection," *IEEE Transactions on Automatic Control*, vol. 64, no. 2, pp. 538–552, 2018.

[21] F. Cacace, F. Conte, M. d'Angelo, and A. Germani, "Feedback polynomial filtering and control of non-gaussian linear time-varying systems," *Systems & Control Letters*, vol. 123, pp. 108–115, 2019.

[22] S. Battilotti and M. d'Angelo, "Stochastic output delay identification of discrete-time Gaussian systems," *Automatica*, vol. 109, p. 108499, 2019.

[23] S. Battilotti, F. Cacace, M. d'Angelo, A. Germani, and B. Sinopoli, "LQ non-Gaussian regulator with Markovian control," *IEEE Control Systems Letters*, vol. 3, no. 3, pp. 679–684, 2019.

[24] S. Battilotti, F. Cacace, and M. d'Angelo, "A stability with optimality analysis of consensus-based distributed filters for discrete-time linear systems," *Automatica*, vol. 129, p. 109589, 2021.