

A Consensus Kalman Filter on L_2 spaces

S. Battilotti^a A. Borri^b F. Cacace^c M. d'Angelo^{b,d} A. Germani^e

^a*Dipartimento di Ingegneria Informatica, Automatica e Gestionale, Università di Roma La Sapienza, Rome, Italy*

^b*CNR-IASI, National Research Council of Italy, Rome, Italy*

^c*Università Campus Bio-Medico di Roma, Rome, Italy*

^d*Università Mercatorum, Rome, Italy*

^e*DISIM, Università degli Studi dell'Aquila, L'Aquila, Italy*

1 Introduction

This paper addresses the problem of filtering infinite dimensional systems over sensor networks. We first recall the main results of the two distinct area of filtering infinite dimensional systems and distributed filtering over sensor networks as well as the motivations and the challenges underlying their integration. Distributed filtering algorithms are an active area of research and one of the main applications of wireless sensor networks (WSNs), consisting of numerous sensors distributed across different geographical locations, each operating under low power constraints and having limited computational capabilities (Akyildiz et al., 2002). In the context of distributed estimation for WSNs, each sensor possesses a subset of observations that are shared among neighboring nodes to cooperatively perform large-scale sensing tasks that cannot be accomplished by individual devices (He et al., 2020). The large majority of these distributed filtering algorithms trace their origins back to the seminal works (Olfati-Saber, 2005, 2009) on the consensus-based paradigm. They can be broadly classified based on the information on which consensus is reached (He et al., 2020): state estimate fusion (Talebi et al., 2021; Yang et al., 2019), measurement vector fusion (Ge et al., 2017; Kamgarpour and Tomlin, 2008; Li et al., 2018; Wan et al., 2018; Wu et al., 2018), and information vector fusion (Battistelli and Chisci, 2014; Kamal et al., 2013; Talebi and Werner, 2019; Wang and Ren, 2017; Wei et al., 2018). In our previous work (Battilotti et al., 2020) we proved that in the state estimate fusion approach at continuous time a suitable choice of the consensus gain allows to approximate as much as desired

the performance of the centralized Kalman filter, i.e. the filter that uses in a centralized way all the sensing formation, whereas for discrete-time systems this can be accomplished by increasing the number of intermediate consensus steps (Battilotti et al., 2021).

The availability of low-cost sensors and the diffusion of wireless networks has contributed in recent years to the development of applications such as weather forecasting (Fisher et al., 2009), diffusion of pollutants (Ortner et al., 2007), smart grids (Moura et al., 2013) and environmental monitoring (Silva et al., 2015) where the state of the system to be estimated is inherently infinite dimensional. The filtering problem of infinite dimensional systems attracted much attention in the years following the introduction of the Kalman filter (Kalman, 1960; Kalman and Bucy, 1961) and found a quite early solution in the Kalman-Bucy infinite dimensional filter of Falb (Falb, 1967), that includes distributed parameter systems with bounded integral operators but it is not applicable to the more interesting cases of partial differential operators. Remarkably, most of the subsequent efforts on filtering infinite dimensional systems that were developed in the 70's were focused on the case of continuous-time systems (Balakrishnan (1974); Bensoussan (2011); Curtain (1975b,a) and the references therein), and only a few recent works take into consideration the case of sampled-data measurements (Sallberg et al., 2010a,b) that is of great relevance in modern digital applications. Analogously, most recent papers on multi-agents cooperative estimation of continuous-time systems propose observers for specific deterministic partial differential equations (PDEs). Demetriou (2010) considers general infinite dimensional deterministic linear systems in Hilbert space and proposes observers with a consensus term obtained by exchanging estimates. Modal decomposition is used in Wang and Wu (2015) to design a H_∞ observer for a deterministic PDE with nonlinear disturbance. The use of modal

Email addresses: `battilotti@diag.uniroma1.it` (S. Battilotti), `alessandro.borri@iasi.cnr.it` (A. Borri), `f.cacace@unicampus.it` (F. Cacace), *corresponding author:* `massimiliano.dangelo@unimercatorum.it` (M. d'Angelo), `alfredo.germani@univaq.it` (A. Germani).

decomposition allows to reach consensus by exchanging finite dimensional information among agents. In Wang et al. (2017) a multi-agent problem on a specific deterministic PDE is solved through an array of observers that operate on spatially separated domains without a consensus protocol. Modal decomposition is used also in Demetriou (2022) to solve the problem of cooperative estimation for a deterministic parabolic PDE with both point-wise and integral measurements. The approach in Xia and Qian (2022) considers a specific PDE with nonlinear term, disturbances and possible loss of sensor measurements. An adaptive cooperative observer is proposed in Cai et al. (2024) for a deterministic parabolic PDE with nonlinear uncertainties. In both Xia and Qian (2022) and Cai et al. (2024) consensus is obtained by exchanging infinite dimensional estimates among agents. In contrast to these observers, we study the case of general discrete-time stochastic linear systems in an optimal filtering setting. Our framework is similar to Khan and Moura (2008) and Battistelli et al. (2016), where the idea is to use a discrete-time filter on the discretized solution of a PDE by decomposing the spatial domain into subdomains assigned to different processing nodes interconnected to form a network. This setting is well suited to the case when the integral operator is the semi-group generated by an underlying continuous-time dynamics and it therefore allows to filter systems with a continuous-time dynamics represented through a partial differential equation and sampled measurements, without resorting to finite-dimensional projection or modal decomposition. Our approach is more general because the system is not necessarily the discretization of a PDE and sensors do not necessarily refers to separated spatial domains. It is worth mentioning that large-scale Kalman filters with consensus dynamics have also been proposed for approximating the estimate of the posterior density of non-linear finite dimensional systems at discrete time through a kernel mean embedding approach evolving in an infinite dimensional reproducing Kernel Hilbert space Guo et al. (2024).

An important problem of many of the cited consensus estimation algorithms is represented by communication constraints, since consensus algorithms that require to exchange estimates like Demetriou (2010), Xia and Qian (2022) or Cai et al. (2024) are hardly applicable, and particularly at continuous-time, in a context where estimates are infinite dimensional. To address this problem we propose to extend to the infinite dimensional case a new algorithm (Battilotti et al., 2024). The main advantage of this method is that at run-time the nodes need only to exchange finite dimensional information but the accuracy of the centralized Kalman filter can be approximated as much as desired by increasing the number of consensus steps. In the actual filtering phase the exchange of information is the same as in the finite dimensional case and the computational burden is moderate.

The outline and contributions of the paper can be sum-

marized as follows. After introducing the required notation and reviewing some preliminaries on linear stochastic systems in Section 2, in Section 3 we provide a formulation of the Kalman filter for discrete-time stochastic systems evolving in L_2 that is amenable of an efficient computation. Although the theory is already well established, we claim that our new formulation is useful and relevant even in the centralized case for the practical implementation of filters for systems obeying partial differential equations with discrete-time measurements. In Section 4 we describe the version over sensor networks of this filter based on the exchange of finite dimensional information and analyze its properties, in particular the convergence to the accuracy of the centralized Kalman filter. Section 5 reports numerical results on a case study.

2 System model

2.1 Preliminaries

Notation. The norm $\|\cdot\|$ without specifications is reserved to the Euclidean norm in $\mathbb{R}^n$ and to the operator norm, $\|A\| = \sup_{\|x\|=1} \|Ax\|$. For $M \in \mathbb{R}^{m \times n}$, $\text{tr}(M)$ is the trace of M , $\text{st}(M) \in \mathbb{R}^{mn}$ denotes the vertical stack of M , and, for $v \in \mathbb{R}^{mn}$, $\text{st}_m^{-1}(v) \in \mathbb{R}^{m \times n}$ is the inverse operation. By $\text{col}_i(M_i)$, $\text{row}_i(M_i)$, $\text{diag}_i(M_i)$ we denote, respectively, the vertical, horizontal and diagonal composition of vectors and matrices. $\mathbf{1}_N$ is vector of size N with entries 1. $\mathcal{H}^n = L_2(\Theta; \mathbb{R}^n)$, $n \in \mathbb{N}$, is the Hilbert space of square integrable functions from a compact domain $\Theta \subset \mathbb{R}^s$ to $\mathbb{R}^n$ with respect to the standard Lebesgue measure. $\mathcal{H}^n$ is endowed with the standard inner product $[x, y]_{\mathcal{H}^n} = \int_{\Theta} x^\top(s)y(s) ds$. The inner product $[\cdot, \cdot]$ will be tacitly assumed to be in $\mathcal{H}^n$ if not otherwise specified. Operators on $\mathcal{H}^n$ are in boldface. $\mathbf{I}$ denotes the identity in $\mathcal{H}^n$. We denote with $\mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}} = L_2(\Omega; \mathcal{F}; \mathcal{H}^n; P)$ the Hilbert space of square integrable $\mathcal{F}$ -measurable functions on a probability space $(\Omega, \mathcal{F}, P)$, with values in $\mathcal{H}^n$. $\mathbb{E}[x] \in \mathcal{H}^n$ denotes the expected value of $x \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$, $\mathbb{E}[x] = \int_{\Omega} x(\omega) dP$. $\mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$ is endowed with the inner product $[x_1, x_2]_{\mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}} = \mathbb{E}[[x_1, x_2]_{\mathcal{H}^n}]$. The covariance of an element of $\mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$ is a bounded linear self-adjoint non-negative operator $\mathcal{H}^n \rightarrow \mathcal{H}^n$, see Appendix A.

Mixing operators and matrices. Given a vector $x = \text{col}_{i=1}^N(x_i)$ of N functions $x_i \in \mathcal{H}^n$ and a vector $m \in \mathbb{R}^N$ we denote by $m^\top x$ the function in $\mathcal{H}^n$ defined as $(m^\top x)(\xi) = \sum_{i=1}^N m_i x_i(\xi)$. Given a matrix $M \in \mathbb{R}^{N \times N}$, $M = \text{col}_{i=1}^N(m_i^\top)$, we denote by Mx the vector of N functions in $\mathcal{H}^n$

$$Mx := \text{col}_{i=1}^N(m_i^\top x). \quad (1)$$

The stack $\text{col}_{i=1}^N(\mathbf{P}_i)$ of N linear operators $\mathbf{P}_i : \mathcal{H}^n \rightarrow \mathcal{H}^n$ is an operator $\mathcal{H}^{nN} \rightarrow \mathcal{H}^{nN}$ such that

$$\text{col}_{i=1}^N(\mathbf{P}_i)x := \text{col}_{i=1}^N(\mathbf{P}_i x_i).$$

Consequently, given $M \in \mathbb{R}^{N \times N}$, $M \text{col}_{i=1}^N(\mathbf{P}_i)$ is the operator $\mathcal{H}^{nN} \rightarrow \mathcal{H}^{nN}$ such that

$$M \text{col}_{i=1}^N(\mathbf{P}_i)x = M \text{col}_{i=1}^N(\mathbf{P}_i x_i).$$

In particular, it is easy to check that when $\mathbf{P}_i = \mathbf{P}$

$$M \text{col}_{i=1}^N(\mathbf{P})x = \text{col}_{i=1}^N(\mathbf{P})Mx. \quad (2)$$

2.2 Linear stochastic systems in L_2

In a probability space $(\Omega, \mathcal{F}, P)$ we consider the following model of infinite dimensional stochastic discrete-time systems with finite dimensional measurements,

$$x_{k+1}(\omega) = \mathbf{A}x_k(\omega) + \mathbf{F}n_k(\omega), \quad (3)$$

$$y_k(\omega) = \mathbf{C}x_k(\omega) + Gn_k(\omega), \quad (4)$$

where $\omega \in \Omega$ and:

- $x_k \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$, i.e. $x_k(\omega) \in \mathcal{H}^n$.
- $y_k \in \mathcal{W}_{\mathbb{R}^q}^{\mathcal{F}}$, i.e. $y_k(\omega) \in \mathbb{R}^q$.
- $n_k \in \mathcal{W}_{\mathbb{R}^p}^{\mathcal{F}}$, i.e. $n_k(\omega) \in \mathbb{R}^p$, $\mathbb{E}[n_k n_j^*] = 0$ for $k \neq j$ and $\mathbb{E}[n_k n_k^*] = I_p$.
- $\mathbf{A} : \mathcal{H}^n \rightarrow \mathcal{H}^n$ is a linear operator

$$(\mathbf{A}x)(\xi) = \int_{\Theta} \kappa_A(\xi, s)x(s) ds, \quad (5)$$

with kernel $\kappa_A \in \mathcal{H}_{(2)}^{n \times n} = L_2(\Theta \times \Theta; \mathbb{R}^{n \times n})$, see Appendix A.

- $\mathbf{C} : \mathcal{H}^n \rightarrow \mathbb{R}^q$ is a linear operator

$$\mathbf{C}x = \int_{\Theta} \kappa_C(s)x(s) ds, \quad (6)$$

with $\kappa_C \in \mathcal{H}^{q \times n} = L_2(\Theta; \mathbb{R}^{q \times n})$, see Appendix A.

- $\mathbf{F} : \mathbb{R}^p \rightarrow \mathcal{H}^n$ is defined for all $u \in \mathbb{R}^p$, by $(\mathbf{F}u)(\xi) = \bar{F}(\xi)u$ with $\bar{F} \in \mathcal{H}^{n \times p} = L_2(\Theta; \mathbb{R}^{n \times p})$.
- $G \in \mathbb{R}^{q \times p}$ is a matrix such that $GG^{\top}$ is positive definite and, $\forall \xi \in \Theta$, $\bar{F}(\xi)G^{\top} = 0$.

In the sequel we shall usually omit the dependence from ω for concision. Notice that both the state noise and the measurement noise are finite dimensional white sequences, and the operator $\mathbf{F}$ “spreads” n_k on all Θ . It is easy to check that the requirement for $\mathbf{A}$, $\mathbf{C}$, $\mathbf{F}$ to have square-integrable kernels implies that:

- $\mathbf{A}$ and $\mathbf{F}$ are bounded operators and if $x \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$ has bounded covariance operator $\mathbf{P}$, then $\mathbf{A}x$ has bounded covariance $\mathbf{A}\mathbf{P}\mathbf{A}^* : \mathcal{H}^n \rightarrow \mathcal{H}^n$ (see Lemma 10) with kernel

$$\kappa_{\mathbf{A}\mathbf{P}\mathbf{A}^*}(\xi, \zeta) = \int_{\Theta \times \Theta} \kappa_A(\xi, s)\kappa_P(s, t)\kappa_A^{\top}(t, \zeta) ds dt. \quad (7)$$

- $\mathbf{C}$ is a bounded operator and if $x \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$ has bounded covariance operator $\mathbf{P}$, then the covariance $\mathbf{C}\mathbf{P}\mathbf{C}^*$ of $\mathbf{C}x$ is a matrix in $\mathbb{R}^{q \times q}$ with bounded norm.
- The process and measurement noises are uncorrelated (see Definition 6 in Appendix A), that is, $\mathbf{F}n_k \perp Gn_k$ because $\mathbb{E}[(\mathbf{F}n_k)(Gn_k)^*] = 0$.
- The covariance $\mathbf{Q} = \mathbf{F}\mathbf{F}^*$ of the process noise $\mathbf{F}n_k \in \mathcal{H}^n$ is a bounded operator $\mathcal{H}^n \rightarrow \mathcal{H}^n$ with kernel $\kappa_Q(\xi, \zeta) = \bar{F}(\xi)\bar{F}(\zeta)^{\top}$, $\kappa_Q \in \mathcal{H}_{(2)}^{n \times n}$.
- The measurement noise covariance $R := \mathbb{E}[Gn_k(Gn_k)^{\top}]$ is positive definite.

The definition of observability and controllability for discrete-time systems in L_2 is similar to the finite dimensional case (see Fuhrmann (1976), and Balakrishnan (2012), Sec 4.9-4.10).

Definition 1 *Given the system*

$$x_{k+1} = \mathbf{A}x_k + \mathbf{B}u_k, \quad (8)$$

with $x_k \in \mathcal{H}^n$, $u_k \in \mathcal{H}^p$, and $\mathbf{A} : \mathcal{H}^n \rightarrow \mathcal{H}^n$, $\mathbf{B} : \mathcal{H}^p \rightarrow \mathcal{H}^n$ bounded linear operators, then the pair $(\mathbf{A}, \mathbf{B})$ is said to be controllable if for any $\bar{x} \in \mathcal{H}^n$ there exists a time $\bar{T} > 0$ such that for any $T \geq \bar{T}$ there exists a sequence $\{u_k\}_{k=0}^T$ such that $x_T = \bar{x}$.

Definition 2 *Consider system (3)–(4) with $\mathbf{F} = \mathbf{0}$ and $G = 0$. The pair $(\mathbf{A}, \mathbf{C})$ is said to be observable if there exists a time $\bar{T} > 0$ such that, for any output sequence, $\|y\|_{\ell_2([0, \bar{T}])}^2 = \sum_{k=0}^{\bar{T}} \|y_k\|^2 = 0$ implies $x_0 = 0$.*

Notice that for an observable system, $\sum_{k=0}^T \|y_k\|^2 = 0 \Rightarrow x_0 = 0$ holds for any $T > \bar{T}$. The following characterization of controllable and observable systems is derived from results in the literature (see Theorem 3.2 and 3.4 in Fuhrmann (1976), Corollary 4.9.1 in Balakrishnan (2012)).

Lemma 1 *The pair $(\mathbf{A}, \mathbf{B})$, with $\mathbf{A} : \mathcal{H}^n \rightarrow \mathcal{H}^n$ and $\mathbf{B} : \mathcal{H}^p \rightarrow \mathcal{H}^n$ is controllable if and only if $\exists \bar{T} > 0$ such that for all $T > \bar{T}$ the operator $\mathcal{C} : \mathcal{H}^n \rightarrow \mathcal{H}^n$ defined by*

$$\mathcal{C} = \sum_{k=0}^{T-1} \mathbf{A}^k \mathbf{B} \mathbf{B}^* \mathbf{A}^{*k}$$

is positive definite.

Lemma 2 The pair (A, C) , with $A : \mathcal{H}^n \rightarrow \mathcal{H}^n$ and $C : \mathcal{H}^n \rightarrow \mathbb{R}^q$ is observable if and only if $\exists \bar{T} > 0$ such that for all $T > \bar{T}$ the operator $\mathcal{O} : \mathcal{H}^n \rightarrow \mathcal{H}^n$ defined by

$$\mathcal{O} = \sum_{k=0}^{T-1} A^{k*} C^* C A^k$$

is positive definite and in that case, for $T \geq \bar{T}$

$$x_0 = \mathcal{O}^{-1} \sum_{k=0}^{T-1} A^{k*} C^* y(k).$$

3 The centralized Kalman filter on L_2 spaces

3.1 Optimal linear estimate

Definition 3 Given $X \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$ and a subspace $\mathcal{S}$ of $\mathcal{H}^n$, the projection of X on $\mathcal{S}$, denoted $\Pi[X|\mathcal{S}]$ is the unique vector $Z \in \mathcal{W}_{\mathcal{S}}^{\mathcal{F}}$ such that $[X - Z, W]_{\mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}} = 0, \forall W \in \mathcal{W}_{\mathcal{S}}^{\mathcal{F}}$.

Let $\epsilon_X = X - \Pi[X|\mathcal{S}]$. Since $\Pi[X|\mathcal{S}] \in \mathcal{S}$,

$$[\epsilon_X, \Pi[X|\mathcal{S}]]_{\mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}} = 0.$$

The optimal estimate $\hat{x}_k$ of $x_k \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$ in the minimum variance sense is the conditional expectation with respect to the σ -algebra $\mathcal{F}_k^y$ generated by the output sequence $Y^k = \{y_0, \dots, y_k\}$, that is, $\hat{x}_k = E[x_k | \mathcal{F}_k^y]$. The conditional expectation can be equivalently expressed as the projection of $x_k \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$ on the subspace of the $\mathcal{F}_k^y$ -measurable functions $\Omega \rightarrow \mathcal{H}^n$, that is, $\hat{x}_k = \Pi[x_k | \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}_k^y}]$. In the hypothesis that the sequences x_k and y_k are jointly Gaussian, the projection of x_k onto $\mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}_k^y}$ is identical to the projection on the space $\mathcal{L}_k^y$ of the affine functions of Y^k having value in $\mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}_k^y}$,

$$\mathcal{L}_k^y = \{z \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}_k^y} : z = \sum_{i=0}^k K_i y_i + f, f \in \mathcal{H}^n\},$$

where $K_i : \mathbb{R}^q \rightarrow \mathcal{H}^n$ are bounded linear operators defined by $(K_i y)(\xi) = \bar{K}_i(\xi) y$ with $\bar{K}_i \in \mathcal{H}^{n \times q}$. We can then write

$$\hat{x}_k = \Pi[x_k | \mathcal{L}_k^y], \quad (9)$$

and the centralized Kalman filter algorithm consists in the recursive computation of (9). When the sequences x_k and y_k are not Gaussian, the conditional expectation $E[x_k | \mathcal{F}_k^y]$ does not coincide with (9), that however provides the best linear un-biased estimator in the minimum variance sense.

3.2 Structure of the centralized Kalman filter

The derivation of the Kalman filter in our setting follows standard arguments used in the finite dimensional case that we briefly recall. We define the space $\mathcal{H}_k^y$ of the linear bounded transformations of $Y^k \in \mathcal{W}_{\mathbb{R}^{kq}}^{\mathcal{F}_k^y}$,

$$\mathcal{H}_k^y = \{z \in \mathcal{W}_{\mathbb{R}^q}^{\mathcal{F}_k^y} : z = \sum_{i=0}^k M_i y_i, M_i \in \mathbb{R}^{q \times q}\},$$

and the output innovation sequence $\nu_k^o \in \mathcal{W}_{\mathbb{R}^q}^{\mathcal{F}_k^y}$ as $\nu_0^o = y_0$,

$$\nu_k^o = y_k - \Pi(y_k | \mathcal{H}_{k-1}^y),$$

and by $\mathcal{H}_k^{\nu^o}$ the space of the linear bounded transformations of ν_k^o , $\mathcal{H}_k^{\nu^o} = \{M \nu_k^o | M \in \mathbb{R}^{q \times q}\}$. Since $\nu^k \perp y_{k-1}$ because $y_{k-1} \in \Pi(y_k | \mathcal{H}_{k-1}^y)$, the sub-spaces $\mathcal{H}_k^{\nu^o}$ and $\mathcal{H}_{k-1}^y$ are orthogonal and $\mathcal{H}_k^y = \mathcal{H}_{k-1}^y \oplus \mathcal{H}_k^{\nu^o}$. Consequently, $\mathcal{H}_k^y$ coincides with the space of the linear bounded transformations on the output innovation sequence $\{\nu_0^o, \dots, \nu_k^o\}$ in $\mathcal{W}_{\mathcal{H}^q}^{\mathcal{F}_k^y}$.

Proposition 11 in Appendix states that orthogonality is preserved under linear transformation from $\mathbb{R}^q$ to $\mathcal{H}^n$, thus $\mathcal{L}_k^y = \mathcal{L}_{k-1}^y \oplus \mathcal{L}_k^{\nu^o}$, where this time $\mathcal{L}_k^{\nu^o}$ is the space of the linear bounded transformations on the output innovation sequence to $\mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}_k^y}$. By exploiting the property of projection over orthogonal subspaces we obtain

$$\hat{x}_k = \Pi(x_k | \mathcal{L}_k^y) = \Pi(x_k | \mathcal{L}_{k-1}^y) + \Pi(x_k | \mathcal{L}_k^{\nu^o}).$$

Now:

- (1) Since $F n_{k-1}$ and the elements of $\mathcal{L}_{k-1}^y$ are uncorrelated,
$$\begin{aligned} \Pi(x_k | \mathcal{L}_{k-1}^y) &= \Pi(A x_{k-1} + F n_{k-1} | \mathcal{L}_{k-1}^y) \\ &= A \Pi(x_{k-1} | \mathcal{L}_{k-1}^y) + \Pi(F n_{k-1} | \mathcal{L}_{k-1}^y) \\ &= A \hat{x}_{k-1}. \end{aligned}$$
- (2) $\Pi(x_k | \mathcal{L}_k^{\nu^o}) = K_k \nu_k^o$, where $K_k : \mathbb{R}^q \rightarrow \mathcal{H}^n$ is the linear operator generated by a $\bar{K}_k \in \mathcal{H}^{n \times q}$ and it corresponds to the Kalman gain.
- (3) $\Pi(y_k | \mathcal{L}_{k-1}^y) = \Pi(C x_k + G n_k | \mathcal{L}_{k-1}^y) = (C A) \hat{x}_{k-1}$, due to the uncorrelation between $G n_k$ and the elements of $\mathcal{L}_{k-1}^y$, where $C A : \mathcal{H}^n \rightarrow \mathbb{R}^q$ is the composition of C and A .

We thus obtain the classic recursive equation

$$\hat{x}_k = A \hat{x}_{k-1} + K_k (y_k - (C A) \hat{x}_{k-1}) \quad (10)$$

in the unknown $K_k : \mathbb{R}^q \rightarrow \mathcal{H}^n$. We are left with the task of deriving the time-varying matrix function $\bar{K}_k \in \mathcal{H}^{n \times q}$ of the Kalman gain operator K_k .

Let us define the *state estimation error* $e_k = x_k - \hat{x}_k \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$, the *prediction error* $e_k^p = x_k - \mathbf{A}\hat{x}_{k-1} \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$, and the *state innovation* $\nu_k^s \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}_k^y}$

$$\nu_k^s := \hat{x}_k - \mathbf{A}\hat{x}_{k-1} = \mathbf{K}_k \nu_k^o, \quad (11)$$

where the last equality holds by Equation (10).

Definition 4 *The state innovation, output innovation, prediction error covariance and estimation error covariance are defined, respectively, as follows*

$$\begin{aligned} \mathbf{P}_k^{s,o} &:= \mathbb{E}[\nu_k^s \nu_k^{o*}] : \mathbb{R}^q \rightarrow \mathcal{H}^n, \\ \mathbf{P}_k^o &:= \mathbb{E}[\nu_k^o \nu_k^{o*}] \in \mathbb{R}^{q \times q}, \\ \mathbf{P}_k^p &:= \mathbb{E}[e_k^p e_k^{p*}] : \mathcal{H}^n \rightarrow \mathcal{H}^n, \\ \mathbf{P}_k &:= \mathbb{E}[e_k e_k^*] : \mathcal{H}^n \rightarrow \mathcal{H}^n. \end{aligned}$$

Theorem 1 *The output innovation, prediction error covariance and estimation error covariance evolve as*

$$\mathbf{P}_k^p = \mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^* + \mathbf{Q}, \quad (12)$$

$$\mathbf{P}_k^o = \mathbf{C}\mathbf{P}_k^p\mathbf{C}^* + \mathbf{R}, \quad (13)$$

$$\mathbf{P}_k^{s,o} = \mathbf{P}_k^p\mathbf{C}^* = \mathbf{K}_k\mathbf{P}_k^o, \quad (14)$$

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{C})\mathbf{P}_k^p. \quad (15)$$

Moreover, $\mathbf{P}_k^o$ is non-singular and if $\mathbf{P}_{k-1}$ is bounded then $\mathbf{P}_k^p$ and $\mathbf{P}_k^o$ are bounded. From (14) it follows that

$$\mathbf{K}_k = \mathbf{P}_k^p\mathbf{C}^*(\mathbf{P}_k^o)^{-1}. \quad (16)$$

The proof is reported in Appendix B.

Corollary 1 *The recursive covariance Equations (12), (15) and (16) can be written explicitly in integral forms with respect to the integral kernels as*

$$\begin{aligned} \kappa_{P_k^p}(\xi, \zeta) &= \int_{\Theta^2} \kappa_A(\xi, s) \kappa_{P_{k-1}}(s, t) \kappa_A(t, \zeta)^\top ds dt \\ &\quad + \bar{F}(\xi) \bar{F}(\zeta)^\top, \end{aligned} \quad (17)$$

$$\begin{aligned} \bar{K}_k(\xi) &= \int_{\Theta} \kappa_{P_k^p}(\xi, s) \kappa_C(s)^\top ds \\ &\quad \times \left(\int_{\Theta^2} (\kappa_C(s) \kappa_{P_k^p}(s, t) \kappa_C(t)^\top + R) ds dt \right)^{-1}, \end{aligned} \quad (18)$$

$$\kappa_{P_k}(\xi, \zeta) = \kappa_{P_k^p}(\xi, \zeta) - \bar{K}_k(\xi) \int_{\Theta} \kappa_C(t) \kappa_{P_k^p}(t, \zeta) dt. \quad (19)$$

Proof. Equations (17) and (18) are obtained by explicitly writing the expression of $\mathbf{P}_k^p x$ given by (12) for $x \in \mathcal{H}^n$

and of $\mathbf{K}_k y$ given by (16) for $y \in \mathbb{R}^q$,

$$\bar{K}(\xi)y = (\mathbf{K}_k y)(\xi) = \int_{\Theta} \kappa_{P_k^p}(\xi, s) \kappa_C(s)^\top ds (\mathbf{P}_k^o)^{-1} y, \quad (20)$$

with $\mathbf{P}_k^o$ computed as in (B.1). Eq. (19) descends from

$$\begin{aligned} (\mathbf{P}_k x)(\xi) &= \int_{\Theta} \kappa_{P_k^p}(\xi, s) x(s) ds \\ &\quad - \bar{K}_k(\xi) \int_{\Theta^2} \kappa_C(t) \kappa_{P_k^p}(t, s) x(s) dt ds, \end{aligned} \quad (21)$$

with $x \in \mathcal{H}^n$. $\square$

Equations (17), (18) and (19) provide the implementation of the Kalman filter for the distributed parameter system (3)-(4). At each time-step the matrix functions $\kappa_{P_k^p}$, κ_{P_k} and the vector function $\bar{K}_k$ are updated through these equations that involve integrals over the spatial domain. In order to reduce the computational burden at run-time it is of interest to perform offline the computation of the asymptotic value of these functions, as discussed in the next section.

3.3 Properties of the centralized Kalman filter

Lemma 3 *Let $\mathbf{S} = \mathbf{S}^* = \mathbf{C}^* \mathbf{R}^{-1} \mathbf{C} : \mathcal{H}^n \rightarrow \mathcal{H}^n$. Then:*

$$\mathbf{K}_k \mathbf{C} = \mathbf{P}_k \mathbf{S}, \quad (22)$$

$$\mathbf{K}_k = \mathbf{P}_k \mathbf{C}^* \mathbf{R}^{-1}. \quad (23)$$

Proof. We start by proving (23). Eq. (15) implies that $\mathbf{P}_k^p \mathbf{C}^* = \mathbf{P}_k \mathbf{C}^* + \mathbf{K}_k \mathbf{C} \mathbf{P}_k^p \mathbf{C}^*$. By rewriting (14) as

$$\mathbf{K}_k \mathbf{C} \mathbf{P}_k^p \mathbf{C}^* + \mathbf{K}_k \mathbf{R} = \mathbf{P}_k^p \mathbf{C}^* = \mathbf{P}_k \mathbf{C}^* + \mathbf{K}_k \mathbf{C} \mathbf{P}_k^p \mathbf{C}^*$$

we obtain $\mathbf{K}_k \mathbf{R} = \mathbf{P}_k \mathbf{C}^*$, that is, (23). Consequently, $\mathbf{K}_k \mathbf{C} = \mathbf{P}_k \mathbf{C}^* \mathbf{R}^{-1} \mathbf{C} = \mathbf{P}_k \mathbf{S}$. $\square$

Lemma 4 *The evolution of the estimation error covariance matrix $\mathbf{P}_k$ obeys the Riccati equation*

$$\begin{aligned} \mathbf{P}_k &= (\mathbf{I} - \mathbf{P}_k \mathbf{S})(\mathbf{A} \mathbf{P}_{k-1} \mathbf{A}^* + \mathbf{Q})(\mathbf{I} - \mathbf{P}_k \mathbf{S})^* \\ &\quad + \mathbf{P}_k \mathbf{S} \mathbf{P}_k. \end{aligned} \quad (24)$$

Proof. Eq. (24) is obtained from the first equation in (B.4) with $\mathbf{K}_k \mathbf{C} = \mathbf{P}_k \mathbf{S}$ and $\mathbf{K}_k \mathbf{R} \mathbf{K}_k^* = \mathbf{P}_k \mathbf{C}^* \mathbf{R}^{-1} \mathbf{C} \mathbf{P}_k = \mathbf{P}_k \mathbf{S} \mathbf{P}_k$. $\square$

The existence of a steady-state solution of (24) can be established as in Hager and Horowitz (1976).

Proposition 1 *If the pairs $(\mathbf{A}, \mathbf{C})$ and $(\mathbf{A}, \mathbf{F})$ are respectively observable and controllable then there exists $\mathbf{P} = \lim_{k \rightarrow \infty} \mathbf{P}_k > 0$ satisfying*

$$\mathbf{P} = (\mathbf{I} - \mathbf{P}\mathbf{S})(\mathbf{A}\mathbf{P}\mathbf{A}^* + \mathbf{Q})(\mathbf{I} - \mathbf{P}\mathbf{S})^* + \mathbf{P}\mathbf{S}\mathbf{P}. \quad (25)$$

The following results are instrumental to deriving the properties of the distributed Kalman filter introduced in the next section. The proofs are reported in Appendix B.

Lemma 5 *In the hypotheses of Proposition 1, let $\mathbf{P} = \lim_{k \rightarrow \infty} \mathbf{P}_k$ and $\bar{\mathbf{A}} := (\mathbf{I} - \mathbf{P}\mathbf{S})\mathbf{A} : \mathcal{H}^n \rightarrow \mathcal{H}^n$. Then the system $z_{k+1} = \bar{\mathbf{A}}z_k$ is asymptotically stable, i.e. $\lim_{k \rightarrow \infty} \|z_k\| = 0$.*

Lemma 6 *In the hypotheses of Proposition 1, let $\mathbf{P} = \lim_{k \rightarrow \infty} \mathbf{P}_k$ and $\mathbf{S} = \mathbf{C}^* \mathbf{R}^{-1} \mathbf{C}$. Then:*

- (1) *For $\alpha \in [0, 1]$ the system $z_{k+1} = (\mathbf{I} - \alpha \mathbf{P}\mathbf{S})z_k$ is simply stable, i.e. $\forall z_0 \exists \bar{z} < \infty : \|z_k\| \leq \bar{z}, \forall k$.*
- (2) *For $\alpha \in (0, 1]$, $\lim_{k \rightarrow \infty} (\mathbf{I} - \alpha \mathbf{P}\mathbf{S})^k z_0 = \Pi z_0 \in \text{Ker}(\mathbf{S})$ where Π is a (non-orthogonal) projector, i.e. an idempotent operator $\Pi : \mathcal{H}^n \rightarrow \text{Ker}(\mathbf{S})$ that does not depend on α ,*

$$\Pi = \lim_{k \rightarrow \infty} (\mathbf{I} - \alpha \mathbf{P}\mathbf{S})^k = \lim_{k \rightarrow \infty} (\mathbf{I} - \mathbf{P}\mathbf{S})^k. \quad (26)$$

- (3) *the adjoint Π^* of Π is an idempotent operator $\Pi^* : \mathcal{H}^n \rightarrow \text{Ker}(\mathbf{S}\mathbf{P})$.*

Lemma 7 *Let the pairs $(\mathbf{A}, \mathbf{C})$ and $(\mathbf{A}, \mathbf{F})$ be observable and controllable respectively, and let Π be the projector defined in (26). Then, for any $x \in \mathcal{H}^n$, $\lim_{k \rightarrow \infty} \|(\Pi \mathbf{A})^k x\| = 0$, that is, $\Pi \mathbf{A}$ generates an asymptotically stable dynamics.*

Lemma 8 *Let Π be the projector defined in (26) and $\mathbf{S} = \sum_{i=1}^N \mathbf{S}_i = \sum_{i=1}^N \mathbf{C}_i^* \mathbf{R}_i^{-1} \mathbf{C}_i$. Then, the following properties hold true: (i) $\mathbf{S}\Pi = 0$; (ii) $\mathbf{S}_i \Pi = 0$; (iii) $\Pi \mathbf{P}\mathbf{S}_i = 0$; (iv) $\Pi \mathbf{P}\mathbf{S} = 0$.*

The following sections will delve into the concepts, algorithms, and applications of the distributed filtering over sensor networks.

4 Kalman filter on L_2 spaces over sensor networks

In the context of sensor networks, each sensor or node is responsible for measuring a transformation of the overall state of the system. The objective for each node is to reconstruct the entire state of the system based solely on the measurements from each own individual sensor and its neighbors. The network of sensors is modeled by an undirected connected graph with N nodes and

Algorithm Broadcast Push-Sum

- 1: In all nodes set $s_{0,i} = 1$ and $w_{0,i} = 0$, except for $w_{0,1} = 1$.
- 2: At step 0 each nodes sends $(s_{0,i}, w_{0,i})$ to itself.
- 3: At step k each node executes:
 1. Let $\{s_r, w_r\}$ be the pairs sent to i in round $k-1$.
 2. Let $s_{k,i} = \sum_r s_r$, $w_{k,i} = \sum_r w_r$.
 3. Send to all neighbors and to i (yourself): $(s_{k,i}/(d_i + 1), w_{k,i}/(d_i + 1))$.
 4. if $w_{k,i} \neq 0$ then $s_{k,i}/w_{k,i}$ is the estimate of N at step k .

Fig. 1. A broadcast version of the *Push-Sum* algorithm for the distributed computation of N , $d_i = |\mathcal{N}^{(i)}|$.

Laplacian matrix $\mathcal{L} = \mathcal{D} - \mathcal{A} \in \mathbb{R}^{N \times N}$, where $\mathcal{D} = \text{diag}_{i=1}^N(d_i)$, $d_i = |\mathcal{N}^{(i)}|$, is the degree matrix and $\mathcal{A}$ is the adjacency matrix. In this framework, the system over a sensor network is represented by

$$x_{k+1} = \mathbf{A}x_k + \mathbf{F}n_k, \quad (27)$$

$$y_k^{(i)} = \mathbf{C}_i x_k + G_i n_k^{(i)}. \quad (28)$$

We assume that some nodes $i \in \mathcal{S}$ have sensing capabilities modeled by the linear bounded operators $\mathbf{C}_i : \mathcal{H}^n \rightarrow \mathbb{R}^{q_i}$. The local measurement noise covariance matrix $\mathbf{R}_i = G_i G_i^\top \in \mathbb{R}^{q_i \times q_i}$ is positive definite and, as in Section 2, $\bar{F}(\xi) G_i^\top = 0$ to ensure uncorrelation between measurement and process noises. Finally, the output noise $n_k^{(i)}$ and $n_k^{(j)}$ are also uncorrelated at each time k . We make the following assumption on the information available at the nodes.

Assumption 1 *Each node knows the kernels of the operators $\mathbf{A}$, $\mathbf{Q}$, $\mathbf{C} = \text{col}_{i \in \mathcal{S}}(\mathbf{C}_i)$ and the matrix $\mathbf{R}_i$. It also knows the kernel of $\mathbf{S} = \sum_{i \in \mathcal{S}} \mathbf{C}_i^* \mathbf{R}_i^{-1} \mathbf{C}_i : \mathcal{H}^n \rightarrow \mathcal{H}^n$ and the number of nodes N .*

The operators $\mathbf{A}$, $\mathbf{Q}$, and $\mathbf{R}_i$ concern the system to be estimated and the local sensor. Conversely, $\mathbf{C}$, $\mathbf{S}$ and N are non-local information that however can be computed once for all in finite time in a pre-processing phase. In particular, notice that the information needed for $\mathbf{S}$ and $\mathbf{C}$ is in essence the same because $\mathbf{S}x \in \mathcal{H}^n$ and

$$(\mathbf{S}x)(\xi) = \sum_{i \in \mathcal{S}} \kappa_{\mathbf{C}_i}^\top(\xi) \int_{\Theta} \mathbf{R}_i^{-1} \kappa_{\mathbf{C}_i}(s) x(s) ds. \quad (29)$$

Consequently, the knowledge of all the kernel functions $\kappa_{\mathbf{C}_i}$ and of the finite dimensional matrices $\mathbf{R}_i \in \mathbb{R}^{q_i \times q_i}$, $i \in \mathcal{S}$, allows to reconstruct both $\mathbf{S}$ and $\mathbf{C}$. It is sufficient to this end for each node to broadcast iteratively to its neighbors the list of known pairs $(\kappa_{\mathbf{C}_i}, \mathbf{R}_i)$ and to update it from received messages. The distributed computation of N is achieved in the pre-processing phase by the broadcast version of the Push-Sum algorithm of Kempe et al. (2003) reported in Fig. 1.

4.1 The algorithm

Let $\hat{x}_{k|k-1}^{(i)}$ and $\hat{x}_k^{(i)}$ be respectively the local prediction and local estimate of x_k at node i . $z_{k,h}^{(i)}$ represent the intermediate consensus estimate at node i and consensus step h . The Kalman filter on L_2 spaces over sensor networks with γ consensus steps iterates over time at each node the following operations

(1) Local prediction:

$$\hat{x}_{k+1|k}^{(i)} = \mathbf{A}\hat{x}_k^{(i)}. \quad (30)$$

(2) Local correction:

$$z_{k+1,0}^{(i)} = \hat{x}_{k+1|k}^{(i)} + \mathbf{K}_i(y_{k+1}^{(i)} - \mathbf{C}_i\hat{x}_{k+1|k}^{(i)}), \quad (31)$$

$$\mathbf{K}_i = \mathbf{N}\mathbf{P}\mathbf{C}_i^* \mathbf{R}_i^{-1}, \quad (32)$$

where $\mathbf{P}$ solves (25). For concision, let $\mathbf{K}_i = 0$, $\mathbf{C}_i = 0$ whenever $i \notin \mathcal{S}$. Thus, $z_{k+1,0}^{(i)} = \hat{x}_{k+1|k}^{(i)}$ when $i \notin \mathcal{S}$.

(3) Consensus iterations. Set $h = 0$ and iterate while $h \leq \gamma$:

3.1 Compute

$$\bar{y}_{k+1,h}^{(i)} = \mathbf{C}z_{k+1,h}^{(i)} \quad (33)$$

and send it to the neighbors. Receive $\bar{y}_{k+1,h}^{(j)}$ and compute

$$z_{k+1,h+1}^{(i)} = z_{k+1,h}^{(i)} + \kappa \mathbf{K} \sum_{j \in \mathcal{N}^{(i)}} (\bar{y}_{k+1,h}^{(j)} - \bar{y}_{k+1,h}^{(i)}), \quad (34)$$

where $\mathbf{K} = \mathbf{P}\mathbf{C}^* \mathbf{R}^{-1}$ and κ is a design parameter.

3.2 Set $h \leftarrow h + 1$.

(4) Set

$$\hat{x}_{k+1}^{(i)} = z_{k+1,\gamma}^{(i)}. \quad (35)$$

At each node the estimate is initialized as

$$\hat{x}_0^{(i)} = \mathbb{E}[x_0] + \mathbf{K}_i(y_0^{(i)} - \mathbf{C}^{(i)}\mathbb{E}[x_0]), \quad (36)$$

where $\mathbf{K}_i$ is defined in (32).

Remark 1 Similarly to consensus on measurement filters, the nodes in this algorithm exchange information that belongs to the measurement space (see (33)). Although the intermediate estimate $z_{k,h}^{(i)}$ is in $\mathcal{H}^n$, $\bar{y}_{k,h}^{(i)}$ belongs to $\mathbb{R}^q$ and is finite dimensional.

4.2 Analysis of the estimation error

Define the local estimation error $\varepsilon_{k,\gamma}^{(i)} = x_k - \hat{x}_k^{(i)}$ and intermediate estimation error $\varepsilon_{k,h}^{(i)} = x_k - z_{k,h}^{(i)}$. The corresponding global errors are $\varepsilon_{k,\gamma} = \text{col}_{i=1}^N(\varepsilon_{k,\gamma}^{(i)}) \in \mathcal{H}^{nN}$ and $\varepsilon_{k,h} = \text{col}_{i=1}^N(\varepsilon_{k,h}^{(i)}) \in \mathcal{H}^{nN}$. Then,

$$\varepsilon_{k+1,0}^{(i)} = (\mathbf{I} - \mathbf{K}_i \mathbf{C}_i)(\mathbf{A}\varepsilon_k^{(i)} + \mathbf{F}n_k) - \mathbf{K}_i \mathbf{G}_i n_{k+1}, \quad (37)$$

$$\bar{y}_{k+1,h}^{(i)} = \mathbf{C}x_{k+1} - \mathbf{C}\varepsilon_{k+1,h}^{(i)}. \quad (38)$$

By using (38), Eq. (34) becomes

$$z_{k+1,h+1}^{(i)} = (\mathbf{I} - \kappa d_i \mathbf{K} \mathbf{C}) z_{k+1,h}^{(i)} + \kappa d_i \mathbf{K} \mathbf{C} x_{k+1} - \kappa \mathbf{K} \mathbf{C} \sum_{j \in \mathcal{N}^{(i)}} \varepsilon_{k+1,h}^{(j)}, \quad (39)$$

$$\varepsilon_{k+1,h+1}^{(i)} = (\mathbf{I} - \kappa d_i \mathbf{K} \mathbf{C}) \varepsilon_{k+1,h}^{(i)} + \kappa \mathbf{K} \mathbf{C} \sum_{j \in \mathcal{N}^{(i)}} \varepsilon_{k+1,h}^{(j)},$$

$$\varepsilon_{k+1,h+1} = \text{col}_i(\mathbf{I} - \kappa d_i \mathbf{K} \mathbf{C}) \varepsilon_{k+1,h} + \kappa \mathbf{A} \text{col}_{i=1}^N(\mathbf{K} \mathbf{C}) \varepsilon_{k+1,h} = \Psi \varepsilon_{k+1,h}, \quad (40)$$

where the operator $\Psi : \mathcal{H}^{nN} \rightarrow \mathcal{H}^{nN}$ is defined, using the notation introduced in (1)–(2), as

$$\Psi = \mathbf{I}_{\mathcal{H}^{nN}} - \kappa \mathcal{L} \text{col}_{i=1}^N(\mathbf{P} \mathbf{S}). \quad (41)$$

From (37) and (40) we obtain, $i = 1, \dots, N$,

$$\varepsilon_{k+1,\gamma} = \Psi^\gamma \left(\tilde{\mathbf{A}} \varepsilon_{k,\gamma} + \tilde{\mathbf{D}} \text{col}_i(\mathbf{F} n_k) - \tilde{\mathbf{G}} \text{col}_i(n_{k+1}) \right), \quad (42)$$

$$\tilde{\mathbf{A}} = \text{col}_i(\mathbf{D}_i \mathbf{A}),$$

$$\tilde{\mathbf{D}} = \text{col}_i(\mathbf{D}_i),$$

$$\mathbf{D}_i = \mathbf{I} - \mathbf{K}_i \mathbf{C}_i = \mathbf{I} - \mathbf{N} \mathbf{P} \mathbf{S}_i,$$

$$\tilde{\mathbf{G}} = \text{col}_{i=1}^N(\mathbf{K}_i \mathbf{G}_i).$$

Since the three terms in (42) are mutually uncorrelated, the evolution of the overall error covariance matrix $\mathbf{P}_k^{(\gamma)} = \mathbb{E}[\varepsilon_{k,\gamma} \varepsilon_{k,\gamma}^*]$ of $\varepsilon_{k,\gamma}$ can be compactly represented as

$$\mathbf{P}_{k+1}^{(\gamma)} = \Psi^\gamma \left(\tilde{\mathbf{A}} \mathbf{P}_k^\gamma \tilde{\mathbf{A}}^* + \tilde{\mathbf{D}} \tilde{\mathbf{Q}} \tilde{\mathbf{D}}^* + \tilde{\mathbf{G}} \tilde{\mathbf{G}}^* \right) \Psi^{\gamma*}, \quad (43)$$

where $\tilde{\mathbf{Q}} : \mathcal{H}^{nN} \rightarrow \mathcal{H}^{nN}$ is the covariance of $\text{col}_{i=1}^N(\mathbf{F} n_k)$. To investigate the error dynamics when the number γ of consensus steps increases we use the following crucial lemma.

Lemma 9 Let $x = \text{col}_{i=1}^N(x_i)$, $x_i \in \mathcal{H}^n$, and let $\Psi : \mathcal{H}^{nN} \rightarrow \mathcal{H}^{nN}$ be as in (41). Let $u_1 := \frac{1}{\sqrt{N}} \mathbf{1}_N$ and $\bar{x} :=$

$\frac{1}{N} \sum_{i=1}^N x_i$. If $\kappa\lambda_k < 1$ for all $\lambda_k \in \sigma(\mathcal{L})$, then

$$\lim_{\gamma \rightarrow \infty} \Psi^\gamma x = \mathcal{P}_1 x + \mathcal{P}_2 x, \quad (44)$$

$$\mathcal{P}_1 x := \text{col}_{i=1}^N(\bar{x}) = u_1 u_1^\top x, \quad (45)$$

$$\mathcal{P}_2 x := \text{col}_{i=1}^N(\Pi(x_i - \bar{x})) = (I_N - u_1 u_1^\top) \text{col}_i(\Pi)x, \quad (46)$$

where Π is defined in (26) with $\alpha = \kappa\lambda_k < 1$. Moreover,

$$[\mathcal{P}_1 x, \mathcal{P}_2 y] = 0 \quad \text{for all } x, y \in \mathcal{H}^{nN}, \quad (47)$$

$$\mathcal{P}_1 = \mathcal{P}_1^*, \quad (48)$$

$$\mathcal{P}_2^* x = \text{col}_{i=1}^N(\Pi^*(x_i - \bar{x})). \quad (49)$$

The proof is reported in Appendix B. We can now state the main theorem.

Theorem 2 *If the pairs $(\mathbf{A}, \mathbf{C})$ and $(\mathbf{A}, \mathbf{F})$ are respectively observable and controllable and $\kappa\lambda_k < 1$ for all $\lambda_k \in \sigma(\mathcal{L})$, then the error covariance at each node becomes identical to the covariance of the centralized Kalman filter when $\gamma \rightarrow \infty$.*

Proof. Based on Lemma 9 when $\gamma \rightarrow \infty$, the aggregate error $\varepsilon_{k+1, \gamma}$ in (42) can be written as

$$\varepsilon_{k+1} = (\mathcal{P}_1 + \mathcal{P}_2) \tilde{\mathbf{A}} \varepsilon_k + \nu_k,$$

where $\nu_k = (\mathcal{P}_1 + \mathcal{P}_2)(\tilde{\mathbf{D}} \text{col}_{i=1}^N(\mathbf{F}n_k) - \tilde{\mathbf{G}} \text{col}_{i=1}^N(n_{k+1}))$. Let $M = M^\top = I_N - u_1 u_1^\top \in \mathbb{R}^{N \times N}$. Then,

$$\mathcal{P}_2 \tilde{\mathbf{A}} x = \text{col}_i(\Pi \mathbf{A}) M x. \quad (50)$$

The proof of (50) descends from (46), the property $\Pi \mathbf{P} \mathbf{S}_i = 0$ (see Lemma 8), and (2):

$$\begin{aligned} \mathcal{P}_2 \tilde{\mathbf{A}} x &= M \text{col}_i(\Pi(I - N \mathbf{P} \mathbf{S}_i) \mathbf{A} x_i) \\ &= M \text{col}_i(\Pi \mathbf{A} x_i) = \text{col}_i(\Pi \mathbf{A}) M x. \end{aligned}$$

M has a base of orthogonal normalized eigenvectors u_i with $u_1 = \frac{1}{\sqrt{N}} \mathbf{1}_N$, and with $U = \text{row}_{i=1}^N(u_i)$, $U^{-1} = U^\top$,

$$U^\top M = \begin{pmatrix} 0 & 0 \\ 0 & I_{N-1} \end{pmatrix} U^\top := D_{01} U^\top,$$

that is, the eigenvalue of u_1 is 0 and the remaining eigenvalues are 1. Consider now the change of coordinates $\eta_k = \text{col}_{i=1}^N(\eta_{i,k}) = U^\top \varepsilon_k$, so that

$$\eta_{k+1} = U^\top (\mathcal{P}_1 + \mathcal{P}_2) \tilde{\mathbf{A}} U \eta_k + U^\top \nu_k.$$

From (50) we obtain

$$\begin{aligned} U^\top \mathcal{P}_2 \tilde{\mathbf{A}} U \eta_k &= U^\top \text{col}_i(\Pi \mathbf{A}) U D_{01} U^\top U \eta_k \\ &= \text{col}_i(\Pi \mathbf{A}) D_{01} \eta_k = \begin{pmatrix} 0 \\ \text{col}_{i=2}^N \Pi \mathbf{A} \eta_{k,i} \end{pmatrix}. \end{aligned}$$

From $U^\top \mathcal{P}_1 = \text{col}(u_1^\top, 0, \dots, 0)$ and $\varepsilon_{k,i} = \frac{1}{\sqrt{N}} \eta_{k,1} + \left(\sum_{\ell=2}^N u_\ell \eta_{k,\ell} \right)_i$ we obtain $U^\top \mathcal{P}_1 \tilde{\mathbf{A}} U \eta_k = \text{col}(\tilde{\eta}_1, 0, \dots, 0)$,

$$\tilde{\eta}_1 := (I - \mathbf{P} \mathbf{S}) \mathbf{A} \eta_{k,1} - \sqrt{N} \sum_{i=1}^N \mathbf{P} \mathbf{S}_i \mathbf{A} \left(\sum_{j=2}^N u_j \eta_{k,j} \right)_i.$$

Moreover, $U^\top \nu_k = \text{col}(\tilde{\nu}_{k,1}, 0, \dots, 0)$,

$$\tilde{\nu}_{k,1} := \sqrt{N} ((I - \mathbf{P} \mathbf{S}) \mathbf{F} n_k - \sum_{i=1}^N \mathbf{P} \mathbf{C}_i^* R_i^{-1} G_i n_{k+1}),$$

where

$$\mathbb{E}[\tilde{\nu}_{k,1} \tilde{\nu}_{k,1}^*] = N((I - \mathbf{P} \mathbf{S}) \mathbf{Q} (I - \mathbf{P} \mathbf{S})^* + \mathbf{P} \mathbf{S} \mathbf{P}) := \mathcal{Q}. \quad (51)$$

Thus the error system in the new coordinates can be represented as

$$\begin{aligned} \eta_{k+1,1} &= (I - \mathbf{P} \mathbf{S}) \mathbf{A} \eta_{k,1} \\ &\quad - \sqrt{N} \left(\sum_{i=1}^N \mathbf{P} \mathbf{S}_i \mathbf{A} \left(\sum_{j=2}^N u_j \eta_{k,j} \right)_i \right) + \tilde{\nu}_{k,1}, \\ \eta_{k+1,i} &= \Pi \mathbf{A} \eta_{k,i}, \quad i = 2, \dots, N. \end{aligned}$$

Since Lemma 7 states that $\Pi \mathbf{A}$ generates an asymptotically stable dynamics, $\eta_{k,i} \rightarrow 0$ for $i \neq 1$. The error dynamics is concentrated in $\eta_{k,1}$ and its dynamics tends asymptotically to

$$\eta_{k+1,1} = (I - \mathbf{P} \mathbf{S}) \mathbf{A} \eta_{k,1} + \tilde{\nu}_{k,1}.$$

Moreover,

$$\varepsilon_k = U \eta_k = u_1 \eta_{k,1} = \text{col}_{i=1}^N \left(\frac{1}{\sqrt{N}} \eta_{k,1} \right).$$

The asymptotic dynamics of $\varepsilon_{k,i}$ at the generic node i is therefore

$$\varepsilon_{k+1,i} = (I - \mathbf{P} \mathbf{S}) \mathbf{A} \varepsilon_{k,i} + \frac{1}{\sqrt{N}} \tilde{\nu}_{k,1}.$$

From (51), (25) and Lemma 5 it is easy to verify that, for all (i, j) and as $k \rightarrow +\infty$, $\mathbb{E}[\varepsilon_{k,i} \varepsilon_{k,j}^*] \rightarrow \mathbf{P}$ solution of the algebraic Riccati equation (25) of the centralized Kalman filter. $\square$

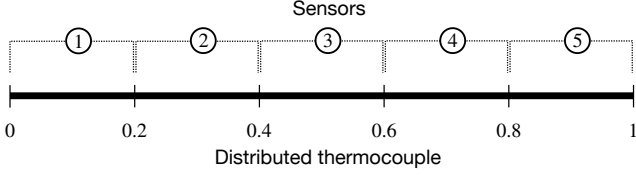


Fig. 2. Sketch of the distributed thermocouple and sensors.

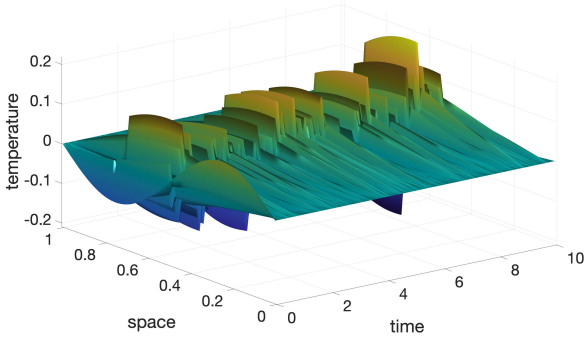


Fig. 3. Sample representation of the discrete-time evolution of the heated bar with boundary conditions $x(t, 0) = x(t, 1) = 0$ and initial condition $x_0(\xi) = 0.1 \sin(2\pi\xi)$, $\xi \in [0, 1]$.

Since in practice $\gamma < \infty$, we can use (42) to state the following sufficient condition.

Theorem 3 *If the pairs (A, C) and (A, F) are respectively observable and controllable and $\kappa\lambda_k < 1$ for all $\lambda_k \in \sigma(\mathcal{L})$, then the error covariance at each node is uniformly bounded when γ is such that the operator $\Psi^\gamma \tilde{A}$ is asymptotically stable, i.e. for all $x_0 \in \mathcal{H}^{nN}$, $\lim_{k \rightarrow \infty} \|(\Psi^\gamma \tilde{A})^k x_0\| = 0$.*

5 Numerical example

We illustrate an application of the consensus Kalman Filter on $L_2([0, 1])$ to the conceptual design of a distributed thermocouple that solves the problem of estimating the temperature profile of a bar of length $\ell = 1$ from measurements of the voltage when a current flows through it. The centralized Kalman filter (CKF), namely the filter that employs the measurements y_i of all the sensors, is given by (10), while the distributed Kalman consensus filter (DKF) for each sensor is given by the algorithm in Section 4.1 (equations (30)–(35)). Specifically, given the thermocouple distributed on the spatial domain $[0, 1]$, we consider 5 sensors positioned in the segments $[0, 0.2]$, $[0.2, 0.4]$, $[0.4, 0.6]$, $[0.6, 0.8]$, $[0.8, 1]$. 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 $m = 1, 2, \dots$, constitute an orthonormal basis for $L_2([0, 1])$. Given $x = x(t, \xi)$ we

have

$$\frac{\partial x}{\partial t} = \frac{\partial^2 x}{\partial \xi^2} \quad \text{with} \quad x(t, 0) = x(t, 1) = 0 \quad \forall t. \quad (52)$$

The solution to (52) with the initial condition

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

is given by

$$x(t, \xi) = \sum_{m=1}^{+\infty} e^{-t\lambda_m} [\phi_m, x_0] \phi_m(\xi), \quad (53)$$

where $\lambda_m = (m\pi)^2$ (Curtain and Zwart (2012)). Thus, by the definition of the inner product, the solution (53) is

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

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

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

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

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

where we recognize the kernel of the underlying operator A

$$\kappa_A(\xi, \eta) = \sum_{m=1}^{+\infty} e^{-\Delta\lambda_m} \phi_m(\xi) \phi_m(\eta). \quad (54)$$

We consider the case in which a current source is applied to the bar of length $\ell = 1$ and sensors measure the voltage at the endpoints of segments as in Fig. 2. Thus, the measurement $z_i(t)$ of the sensor i -th is given by $z_i(t) = I r_i(t)$ where I is the applied current intensity, and $r_i(t)$ is the resistance of the bar that depends on the temperature through the equation

$$r_i(t) = \int_{a_i}^{b_i} \beta(1 + \alpha x(t, \xi)) d\xi,$$

where α and β are some positive coefficients depending on the bar, and a_i, b_i are the endpoints of the segment

of sensor i -th (*e.g.* for sensor 3, $a_i = 0.4$, $b_i = 0.6$). By setting $y_i(t) := z_i(t) - c_0$, with $c_0 = 0.2I\beta$, and $c = I\beta\alpha$ we can define the new measurement equation

$$y_i(t) = \int_{a_i}^{b_i} c x(t, \xi) d\xi. \quad (55)$$

In practice, measurements are sampled in time and thus it is reasonable to consider the discrete-time model associated to (55), which, with the previously defined sequence $\{t_k\}$ and discretization step Δ , and the addition of an output noise term $G_i n_k$, results into

$$y_k^{(i)} = \int_{a_i}^{b_i} c x_k(\xi) d\xi + G_i n_k^{(i)} = C_i x_k + G_i n_k^{(i)},$$

where we recognize the kernels κ_{C_i} of the operators C_i :

$$\kappa_{C_i}(\xi) = \begin{cases} c & \xi \in [a_i, b_i], \\ 0 & \text{otherwise.} \end{cases}$$

We also assume $n_k \in \mathbb{R}^2$ to be a Gaussian white noise vector, $F = (f(\xi), 0)$ and $G_i = (0, g_i)$ (such that $F G_i^\top = 0$). We suppose that the bar is affected by a random thermal noise for $\xi \in [0.5, 0.75]$. The state and measurement noise amplitudes are

$$f(\xi) = \begin{cases} 0.2 \cdot \xi(\xi - 1), & \xi \in [0.5, 0.75] \\ 0, & \text{otherwise} \end{cases}$$

$$g_i = 0.1 \quad \text{for all } i \in \{1, \dots, 5\}.$$

Moreover, the final simulations time is $\bar{t} = 10$, $\Delta_t = 2 \cdot 10^{-2}$ (*i.e.* 501 time points), while the space is approximated with increment $\Delta_\xi = 10^{-2}$ (*i.e.* 101 space points). The measurement parameters are $I = 5$, $\alpha = \beta = 1$, that yield $c = 5$. The initial condition is set to

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

In order to obtain an efficient implementation of the filter, we followed the steps of Corollary 1. Note that, the gain K_i of Eq. (32) can be expressed through $\bar{K}_i : \Theta \rightarrow \mathbb{R}$ as for the operator K through the gain $\bar{K} : \Theta \rightarrow \mathbb{R}^5$ of Eq. (18). In fact, each node compute off-line the gain $\bar{K} : \Theta \rightarrow \mathbb{R}^5$ (eq. (18)) and $\bar{K}_i : \Theta \rightarrow \mathbb{R}$ by iterating the computation of P_k^p , $\bar{K}_k$, and P_k through (17), (18), (19). The existence of K and K_i descends from the existence of P (Proposition 1), since system (3) (without noise) with κ_A given by (54) is asymptotically stable. In the numerical computation we took the first 10 eigenfunctions ϕ_m (*e.g.*, $m = 1, 2, \dots, 10$) and we used the following terminal condition for the numerical solution of the iterations (17)–(18)–(19) ($\kappa_{P_k^p}$, $\bar{K}_k$, and

Table 1

The agents' execution time of the proposed filter grows linearly with γ (number of consensus step iterations).

$\gamma = 1$	$\gamma = 3$	$\gamma = 6$	$\gamma = 20$	$\gamma = 50$	$\gamma = 150$
2.7s	2.9s	3.4s	5.2s	9.3s	22.9s

κ_{P_k})

$$\left(\sum_{i,j=1}^N (P_{k+1}^d(\xi_i, \xi_j) - P_k^d(\xi_i, \xi_j))^2 \right)^{\frac{1}{2}}$$

$$\leq 10^{-10} \cdot \left(\sum_{i,j=1}^N P_k^d(\xi_i, \xi_j)^2 \right)^{\frac{1}{2}}$$

where P_k^d denotes the discretized approximation of P_k . With the specified parameters, the terminal condition is reached after 32 iterations, requiring 10 seconds when implemented in MATLAB[®] on a standard laptop. The profile of the approximated steady-state Kalman gain $\bar{K}(\xi)$ obtained through this procedure is shown in Fig. 4 (left). Moreover, we test the numerical accuracy of the estimate of the CKF by comparing in Fig. 4 (right) the value of $\kappa_P(\xi, \xi)$, theoretical variance of the estimation error at point $\xi \in [0, 1]$, with the average mean square error (MSE) across $N_s = 100$ Monte Carlo simulations, computed as

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

where $N_t = \bar{t}/\Delta_t + 1 = 501$ is the number of time samples, $x_k^{\{i\}}(\xi)$ and $\hat{x}_k^{\{i\}}(\xi)$ denote the true and estimated value at point ξ and time t_k for the i -th simulation, and $\tau = 50$ has been chosen to exclude the transient. The plot shows almost perfect agreement between the theoretical and actual error variance throughout the bar.

We now show the performance of the distributed Kalman consensus filter (DKF) of Section 4.1 (equations (30)–(35)). We set $\kappa = \frac{1}{5}$ (Equation (34)) and we show the performance in terms of MSE across $N_s = 10$ Monte Carlo runs of each node for different values of γ of the consensus step iterations. Specifically, Fig. 5 shows the MSE of each node i (MSE_i) of the DKF for the values $\gamma \in \{1, 3, 6, 20, 50, 150\}$. The figure clearly confirms that by increasing the number γ of the consensus step iterations, the error covariance at each node converges to the covariance of the centralized Kalman filter, thereby validating the result of Theorem 2. Finally, the elapsed time (seconds) to execute one simulation of the DKF for the values of γ is summarized in the table below.

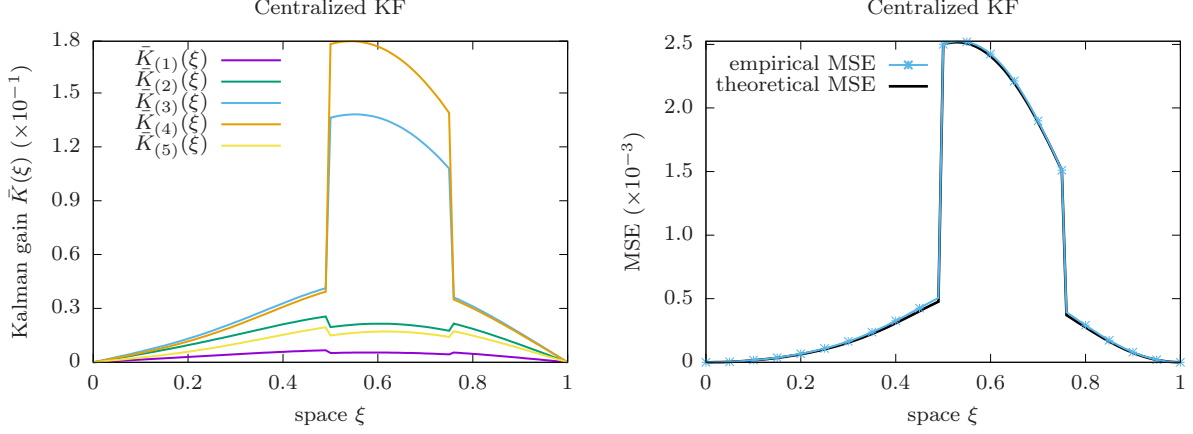


Fig. 4. Components $\bar{K}_{(j)}(\xi) \in \mathbb{R}$ for $j \in \{1, \dots, 5\}$ of the steady-state Kalman gain $\bar{K}(\xi) \in \mathbb{R}^5$ defined in (18) (left panel). Agreement between theoretical and empirical MSE across 100 Monte Carlo runs (right panel).

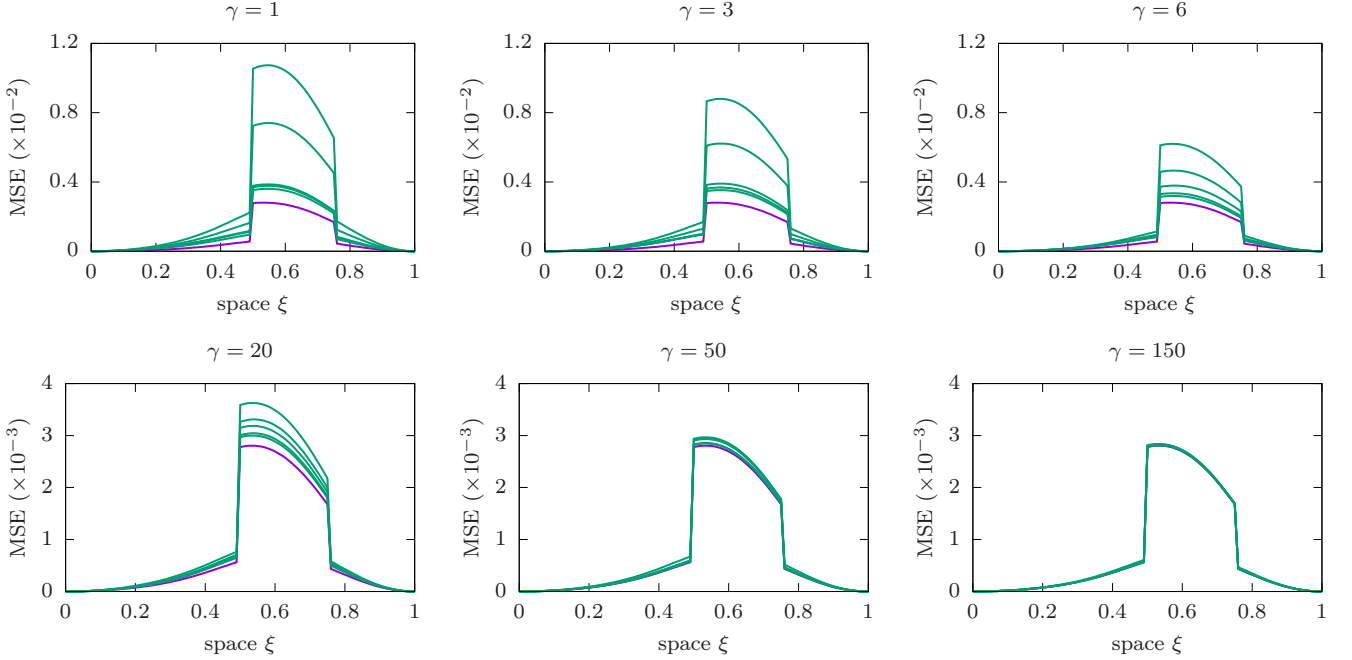


Fig. 5. MSE_i across 10 Monte Carlo runs of each node $i \in \{1, \dots, 5\}$ of the DKF for the values $\gamma \in \{1, 3, 6, 20, 50, 150\}$ over the interval $[0, 1]$. The purple line is the covariance of the centralized Kalman filter.

6 Conclusions

The approach presented in this paper provides an algorithm for the cooperative estimation of systems described by partial differential equations that is straightforward to implement on digital devices. The integration of the system over a discretization interval yields a discrete-time system that is amenable of being processed by a distributed filter. There are several extensions already proposed in the finite-dimensional case that are worth investigating in this context, for example to the case of switching or randomly varying network connec-

tions (Matei et al., 2008; Gan et al., 2024), random packet losses (Battilotti et al., 2023a) and distributed output feedback control Battilotti et al. (2023b).

Appendix

A Spaces and covariance operators

Let:

- $\mathcal{H}^{n \times m} = L_2(\Theta; \mathbb{R}^{n \times m})$, $n \in \mathbb{N}$, be the Hilbert space of square integrable, with respect to the standard

Lebesgue measure, matrix functions from a compact domain $\Theta \subset \mathbb{R}^s$ to $\mathbb{R}^{n \times m}$. $\mathcal{H}^{n \times m}$ is endowed with the inner product $[A, B]_{\mathcal{H}^{n \times m}} = \int_{\Theta} \text{tr}(A(s)B^\top(s)) ds$.

- $\mathcal{H}_{(2)}^{n \times m} = L_2(\Theta \times \Theta; \mathbb{R}^{n \times m})$, $n, m \in \mathbb{N}$, be the Hilbert space of matrix functions from a compact domain $\Theta \times \Theta \subset \mathbb{R}^s \times \mathbb{R}^s$ to $\mathbb{R}^{n \times m}$. $\mathcal{H}_{(2)}^{n \times m}$ is endowed with the inner product $[A, B]_{\mathcal{H}_{(2)}^{n \times m}} = \int_{\Theta \times \Theta} \text{tr}(A(s, t)B^\top(s, t)) ds dt$.

Given a random function $x \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$, let us denote $\hat{x} = x - \mathbb{E}[x]$ the corresponding 0-mean function. For any two elements $z, w \in \mathcal{H}^n$, we have that $\mathbb{E}[\hat{x}, z]_{\mathcal{H}^n} [\hat{x}, w]_{\mathcal{H}^n}$ is a bilinear form over $\mathcal{H}^n$ which is bounded. Hence, there exists a bounded linear self-adjoint non-negative operator $\mathbf{P} : \mathcal{H}^n \rightarrow \mathcal{H}^n$ such that $\mathbb{E}[\hat{x}, z]_{\mathcal{H}^n} [\hat{x}, w]_{\mathcal{H}^n} = [\mathbf{P}z, w]_{\mathcal{H}^n}$ which we shall term the *covariance* of x (Balakrishnan, 2012).

Definition 5 The covariance of $x \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$, denoted $\mathbb{E}[xx^*]$, is the bounded linear self-adjoint non-negative operator $\mathcal{H}^n \rightarrow \mathcal{H}^n$ such that, with $\hat{x} = x - \mathbb{E}[x]$, $\mathbb{E}[\hat{x}, z]_{\mathcal{H}^n} [\hat{x}, w]_{\mathcal{H}^n} = [\mathbb{E}[xx^*]z, w]_{\mathcal{H}^n}$ for any $z, w \in \mathcal{H}^n$. The covariance of two random functions $x \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$, $y \in \mathcal{W}_{\mathcal{H}^m}^{\mathcal{F}}$, denoted $\mathbb{E}[xy^*]$, is the bounded linear operator $\mathcal{H}^m \rightarrow \mathcal{H}^n$ such that $\mathbb{E}[\hat{x}, z]_{\mathcal{H}^n} [\hat{y}, w]_{\mathcal{H}^m} = [z, \mathbb{E}[xy^*]w]_{\mathcal{H}^n}$ for any $z \in \mathcal{H}^n$ and $w \in \mathcal{H}^m$.

One can readily verify that $\mathbb{E}[xy^*]^* = \mathbb{E}[yx^*]$.

Definition 6 Given $x \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$ and $y \in \mathcal{W}_{\mathcal{H}^m}^{\mathcal{F}}$, we say that x and y are uncorrelated and write $x \perp y$ when $\mathbb{E}[xy^*] = 0$.

It is easy to check that if $x, y \in \mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}$ are 0-mean, then $x \perp y$ if and only if $[x, y]_{\mathcal{W}_{\mathcal{H}^n}^{\mathcal{F}}} = 0$.

Lemma 10 If $x \in \mathcal{W}_{\mathcal{H}}^{\mathcal{F}}$ has covariance $\mathbf{P}$ and $\mathbf{A} : \mathcal{H} \rightarrow \mathcal{H}$ then $\mathbf{A}x$ has covariance $\mathbf{A}\mathbf{P}\mathbf{A}^*$.

Proof. Since for all z, w : $\mathbb{E}[[x, z][x, w]] = [\mathbf{P}z, w]$, then

$$\mathbb{E}[[\mathbf{A}x, z][\mathbf{A}x, w]] = \mathbb{E}[[x, \mathbf{A}^*z][x, \mathbf{A}^*w]] = [\mathbf{P}\mathbf{A}^*z, \mathbf{A}^*w] = [\mathbf{A}\mathbf{P}\mathbf{A}^*z, w].$$

Lemma 11 If $x, y \in \mathcal{W}_{\mathcal{H}}^{\mathcal{F}}$, $\mathbf{M}, \mathbf{N} : \mathcal{H} \rightarrow \mathcal{H}$ and $x \perp y$, then $\mathbf{M}x \perp \mathbf{N}y$.

Proof. Since for all z, w : $\mathbb{E}[[x, z][y, w]] = 0$, then

$$\mathbb{E}[[\mathbf{M}x, z][\mathbf{N}y, w]] = \mathbb{E}[[x, \mathbf{M}^*z][y, \mathbf{N}^*w]] = 0.$$

B Proofs of auxiliary results

B.1 Proof of Theorem 1

Eq. (12) follows from $e_k^p = \mathbf{A}e_k + \mathbf{F}n_k$ and $e_{k-1} \perp n_k$. Since $\mathbf{A}$ is bounded, $\mathbf{A}\mathbf{P}\mathbf{A}^*$ and $\mathbf{P}_k^p$ are bounded whenever $\mathbf{P}_{k-1}$ is bounded. Further, since $\nu_k^o = \mathbf{C}\mathbf{A}e_{k-1} + \mathbf{C}\mathbf{F}n_{k-1} + Gn_k$ and the three terms are uncorrelated,

$$\begin{aligned} P_k^o &= \mathbb{E}[\nu_k^o \nu_k^{o*}] = \mathbf{C}(\mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^*)\mathbf{C}^* + \mathbf{C}\mathbf{Q}\mathbf{C}^* + R \\ &= \mathbf{C}\mathbf{P}_k^p\mathbf{C}^* + R. \end{aligned} \quad (\text{B.1})$$

From (11) it follows that

$$P_k^{s,o} = \mathbb{E}[\nu_k^s \nu_k^{o*}] = \mathbb{E}[\mathbf{K}_k \nu_k^o \nu_k^{o*}] = \mathbf{K}_k P_k^o. \quad (\text{B.2})$$

Further, (11) yields also $\nu_k^s = \mathbf{A}e_{k-1} + \mathbf{F}n_{k-1} - e_k$ and, since $e_k \perp \nu_k^o$, $e_{k-1} \perp n_k$ and $n_{k-1} \perp n_k$,

$$\begin{aligned} P_k^{s,o} &= \mathbb{E}[\nu_k^s \nu_k^{o*}] = \mathbb{E}[(\mathbf{A}e_{k-1} + \mathbf{F}n_{k-1}) \\ &\quad \cdot (\mathbf{C}\mathbf{A}e_{k-1} + \mathbf{C}\mathbf{F}n_{k-1} + Gn_k)^*] \\ &= \mathbf{A}\mathbf{P}\mathbf{A}^*\mathbf{C}^* + \mathbf{F}\mathbf{F}^*\mathbf{C}^* = \mathbf{P}_k^p\mathbf{C}^*. \end{aligned} \quad (\text{B.3})$$

Comparing (B.2) and (B.3) we obtain (14). Finally, P_k^o is non-singular because is the sum of the positive definite matrix R and the semi-positive definite matrix $\mathbf{C}\mathbf{P}_k^p\mathbf{C}^*$, hence we can write (16). Standard manipulations yield

$$e_k = (\mathbf{I} - \mathbf{K}_k\mathbf{C})\mathbf{A}e_{k-1} + (\mathbf{I} - \mathbf{K}_k\mathbf{C})\mathbf{F}n_{k-1} - \mathbf{K}_kGn_k.$$

By noticing that the three terms are pairwise uncorrelated and $\mathbf{P}_k^p\mathbf{C}^* = \mathbf{K}_kP_k^o$ we obtain

$$\begin{aligned} P_k &= (\mathbf{I} - \mathbf{K}_k\mathbf{C})(\mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^* + \mathbf{Q})(\mathbf{I} - \mathbf{K}_k\mathbf{C})^* \\ &\quad + \mathbf{K}_kR\mathbf{K}_k^* \\ &= (\mathbf{I} - \mathbf{K}_k\mathbf{C})\mathbf{P}_k^p(\mathbf{I} - \mathbf{K}_k\mathbf{C})^* + \mathbf{K}_kR\mathbf{K}_k^* \\ &= (\mathbf{I} - \mathbf{K}_k\mathbf{C})\mathbf{P}_k^p - \mathbf{P}_k^p\mathbf{C}^*\mathbf{K}_k^* + \mathbf{K}_kP_k^o\mathbf{K}_k^* \\ &= (\mathbf{I} - \mathbf{K}_k\mathbf{C})\mathbf{P}_k^p. \end{aligned} \quad (\text{B.4})$$

□

B.2 Proof of Lemma 5

The statement is equivalent to proving that the dual system $w_{k+1} = \bar{\mathbf{A}}^*w_k$ is asymptotically stable. With the Lyapunov function $V(w) = [w, \mathbf{P}w]_{\mathcal{H}^n}$ we obtain

$$\begin{aligned} V(w_{k+1}) - V(w_k) &= [w_k, (\bar{\mathbf{A}}\mathbf{P}\bar{\mathbf{A}}^* - \mathbf{P})w_k] \\ &= -[w_k, \mathbf{M}_\mathbf{P}w_k], \\ \mathbf{M}_\mathbf{P} &:= (\mathbf{I} - \mathbf{P}\mathbf{S})\mathbf{Q}(\mathbf{I} - \mathbf{P}\mathbf{S})^* + \mathbf{P}\mathbf{S}\mathbf{P}. \end{aligned}$$

The self-adjoint operator $\mathbf{M}_\mathbf{P}$ is the sum of two semi-definite positive self-adjoint operators and it is positive

definite, hence $V(w_{k+1}) - V(w_k) < 0$. In fact, if $\mathbf{S}\mathbf{P}w = 0$ we get

$$\begin{aligned} & [w, ((\mathbf{I} - \mathbf{P}\mathbf{S})\mathbf{Q}(\mathbf{I} - \mathbf{P}\mathbf{S})^* + \mathbf{P}\mathbf{S}\mathbf{P})w] \\ &= [w, ((\mathbf{I} - \mathbf{S}\mathbf{P})\mathbf{Q}w)] = [w, \mathbf{Q}w] - [w, \mathbf{P}\mathbf{S}\mathbf{Q}w] \\ &= [w, \mathbf{Q}w] - [\mathbf{S}\mathbf{P}w, \mathbf{Q}w] = [w, \mathbf{Q}w] > 0. \end{aligned}$$

Conversely, if $\mathbf{S}\mathbf{P}w \neq 0$, since $\mathbf{P} > 0$ and $\mathbf{S} \geq 0$,

$$[w, \mathbf{P}\mathbf{S}\mathbf{P}w] = [\mathbf{S}^{\frac{1}{2}}\mathbf{P}w, \mathbf{S}^{\frac{1}{2}}\mathbf{P}w] > 0,$$

and since $(\mathbf{I} - \mathbf{S}\mathbf{P})\mathbf{Q}(\mathbf{I} - \mathbf{S}\mathbf{P})^* \geq 0$ we conclude that $[w, \mathbf{M}\mathbf{P}w]_{\mathcal{H}^n} > 0$. $\square$

B.3 Proof of Lemma 6

(1) With the Lyapunov function $V(z) = [z, \mathbf{P}^{-1}z]_{\mathcal{H}^n}$ we obtain

$$\Delta V(z_k) = V(z_{k+1}) - V(z_k) = [z_k, -\alpha(2\mathbf{S} - \alpha\mathbf{S}\mathbf{P}\mathbf{S})z_k]_{\mathcal{H}^n}. \quad (\text{B.5})$$

Since $\mathbf{P} = \mathbf{P}^* > 0$, $\mathbf{S} = \mathbf{S}^* \geq 0$, $\mathbf{S}\mathbf{P}\mathbf{S} \geq 0$, $0 \leq \alpha \leq 1$,

$$2\mathbf{S} - \alpha\mathbf{S}\mathbf{P}\mathbf{S} \geq \mathbf{S} - \mathbf{S}\mathbf{P}\mathbf{S} = \mathbf{S}^{\frac{1}{2}}(\mathbf{I} - \mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}})\mathbf{S}^{\frac{1}{2}}, \quad (\text{B.6})$$

it is sufficient to prove $\mathbf{I} - \mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}} \geq 0$, where $\mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}}$ is self-adjoint and positive semi-definite. A necessary and sufficient condition for $\mathbf{I} - \mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}} \geq 0$ is that $\|\mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}}\| \leq 1$, because $\forall x$

$$\begin{aligned} & [x, (\mathbf{I} - \mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}})x] \geq 0 \Rightarrow \\ & \|x\|^2 \geq [x, \mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}}x] = \|\mathbf{P}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}x\|^2. \end{aligned}$$

Consequently $\|\mathbf{P}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}\| \leq 1$. But

$$\|\mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}}\| = \|(\mathbf{P}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}})^*(\mathbf{P}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}})\| \leq \|\mathbf{P}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}\|^2 \leq 1,$$

where we have used the property that, for any bounded $\mathbf{M}$, $\|\mathbf{M}\mathbf{M}^*\| = \|\mathbf{M}^*\mathbf{M}\|$. Hence, positive definiteness of $\mathbf{I} - \mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}}$ is equivalent to $\|\mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}}\| \leq 1$. Conversely, from $\|\mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}}\| \leq 1$ it follows

$$\begin{aligned} & [x, (\mathbf{I} - \mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}})x] = \|x\|^2 - [x, \mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}}x] \\ & \geq \|x\|^2 - \|\mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}}\| \|x\|^2 \geq 0. \end{aligned}$$

The Riccati equation implies that $\mathbf{P} - \mathbf{P}\mathbf{S}\mathbf{P} = \mathbf{P}^{\frac{1}{2}}(\mathbf{I} - \mathbf{P}^{\frac{1}{2}}\mathbf{S}\mathbf{P}^{\frac{1}{2}})\mathbf{P}^{\frac{1}{2}} \geq 0$, that is, $\mathbf{I} - \mathbf{P}^{\frac{1}{2}}\mathbf{S}\mathbf{P}^{\frac{1}{2}} \geq 0$. In view of the above, it must be $\|\mathbf{P}^{\frac{1}{2}}\mathbf{S}\mathbf{P}^{\frac{1}{2}}\| \leq 1$. However, by setting $\mathbf{M} = \mathbf{P}^{\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}$, we have $\mathbf{P}^{\frac{1}{2}}\mathbf{S}\mathbf{P}^{\frac{1}{2}} = \mathbf{M}\mathbf{M}^*$ and $\mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}} = \mathbf{M}^*\mathbf{M}$ and we conclude that $\|\mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}}\| =$

$\|\mathbf{P}^{\frac{1}{2}}\mathbf{S}\mathbf{P}^{\frac{1}{2}}\| \leq 1$, hence $\mathbf{I} - \mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}} \geq 0$ and $\Delta V(z_k) \leq 0$ that proves (i).

(2) As for the second property, it is enough to observe that the sequence $V(z_k)$ is positive and non-increasing and it hence has a limit for $k \rightarrow \infty$. In the limit, $\Delta V(z_k) \rightarrow 0$ and we show that this is possible if and only if $\lim_{k \rightarrow \infty} z_k \in \text{Ker}(\mathbf{S})$. Sufficiency descends from (B.5). Necessity is easy to prove since $\Delta V(z) = 0$ implies

$$[z, (2\mathbf{S} - \alpha\mathbf{S}\mathbf{P}\mathbf{S})z] = [\mathbf{S}^{\frac{1}{2}}z, (2\mathbf{I} - \mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}})\mathbf{S}^{\frac{1}{2}}z] = 0.$$

Since we proved $\mathbf{I} - \mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}} \geq 0$ it follows that $2\mathbf{I} - \mathbf{S}^{\frac{1}{2}}\mathbf{P}\mathbf{S}^{\frac{1}{2}} > 0$ and $\Delta V(z) = 0$ only when $z \in \text{Ker}(\mathbf{S}^{\frac{1}{2}}) = \text{Ker}(\mathbf{S}) = \text{Ker}(\mathbf{P}\mathbf{S})$. Finally, we prove that $\lim_{k \rightarrow \infty} z_k$ does not depend on $\alpha \in (0, 1]$. Since $\mathbf{S}$ is self-adjoint and has finite rank, we can represent $\mathbf{S}$ and $\mathbf{P}\mathbf{S}$ as

$$\mathbf{S}x = \sum_{i=1}^q [x, \phi_i^S] \lambda_i^S \phi_i^S, \quad (\text{B.7})$$

$$\mathbf{P}\mathbf{S}x = \sum_{i=1}^q [x, \phi_i^S] \lambda_i^S \phi_i^{PS}, \quad (\text{B.8})$$

where ϕ_i^S , $i = 1, \dots, q$ are orthogonal eigenfunctions of $\mathbf{S}$ and $\phi_i^{PS} = \mathbf{P}\phi_i^S$ are a set of non-orthogonal functions in $\mathcal{H}^n$. Let $V^S = \text{span}_{i=1}^q(\phi_i^{PS})$. For any $x \in \mathcal{H}^n$, it is always possible to represent it as $x = x_1 + x_2$ with $x_1 \in V^S$ and $x_2 \in \text{Ker}(\mathbf{S})$. To this end it is sufficient to choose $x_1 \in \text{span}_{i=1}^q(\phi_i^{PS})$ as

$$\begin{aligned} & x_1 = \sum_{i=1}^q b_i \phi_i^{PS}, \quad x_2 = x - x_1, \\ & b = \text{col}_{i=1}^q(b_i) = \Phi^{-1} \text{col}_{i=1}^q([x, \phi_i^S]), \\ & (\Phi)_{ij} = [\phi_j^S, \phi_i^{PS}], \end{aligned}$$

where Φ is non-singular because $\mathbf{P}$ is positive definite, and check that $\mathbf{S}x = \mathbf{S}x_1$, $\mathbf{S}x_2 = 0$. Hence,

$$(\mathbf{I} - \alpha\mathbf{P}\mathbf{S})x = (\mathbf{I} - \alpha\mathbf{P}\mathbf{S})x_1 + x_2.$$

Since $x_1 \in V^S$ implies $\mathbf{P}\mathbf{S}x_1 \in V^S$ and $(\mathbf{I} - \alpha\mathbf{P}\mathbf{S})x_1 \in V^S$,

$$(\mathbf{I} - \alpha\mathbf{P}\mathbf{S})^\gamma x = (\mathbf{I} - \alpha\mathbf{P}\mathbf{S})^\gamma x_1 + x_2.$$

By using the Lyapunov function of point (i) we conclude that $\lim_{\gamma \rightarrow \infty} (\mathbf{I} - \alpha\mathbf{P}\mathbf{S})^\gamma x_1 = 0$, and, consequently,

$$\lim_{\gamma \rightarrow \infty} (\mathbf{I} - \alpha\mathbf{P}\mathbf{S})^\gamma x = x_2.$$

Observing that x_2 does not depend on α , we conclude that the same holds for the operator $\mathbf{\Pi}$.

(3) We prove the last property similarly by considering the adjoint system $w_{k+1} = (\mathbf{I} - \mathbf{S}\mathbf{P})w_k$ and by noticing

that any $y \in \mathcal{H}^n$ can be written as $y = y_1 + y_2$ with $y_1 \in \text{span}(\phi_i^S)$ and $y_2 \in \text{Ker}(\mathbf{S}\mathbf{P})$ so that $(\mathbf{I} - \mathbf{S}\mathbf{P})y_1 \in \text{span}(\phi_i^S)$ and $(\mathbf{I} - \mathbf{S}\mathbf{P})y_2 = y_2$ and therefore $\mathbf{\Pi}^*y = y_2$.

B.4 Proof of Lemma 7

Lemma 6 can be proved by a Lyapunov argument with the function $V(z) = [z, \mathbf{P}z]_{\mathcal{H}^n}$ on the auxiliary dynamic system $z_{k+1} = \mathbf{A}^*\mathbf{\Pi}^*z_k$. Thus, by rewriting the Riccati equation (25) as

$$\bar{\mathbf{A}}\mathbf{P}\bar{\mathbf{A}}^* - \mathbf{P} = -(\mathbf{I} - \mathbf{P}\mathbf{S})\mathbf{Q}(\mathbf{I} - \mathbf{P}\mathbf{S})^\top - \mathbf{P}\mathbf{S}\mathbf{P}, \quad (\text{B.9})$$

by the property (iv) of Lemma 8 we have $\mathbf{\Pi}(\mathbf{I} - \mathbf{P}\mathbf{S}) = \mathbf{\Pi}$, and thus we obtain

$$\begin{aligned} \Delta V(z_k) &= V(z_{k+1}) - V(z_k) \\ &= [z_k, (\mathbf{\Pi}\bar{\mathbf{A}}\mathbf{P}\bar{\mathbf{A}}^*\mathbf{\Pi}^* - \mathbf{P})z_k]_{\mathcal{H}^n} \\ &= [z_k, (-\mathbf{P} + \mathbf{\Pi}\mathbf{P}\mathbf{\Pi}^* - \mathbf{\Pi}\mathbf{Q}\mathbf{\Pi}^*)z_k]_{\mathcal{H}^n}. \end{aligned}$$

To prove that the operator on the right-hand side is negative definite, we first show that $\mathbf{\Pi}\mathbf{P}\mathbf{\Pi}^* = \mathbf{P}\mathbf{\Pi}^*$. In fact, $\forall x \in \mathcal{H}^n$, $\mathbf{\Pi}^*x \in \text{Ker}(\mathbf{S}\mathbf{P})$ (Lemma 6, iii), that is, $\mathbf{P}\mathbf{\Pi}^*x \in \text{Ker}(\mathbf{S})$. Since $\mathbf{\Pi}$ is a projector of $\text{Ker}(\mathbf{S})$, $\mathbf{\Pi}\mathbf{P}\mathbf{\Pi}^*x = \mathbf{P}\mathbf{\Pi}^*x$, that is, $\mathbf{\Pi}\mathbf{P}\mathbf{\Pi}^* = \mathbf{P}\mathbf{\Pi}^*$. Being $\mathbf{\Pi}\mathbf{P}\mathbf{\Pi}^*$ self-adjoint, $\mathbf{\Pi}\mathbf{P}\mathbf{\Pi}^* = \mathbf{\Pi}\mathbf{P} = \mathbf{P}\mathbf{\Pi}^*$. Thus,

$$\begin{aligned} \mathbf{P} - \mathbf{\Pi}\mathbf{P}\mathbf{\Pi}^* &= \mathbf{P} - \mathbf{\Pi}\mathbf{P} - \mathbf{P}\mathbf{\Pi}^* + \mathbf{\Pi}\mathbf{P}\mathbf{\Pi}^* \\ &= (\mathbf{I} - \mathbf{\Pi})\mathbf{P}(\mathbf{I} - \mathbf{\Pi})^* \geq 0, \end{aligned}$$

and $[(-\mathbf{P} + \mathbf{\Pi}\mathbf{P}\mathbf{\Pi}^*)z, z] \leq 0$ with the equality that holds only when $\mathbf{\Pi}^*z = z$. Thus,

$$\begin{aligned} \Delta V(z_k) &= -[z_k, (\mathbf{I} - \mathbf{\Pi})\mathbf{P}(\mathbf{I} - \mathbf{\Pi})^*z_k] \\ &\quad - [z_k, \mathbf{\Pi}\mathbf{Q}\mathbf{\Pi}^*z_k] < 0, \end{aligned}$$

because $\mathbf{Q} > 0$ and one of the two terms is negative. Then, $\mathbf{\Pi}\mathbf{A}$ generates an asymptotically stable dynamics. $\square$

B.5 Proof of Lemma 8

i) $\mathbf{S}\mathbf{\Pi} = 0$ follows trivially from the fact that $\mathbf{\Pi}x \in \text{Ker}(\mathbf{S})$. (ii) Since $\mathbf{S}$ and $\mathbf{S}_i$ are self-adjoint and semi-positive definite, $\forall x \in \mathcal{H}^n$

$$[\mathbf{S}\mathbf{\Pi}x, \mathbf{\Pi}x]_{\mathcal{H}^n} = \sum_{i=1}^q [\mathbf{S}_i\mathbf{\Pi}x, \mathbf{\Pi}x]_{\mathcal{H}^n} = 0.$$

Since the terms in the sum are non-negative, they need to be 0 and $\mathbf{\Pi}x \in \text{Ker}(\mathbf{S}_i)$, that is, $\mathbf{S}_i\mathbf{\Pi}x = 0$ for all x and consequently $\mathbf{S}_i\mathbf{\Pi} = 0$. Properties (iii) and (iv) follow from the fact that, $\forall x \in \mathcal{H}^n$, $\mathbf{\Pi}^*x \in \text{Ker}(\mathbf{S}\mathbf{P})$, that is,

$\mathbf{S}\mathbf{P}\mathbf{\Pi}^*x = 0$, equivalent to $\mathbf{S}\mathbf{P}\mathbf{\Pi}^* = 0$, $\mathbf{\Pi}\mathbf{P}\mathbf{S}\mathbf{P} = 0$. As in the previous case,

$$[\mathbf{S}\mathbf{P}\mathbf{\Pi}^*x, \mathbf{P}\mathbf{\Pi}^*x]_{\mathcal{H}^n} = \sum_{i=1}^q [\mathbf{S}_i\mathbf{P}\mathbf{\Pi}^*x, \mathbf{P}\mathbf{\Pi}^*x]_{\mathcal{H}^n} = 0,$$

which implies, being $\mathbf{S}_i = \mathbf{S}_i^* \geq 0$, that $\mathbf{S}_i\mathbf{P}\mathbf{\Pi}^*x = 0$ for all x , in turn implying $\mathbf{S}_i\mathbf{P}\mathbf{\Pi}^* = 0$ and $\mathbf{\Pi}\mathbf{P}\mathbf{S}_i = 0$. $\square$

B.6 Proof of Lemma 9

Since $\mathcal{L} = \mathcal{L}^\top$ there is a matrix $U = (u_1, \dots, u_N)$ and a diagonal $\Lambda = \text{diag}_{k=1}^N(\lambda_k)$ such that $\mathcal{L}u_k = \lambda_k u_k$ and $\mathcal{L}U = U\Lambda$, $U^{-1} = U^\top$. Let $\lambda_1 = 0$ be the null eigenvalue of $\mathcal{L}$, which is simple because the graph is connected. The corresponding normalized eigenvector is $u_1 = \frac{1}{\sqrt{N}}\mathbf{1}_N$. Let $z = U^{-1}x = \text{col}_{i=1}^N(x_i)$, $z_i = u_i^\top x$. We obtain

$$\begin{aligned} \Psi x &= \Psi U z = U z - U \Lambda U^{-1} \text{col}_{i=1}^N(\kappa \mathbf{P}\mathbf{S}x_i) \\ &= U(z - \Lambda \text{col}_{i=1}^N(\kappa \mathbf{P}\mathbf{S}u_i^\top x_i)) \\ &= U \text{col}_{i=1}^N(z_i - \kappa \lambda_i \mathbf{P}\mathbf{S}z_i) \\ &= u_1 z_1 + \sum_{i=2}^N u_i (\mathbf{I} - \kappa \lambda_i \mathbf{P}\mathbf{S}) z_i \\ &= \left(u_1 u_1^\top + \sum_{i=2}^N u_i (\mathbf{I} - \kappa \lambda_i \mathbf{P}\mathbf{S}) u_i^\top \right) x. \end{aligned}$$

Due to the orthogonality between u_i and u_j for $i \neq j$,

$$\Psi^\gamma x = \left(u_1 u_1^\top + \sum_{i=2}^N u_i (\mathbf{I} - \kappa \lambda_i \mathbf{P}\mathbf{S})^\gamma u_i^\top \right) x.$$

Bu using the limit in Lemma 6,

$$\begin{aligned} \lim_{\gamma \rightarrow \infty} \Psi^\gamma x &= \left(u_1 u_1^\top + \sum_{i=2}^N u_i \mathbf{\Pi} u_i^\top \right) x \\ &= u_1 u_1^\top x + \left(\sum_{i=2}^N u_i u_i^\top \right) \text{col}_{i=1}^N(\mathbf{\Pi}x_i), \end{aligned}$$

and since $\sum_{i=2}^N u_i u_i^\top = I_N - u_1 u_1^\top$, the last expression is exactly (44). Due to the orthogonality between $u_1 u_1^\top$ and $I_N - u_1 u_1^\top$ is readily verified that $[\mathcal{P}_1 x, \mathcal{P}_2 y] = 0$ for all $x, y \in \mathcal{H}^n$. The proof of (48) is immediate by noticing that, with $M := I_N - u_1 u_1^\top$, $Mx = \text{col}_i(x_i - \bar{x})$ and thus, if $y = \text{col}_i(y_i)$,

$$\begin{aligned} [\mathcal{P}_2 y, x] &= [M \text{col}_i(\mathbf{\Pi}y_i), x] = [\text{col}_i(\mathbf{\Pi}y_i), Mx] \\ &= \sum_{i=1}^N [\mathbf{\Pi}y_i, (Mx)_i] = \sum_{i=1}^N [y_i, \mathbf{\Pi}^*(Mx)_i] \\ &= [y, \text{col}_i(\mathbf{\Pi}^*(Mx)_i)] \\ &= [y, \text{col}_i(\mathbf{\Pi}^*)(x_i - \bar{x})] = [y, \mathcal{P}_2^* x]. \end{aligned}$$

□

References

- Akyildiz, I., Su, W., Sankarasubramaniam, Y., Cayirci, E., 2002. Wireless sensor networks: a survey. *Computer networks* 38 (4), 393–422.
- Balakrishnan, A., 1974. Stochastic optimization theory in Hilbert spaces. *Applied Mathematics and Optimization* 1, 97–120.
- Balakrishnan, A. V., 2012. *Applied Functional Analysis*. Vol. 3. Springer Science & Business Media.
- Battilotti, S., Borri, A., Cacace, F., d’Angelo, M., 2024. Optimal discrete-time distributed Kalman filter with reduced communication. Submitted for publication.
- Battilotti, S., Cacace, F., d’Angelo, M., 2023a. Consensus analysis of random sub-graphs for distributed filtering with link failures. *IEEE Trans. on Automatic Control*.
- Battilotti, S., Cacace, F., d’Angelo, M., 2023b. Distributed optimal control of discrete-time linear systems over networks. *IEEE Trans. on Control of Network Systems*.
- Battilotti, S., Cacace, F., d’Angelo, M., 2021. A stability with optimality analysis of consensus-based distributed filters for discrete-time linear systems. *Automatica* 129, 109589.
- Battilotti, S., Cacace, F., d’Angelo, M., Germani, A., 2020. Asymptotically optimal consensus-based distributed filtering of continuous-time linear systems. *Automatica* 122, 109189.
- Battistelli, G., Chisci, L., 2014. Kullback–Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability. *Automatica* 50 (3), 707–718.
- Battistelli, G., Chisci, L., Forti, N., Pelosi, G., Selleri, S., 2016. Distributed finite-element Kalman filter for field estimation. *IEEE Transactions on Automatic Control* 62 (7), 3309–3322.
- Bensoussan, A., 2011. *Stochastic control by functional analysis methods*. Elsevier.
- Cai, M., Yuan, Y., Luo, B., Xu, X., Dubljevic, S., 2024. Adaptive boundary observer network design for the consensus on the estimation of a class of parabolic partial differential equation systems. *International Journal of Robust and Nonlinear Control* 34 (2), 927–948.
- Curtain, R., 1975a. A survey of infinite-dimensional filtering. *Siam Review* 17 (3), 395–411.
- Curtain, R. F., 1975b. Infinite-dimensional filtering. *SIAM Journal on Control* 13 (1), 89–104.
- Curtain, R. F., Zwart, H., 2012. *An introduction to infinite-dimensional linear systems theory*. Vol. 21. Springer Science & Business Media.
- Demetriou, M. A., 2010. Design of consensus and adaptive consensus filters for distributed parameter systems. *Automatica* 46 (2), 300–311.
- Demetriou, M. A., 2022. Modal consensus observers for distributed parameter systems. *International Journal of Robust and Nonlinear Control* 32 (1), 472–497.
- Falb, P. L., 1967. Infinite-dimensional filtering: The Kalman-Bucy filter in Hilbert space. *Information and Control* 11, 102–137.
- Fisher, M., Nocedal, J., Trémolet, Y., Wright, S. J., 2009. Data assimilation in weather forecasting: a case study in pde-constrained optimization. *Optimization and Engineering* 10 (3), 409–426.
- Fuhrmann, P. A., 1976. Exact controllability and observability and realization theory in Hilbert space. *Journal of Mathematical Analysis and Applications* 53 (2), 377–392.
- Gan, D., Xie, S., Liu, Z., Lü, J., 2024. Stability of ffs-based diffusion adaptive filter under cooperative excitation condition. *IEEE Transactions on Automatic Control* 69 (11), 7479–7492.
- Ge, X., Han, Q.-L., Wang, Z., 2017. A dynamic event-triggered transmission scheme for distributed set-membership estimation over wireless sensor networks. *IEEE Trans. on Cybernetics* 49 (1), 171–183.
- Guo, L., Wang, J., Zhao, Y., Zhang, J.-F., 2024. Consensus-based distributed nonlinear filtering with kernel mean embedding. *IEEE Transactions on Aerospace and Electronic Systems*.
- Hager, W. W., Horowitz, L. L., 1976. Convergence and stability properties of the discrete Riccati operator equation and the associated optimal control and filtering problems. *SIAM Journal on Control and Optimization* 14 (2), 295–312.
- He, S., Shin, H.-S., Xu, S., Tsourdos, A., 2020. Distributed estimation over a low-cost sensor network: A review of state-of-the-art. *Information Fusion* 54, 21–43.
- Kalman, R. E., 1960. A new approach to linear filtering and prediction problems. *Trans. of the ASME - Journal of Basic Engineering* 82 (1), 35–45.
- Kalman, R. E., Bucy, R. S., 1961. New results in linear filtering and prediction theory. *Trans. of the ASME - Journal of Basic Engineering* 83 (1), 95–108.
- Kamal, A., Farrell, J., Roy-Chowdhury, A., 2013. Information weighted consensus filters and their application in distributed camera networks. *IEEE Trans. on Automatic Control* 58 (12), 3112–3125.
- Kamgarpour, M., Tomlin, C., 2008. Convergence properties of a decentralized Kalman filter. In: *Proc. of the 47th IEEE Conf. on Decision and Control*. IEEE, pp. 3205–3210.
- Kempe, D., Dobra, A., Gehrke, J., 2003. Gossip-based computation of aggregate information. In: *Proc. of the 44th IEEE Symposium on Foundations of Computer Science*, 2003. IEEE, pp. 482–491.
- Khan, U. A., Moura, J. M., 2008. Distributing the Kalman filter for large-scale systems. *IEEE Transactions on Signal Processing* 56 (10), 4919–4935.
- Li, W., Wei, G., Ho, D., Ding, D., 2018. A weightedly uniform detectability for sensor networks. *IEEE Trans. on Neural Networks and Learning Systems* 29 (11), 5790–5796.

- Matei, I., Martins, N., Baras, J. S., 2008. Almost sure convergence to consensus in markovian random graphs. In: 2008 47th IEEE Conference on Decision and Control. IEEE, pp. 3535–3540.
- Moura, S., Bendtsen, J., Ruiz, V., 2013. Observer design for boundary coupled PDEs: Application to thermostatically controlled loads in smart grids. In: 52nd IEEE conference on decision and control. IEEE, pp. 6286–6291.
- Olfati-Saber, R., 2005. Distributed Kalman filter with embedded consensus filters. In: 44th IEEE Conf. on Decision and Control (CDC). IEEE, pp. 8179–8184.
- Olfati-Saber, R., 2009. Kalman-consensus filter: Optimality, stability, and performance. In: 48th IEEE Conf. on Decision and Control (CDC). IEEE, pp. 7036–7042.
- Ortner, M., Nehorai, A., Jeremic, A., 2007. Biochemical transport modeling and Bayesian source estimation in realistic environments. *IEEE Transactions on Signal Processing* 55 (6), 2520–2532.
- Sallberg, S. A., Maybeck, P. S., Oxley, M. E., 2010a. Infinite-dimensional sampled-data Kalman filter. In: 49th IEEE Conf. on Decision and Control (CDC). IEEE, pp. 7363–7368.
- Sallberg, S. A., Maybeck, P. S., Oxley, M. E., 2010b. Infinite-dimensional sampled-data Kalman filtering and the stochastic heat equation. In: 49th IEEE Conference on Decision and Control (CDC). IEEE, pp. 5062–5067.
- Silva, B., Fisher, R., Kumar, A., Hancke, G., 2015. Experimental link quality characterization of wireless sensor networks for underground monitoring. *IEEE Trans. on Industrial Informatics* 11 (5), 1099–1110.
- Talebi, S., Werner, S., 2019. Distributed Kalman filtering and control through embedded average consensus information fusion. *IEEE Trans. on Automatic Control* 64 (10), 4396–4403.
- Talebi, S., Werner, S., Gupta, V., Huang, Y.-F., 2021. On stability and convergence of distributed filters. *IEEE Signal Processing Letters* 28, 494–498.
- Wan, C., Gao, Y., Li, X., Song, E., 2018. Distributed filtering over networks using greedy gossip. In: 2018 21st International Conference on Information Fusion. IEEE, pp. 1968–1975.
- Wang, H.-D., Wu, H.-N., 2015. Design of finite dimensional robust h_∞ distributed consensus filters for dissipative pde systems with sensor networks. *International Journal of Robust and Nonlinear Control* 25 (10), 1454–1471.
- Wang, J.-W., Liu, Y.-Q., Hu, Y.-Y., Sun, C.-Y., 2017. A spatial domain decomposition approach to distributed h_∞ observer design of a linear unstable parabolic distributed parameter system with spatially discrete sensors. *International Journal of Control* 90 (12), 2772–2785.
- Wang, S., Ren, W., 2017. On the convergence conditions of distributed dynamic state estimation using sensor networks: A unified framework. *IEEE Trans. on Control Systems Technology* 26 (4), 1300–1316.
- Wei, G., Li, W., Ding, D., Liu, Y., 2018. Stability analysis of covariance intersection-based Kalman consensus filtering for time-varying systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 50 (11), 4611–4622.
- Wu, Z., Fu, M., Xu, Y., Lu, R., 2018. A distributed Kalman filtering algorithm with fast finite-time convergence for sensor networks. *Automatica* 95, 63–72.
- Xia, F., Qian, X., 2022. Design of infinite-dimensional distributed consensus filter for mobile sensor networks with packet losses. *The Journal of Engineering* 2022 (6), 608–617.
- Yang, W., Zhang, Y., Chen, G., Yang, C., Shi, L., 2019. Distributed filtering under false data injection attacks. *Automatica* 102, 34–44.