Asymptotically Optimal Consensus-based Distributed Filtering of Continuous-Time Linear Systems

Article *in* Automatica · December 2020

DOI: 10.1016/j.automatica.2020.109189

4 authors, including:

S. Battilotti
Sapienza University of Rome
119 PUBLICATIONS **808** CITATIONS
SEE PROFILE

Filippo Cacace
Università Campus Bio-Medico di Roma
157 PUBLICATIONS **1,293** CITATIONS
SEE PROFILE

Alfredo Germani
Università degli Studi dell'Aquila
254 PUBLICATIONS **3,356** CITATIONS
SEE PROFILE

Asymptotically Optimal Consensus-based Distributed Filtering of Continuous-Time Linear Systems

S. Battilotti [a] F. Cacace [b] M. d'Angelo [a] A. Germani [c]

[a] *DIAG, Università "Sapienza" di Roma, Roma, Italy, (e-mail: s.battilotti@diag.uniroma1.it, mdangelo@diag.uniroma1.it).*

[b] *Università Campus Bio-Medico di Roma, Roma, Italy (e-mail: f.cacace@unicampus.it).*

[c] *DISIM, Università degli Studi dell'Aquila, L'Aquila, Italy (e-mail: alfredo.germani@univaq.it).*

Abstract

We describe a consensus-based distributed filtering algorithm for linear systems with a parametrized gain and show that when the parameter becomes large the error covariance at each node becomes arbitrarily close to the error covariance of the optimal centralized Kalman filter. The result concerns distributed estimation over a connected un-directed or directed graph and for static configurations it only requires to exchange the estimates among adjacent nodes. A comparison with related approaches confirms the theoretical results and shows that the method can be applied to a wide range of distributed estimation problems.

Key words: Continuous-time filters; Kalman filtering; Distributed filtering; Consensus filters.

1 Introduction

Consensus-based algorithms for distributed filtering of the state of a dynamical target system have attracted considerable research in the last years. Distributed filtering involves state estimation using a set of local filters that communicate with all other nodes through a communication network that constraints the information exchange to the neighbors in the network without requiring a centralized node. This framework encompasses a wide range of cases, such as multi-sensor fusion where a set of identical sensors aims at reaching a common estimate of the target through some distributed information fusion algorithm, or wireless sensor networks where the individual nodes may either be *communication nodes* that only have processing and computation capabilities, but no direct access to the target system, or *sensor nodes* with additional sensing capabilities. The popularity of distributed estimation techniques is due to their scalability for large networks and high fault tolerance. Surveys of distributed state estimation approaches and comparison with centralized and decentralized methods can be found in [4, 14, 18]. In addition, the motivation for reaching consensus and synchronization in the distributed estimates arises in many fields, especially in connection with control problems, where are particularly relevant in the context of cooperative multi-agent systems, see [11, 15, 20, 21] for comprehensive treatments.

The most relevant issues for distributed estimation techniques include: (i) *accuracy*, i.e., reducing the estimation error at each node; (ii) *consensus*, i.e. convergence of estimates across nodes; (iii) *communication*, i.e. reducing the amount of communication burden among nodes; (iv) *observability*, i.e. the capability of dealing with sensors with limited or null system measurements [8]. It is generally assumed the existence of trade-offs among these features. For example, strong consensus may be obtained at the expenses of estimation accuracy at some nodes [2], and increased accuracy might require more intensive communication across nodes. It is moreover tacitly assumed that the accuracy of distributed schemes is always worse than accuracy of a centralized optimal algorithm. These limitations are present in all the approaches reported in the literature. In [14] the system is supposed to be jointly detectable by each node and its neighbors. Recent approaches [7, 10, 19] have relaxed this assumption but a trade-off between internal stability of the filter and accuracy of the estimates still exists.

Our work extends the preliminary version in [1] and shows that these trade-offs can be overcome.

(1) We exhibit a distributed algorithm that requires only collective observability, attains consensus and tends to optimal accuracy with least communication burden.
(2) The proposed algorithm extends the consensus dis-

tributed DKF of [14] to the case of networks with nodes that have limited or null measurement capabilities and provides a more accurate estimate.

(3) Our work extends the results obtained in [16] that uses a similar approach for systems with null or bounded measurement noise.

The Approximate Distributed Kalman Filter (ADKF) is introduced in Section 3 after formally describing the framework in Section 2 and the centralized optimal approach in Section 2.1. The extension to directed graphs and arbitrary gains is considered in Section 4. In Section 5 and Section 6 we compare, from the theoretical and numerical perspective, the proposed algorithm with the centralized solution and other recent approaches, DKF [14], MKCF [17, 7], DKBF [19], and IKCF of [7].

Notation. $\mathbb{R}$ denotes real numbers. For a square matrix A, $\mathrm{tr}(A)$ is the trace, $\sigma(A)$ is the spectrum and $\mu(A)$ the spectral abscissa. If $\mu(A) < 0$ then A is said to be Hurwitz stable. $\mathbb{E}\{\cdot\}$ denotes expectation. $\otimes$ is the Kronecker product. The operators $\mathrm{row}_i()$, $\mathrm{col}_i()$, $\mathrm{diag}_i()$ denote respectively the horizontal, vertical and diagonal compositions of matrices and vectors indexed by i. Let $\mathcal{S}(n) \in \mathbb{R}^{n\times n}$ be the set of symmetric matrices of size n. $\mathcal{P}(n)$ (resp., $\mathcal{P}_+(n)$) $\subset \mathcal{S}(n)$ denotes the set of positive semi-definite (definite) matrices in $\mathcal{S}(n)$. I_n is the identity matrix of size n and $U_n = 1_n 1_n^\top$, $1_n = \mathrm{col}_{i=1}^n(1)$ is the square matrix of size n with all entries 1.

2 Problem formulation and preliminaries

We use a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ to describe the information exchange between the N nodes. $\mathcal{V} = \{1, 2 \ldots, N\}$ is the set of agents and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges. An edge (i, j) of $\mathcal{G}$ represents the exchange information between nodes i and j. The graph is undirected and connected, and the set of neighbors of node i is denoted by $\mathcal{N}^{(i)} := \{j \in \mathcal{V} : (j, i) \in \mathcal{E},\ j \neq i\}$. The adjacency matrix $\mathcal{A}$ of a graph $\mathcal{G}$ is an $N \times N$ matrix, whose (i, j)-th entry is 1 if $(i, j) \in \mathcal{E}$. The degree matrix $\mathcal{G}$ is $\mathcal{D} = \mathrm{diag}_i(|\mathcal{N}^{(i)}|)$. The Laplacian of $\mathcal{G}$ is $\mathcal{L} = \mathcal{D} - \mathcal{A}$. $\mathcal{L}$ is symmetric iff the graph is undirected. Moreover, $0 = \lambda_1(\mathcal{L}) < \lambda_2(\mathcal{L}) \leqslant \cdots \leqslant \lambda_N(\mathcal{L})$, where $\lambda_i(\mathcal{L}) \in \sigma(\mathcal{L})$, iff the graph is connected. An eigenvector associated to $\lambda_1(\mathcal{L})$ is 1_N.

Consider the process

$$\dot{\mathbf{x}}_t = A\mathbf{x}_t + \mathbf{f}_t, \tag{1}$$

$$\mathbf{y}_t^{(i)} = C_i\mathbf{x}_t + \mathbf{g}_t^{(i)}, \quad i = 1, \ldots, N, \tag{2}$$

where $\mathbf{x}_t \in \mathbb{R}^n$, $\mathbf{y}_t^{(i)} \in \mathbb{R}^{q_i}$, $q_i \geqslant 0$, and $\mathbf{f}_t$ and $\mathbf{g}_t^{(i)}$, $i = 1, \ldots, N$, are zero-mean white noises, mutually independent with covariance respectively $Q \in \mathcal{P}_+(n)$, $R_i \in \mathcal{P}_+(q_i)$ $i = 1, \ldots, N$ [6]. The matrices Q and $R = \mathrm{diag}_i(R_i)$ are nonsingular. Also $\mathbf{x}_0$ is random with mean $\bar{\mathbf{x}}_0 := \mathbb{E}\{\mathbf{x}_0\}$ and covariance $\Sigma_{\mathbf{x}_0} := \mathbb{E}\{(\mathbf{x}_0 - \bar{\mathbf{x}}_0)(\mathbf{x}_0 - \bar{\mathbf{x}}_0)^\top\}$. If $C = \mathrm{col}_i(C_i)$ we assume that (C, A) is observable. Also, we use the notation $\mathbf{y}_t = \mathrm{col}_i(\mathbf{y}_t^{(i)})$. Each $\mathbf{y}_t^{(i)}$ represents the data available at node i, $i = 1, \ldots, N$, in the network. We will design a distributed state estimator for the system (1) with the given topology of the network $\mathcal{G}$. The distributed estimators will consist of N local estimators, one for each node, which exchange their local estimate with the neighbors.

2.1 Centralized Kalman-Bucy filter (CKBF)

The equations of the centralized Kalman-Bucy filter (CKBF) for (1) are $\hat{\mathbf{x}}_0 = \bar{\mathbf{x}}_0$, $\mathbf{P}_0 = \Sigma_{\mathbf{x}_0}$,

$$\dot{\hat{\mathbf{x}}}_\mathbf{t} = A\hat{\mathbf{x}}_t + \mathbf{K}_t(\mathbf{y}_t - C\hat{\mathbf{x}}_t), \tag{3}$$

$$\dot{\mathbf{P}}_t = A\mathbf{P}_t + \mathbf{P}_t A^\top + Q - \mathbf{P}_t C^\top R^{-1} C\mathbf{P}_t, \tag{4}$$

with $\mathbf{K}_t = \mathbf{P}_t C^\top R^{-1}$. The matrix $\mathbf{P}_t$ represents the covariance of the estimation error $\mathbb{E}\{(\mathbf{x}_t - \hat{\mathbf{x}}_t)(\mathbf{x}_t - \hat{\mathbf{x}}_t)^\top\}$. The CKBF is optimal and in the Gaussian case it computes the conditional expectation $\mathbb{E}\{\mathbf{x}_t \mid \mathbf{y}_s,\ s \leqslant t\}$, The covariance $\mathbf{P}_t \in \mathcal{P}_+(n)$ is bounded for all $t \geqslant 0$ and $\mathbf{P}_t \to P_\infty$ as $t \to +\infty$ with $P_\infty \in \mathcal{P}_+(n)$ the unique solution of

$$0 = AP_\infty + P_\infty A^\top + Q - P_\infty C^\top R^{-1} C P_\infty. \tag{5}$$

If $A_C := A - K_\infty C$, $K_\infty = P_\infty C^\top R^{-1}$, from (4), we obtain the asymptotically optimal CBKF

$$\dot{\hat{\mathbf{x}}}_{ss,t} = A\hat{\mathbf{x}}_{ss,t} + K_\infty(\mathbf{y}_t - C\hat{\mathbf{x}}_{ss,t}). \tag{6}$$

3 Approximate distributed Kalman filter (ADKF)

3.1 ADKF algorithm

The ADKF consists of one filter for each sensor node of the network. The equations at the i-th sensor node are

$$\begin{aligned}\dot{\hat{\mathbf{x}}}_t^{(i)} =& A\hat{\mathbf{x}}_t^{(i)} + K_i(\mathbf{y}_t^{(i)} - C_i\hat{\mathbf{x}}_t^{(i)}) \\ &+ \gamma P_\infty \sum_{j\in\mathcal{N}^{(i)}} (\hat{\mathbf{x}}_t^{(j)} - \hat{\mathbf{x}}_t^{(i)}),\end{aligned} \tag{7}$$

with $K_i = NP_\infty {C_i}^\top {R_i}^{-1}$, P_∞ solution of (5) and $\gamma > 0$ a filter parameter chosen so that the matrix $A_D(\gamma)$ defined below is Hurwitz stable.

$$A_i := A - K_i C_i,\ i = 1, \ldots, N, \tag{8}$$

$$A_D(\gamma) := \mathrm{diag}_i(A_i) - \gamma(\mathcal{L} \otimes P_\infty) \in \mathbb{R}^{nN\times nN}. \tag{9}$$

Theorem 1 *There exists $\gamma_0 > 0$ such that for all $\gamma > \gamma_0$ $A_D(\gamma)$ is Hurwitz stable.*

Proof. We prove that there exists $X_\infty \in \mathcal{P}_+(nN)$ and $\gamma_0 > 0$ such that for all $\gamma > \gamma_0$

$$X_\infty A_D(\gamma) + A_D^\top(\gamma) X_\infty < 0. \tag{10}$$

If $X_\infty := I_N \otimes P_\infty^{-1}$ and $W_i := P_\infty^{-1} A_i + A_i^\top P_\infty^{-1}$, we have

$$X_\infty A_D(\gamma) + A_D^\top(\gamma) X_\infty = \mathrm{diag}_i(W_i) - 2\gamma(\mathcal{L} \otimes I_n).$$

Since $v \in \mathbb{R}^{nN} \backslash \{0\} : v^\top (\mathcal{L} \otimes I_n) v = 0 \Rightarrow v = 1_N \otimes \bar{v}$, $\bar{v} \in \mathbb{R}^n \backslash \{0\}$, it follows that

$$\begin{aligned} v \in \mathbb{R}^{nN} \backslash \{0\} : \quad & v^\top (\mathcal{L} \otimes I_n) v = 0 \\ \Rightarrow v^\top \mathrm{diag}_i(W_i) v &= N \bar{v}^\top (P_\infty^{-1} A_C + A_C^\top P_\infty^{-1}) \bar{v} < 0. \end{aligned}$$

By Finsler's lemma there exists $\gamma_0 > 0$ such that, for all $\gamma > \gamma_0$, $\mathrm{diag}_i(W_i) - 2\gamma(\mathcal{L} \otimes I_n) < 0$, which proves (10). □

3.2 Properties of the ADKF

Define the local estimation error $\mathbf{e}_t^{(i)} := \mathbf{x}_t - \widehat{\mathbf{x}}_t^{(i)}$, the total estimation error $\mathbf{e}_t := \mathrm{col}_i(\mathbf{e}_t^{(i)})$, the total measurement noise vector $\mathbf{g}_t := \mathrm{col}_i(\mathbf{g}_t^{(i)}) \in \mathbb{R}^{\sum_{i=1}^n q_i}$, the noise vector $\mathbf{h}_t := \mathrm{col}_i(\mathbf{f}_t - K_i \mathbf{g}_t^{(i)})$ with covariance $\Psi_h := \mathbb{E}\{\mathbf{h}_t \mathbf{h}_t^\top\}$, and the estimation error covariance matrix $\mathbf{X}_t := \mathbb{E}\{\mathbf{e}_t \mathbf{e}_t^\top\}$. Clearly $\mathbf{X}_t$ depends on γ, but we omit this dependence for notational simplicity.

$$\dot{\mathbf{e}}_t = A_D(\gamma)\mathbf{e}_t + \mathbf{h}_t \tag{11}$$

$$\Psi_h = U_N \otimes Q + \mathrm{diag}_i(K_i) R \, \mathrm{diag}_i(K_i^\top). \tag{12}$$

Theorem 2 *For all $\gamma > \gamma_0$ the estimation error covariance matrix $\mathbf{X}_t$ is uniformly bounded in time.*

Proof. The result follows from the fact that $A_D(\gamma)$ is Hurwitz for $\gamma > \gamma_0$ and Ψ_h is constant. □

Our purpose is to show the key result that $\mathbf{X}_\infty \to U_N \otimes P_\infty$ when $\gamma \to \infty$ (P_∞ is the asymptotic error covariance of the CKBF). Let $\mathbf{X}_\infty^C := U_N \otimes P_\infty$ be the the asymptotic error covariance of N identical CBKFs. Since $(\mathcal{L} \otimes P_\infty)\mathbf{X}_\infty^C = (\mathcal{L} \otimes P_\infty)(U_N \otimes P_\infty) = 0$,

$$\begin{aligned} 0 = & \left(\mathrm{diag}_{i=1}^N(A_C) - \gamma(\mathcal{L} \otimes P_\infty) \right) \mathbf{X}_\infty^C \\ & \mathbf{X}_\infty^C \left(\mathrm{diag}_{i=1}^N(A_C) - \gamma(\mathcal{L} \otimes P_\infty) \right)^\top + U_N \otimes Q \\ & + \mathrm{diag}_{i=1}^N(K_\infty)(U_N \otimes R) \mathrm{diag}_{i=1}^N(K_\infty^T). \end{aligned} \tag{13}$$

Let $G_i := C_i^\top R_i^{-1} C_i$, $G_d := \mathrm{diag}_i(G_i)$, and $G := C^\top R^{-1} C$. Notice that

$$G = \sum_{i=1}^N G_i = \sum_{i=1}^N C_i^\top R_i^{-1} C_i. \tag{14}$$

By introducing the covariance mismatch $\mathbf{E}_t := \mathbf{X}_t - \mathbf{X}_\infty^C$ we obtain after some manipulations

$$\dot{\mathbf{E}}_t = A_D(\gamma)\mathbf{E}_t + \mathbf{E}_t A_D^\top(\gamma) + \Sigma, \tag{15}$$

$$\begin{aligned} \Sigma := & N^2 (I_N \otimes P_\infty) G_d (I_N \otimes P_\infty) \\ & + U_N \otimes (P_\infty G P_\infty) - N (I_N \otimes P_\infty) G_d (U_N \otimes P_\infty) \\ & - N (U_N \otimes P_\infty) G_d (I_N \otimes P_\infty). \end{aligned} \tag{16}$$

Our main result can thus be stated as follows.

Theorem 3 *As $\gamma \to +\infty$, the covariance matrix of the estimation error of the ADKF (7) tends to the covariance matrix of the estimation error of the CKBF (4) when $t \to +\infty$. In other words, we have*

$$\lim_{\gamma \to +\infty} \lim_{t \to +\infty} \mathbf{X}_t = \mathbf{X}_\infty^C := U_N \otimes P_\infty. \tag{17}$$

In order to prove the above statement, we notice that there always exists a transformation T such that

$$T = \begin{pmatrix} \frac{1}{\sqrt{N}} 1_N^\top \\ t_2 \\ \vdots \\ t_N \end{pmatrix}, \quad T \mathcal{L} T^\top = \mathrm{diag}\{0, \lambda_2, \ldots, \lambda_N\}$$

where $\lambda_2, \ldots, \lambda_N > 0$ are the positive eigenvalues of $\mathcal{L}$ (see Section 2) and $t_i 1_N = \sum_{j=1}^N t_{i,j} = 0, t_i t_j^\top = \delta_{i,j}$ with $i, j = 2, \ldots, N$, i.e. T is orthonormal. Define $S := T \otimes I_n$ and let $\widetilde{\mathbf{E}}_t := S \mathbf{E}_t S^\top$. We have after some manipulations and taking into account (14)

$$\dot{\widetilde{\mathbf{E}}}_t = \widetilde{A}_D(\gamma)\widetilde{\mathbf{E}}_t + \widetilde{\mathbf{E}}_t \widetilde{A}_D^\top(\gamma) + N^2 \widetilde{\Sigma} \tag{18}$$

where for $i, j = 2, \ldots, N$,

$$\widetilde{A}_D(\gamma) = \begin{pmatrix} A_C & \Pi_{1,2} & \cdots & \Pi_{1,N} \\ \Pi_{2,1} & \Pi_{2,2} - \gamma\lambda_2 P_\infty & \cdots & \Pi_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \Pi_{N,1} & \Pi_{N,2} & \cdots & \Pi_{N,N} - \gamma\lambda_N P_\infty \end{pmatrix} \tag{19}$$

$$
\begin{aligned}
\Pi_{1,j} =&\Pi_{j,1} := -\sqrt{N}\sum_{l=1}^{N} t_{j,l} P_\infty G_l, \quad j = 2,\dots,N,\\
\Pi_{i,j} =&\Pi_{j,i} := \delta_{i,j} A - N\sum_{l=1}^{N} t_{i,l} t_{j,l} P_\infty G_l,\\
\widetilde{\Sigma} =& \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & \widetilde{\Sigma}_{2,2} & \cdots & \widetilde{\Sigma}_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \widetilde{\Sigma}_{N,2} & \cdots & \widetilde{\Sigma}_{N,N} \end{pmatrix}\\
\widetilde{\Sigma}_{i,j} =&\widetilde{\Sigma}_{j,i} := \sum_{l=1}^{N} t_{i,l} t_{j,l} P_\infty G_l P_\infty.
\end{aligned}
\tag{20}
$$

The matrix $\widetilde{A}_D(\gamma)$ is Hurwitz $\forall\gamma > \gamma_0$ since $A_D(\gamma)$ is Hurwitz $\forall\gamma > \gamma_0$ by Theorem 1. Moreover, since $\forall\gamma > \gamma_0 : \sigma(\widetilde{A}_D(\gamma)) \cap \sigma(-\widetilde{A}_D^\top(\gamma)) = \varnothing$, for any $\gamma > \gamma_0$ there exists a unique symmetric solution $\widetilde{\mathbf{E}}_\infty(\gamma)$ to

$$
0 = \widetilde{A}_D(\gamma)\widetilde{\mathbf{E}}_\infty(\gamma) + \widetilde{\mathbf{E}}_\infty(\gamma)\widetilde{A}_D^\top(\gamma) + N^2\widetilde{\Sigma} \tag{21}
$$

The proof of Theorem 3 follows directly from the next two lemmas.

Lemma 1 *For each $\gamma > \gamma_0$ the matrix $\widetilde{\mathbf{E}}_t$ satisfies*

$$
\lim_{t\to+\infty} \widetilde{\mathbf{E}}_t = \widetilde{\mathbf{E}}_\infty(\gamma). \tag{22}
$$

Proof. Let $\boldsymbol{\Delta}_t := \widetilde{\mathbf{E}}_t - \widetilde{\mathbf{E}}_\infty(\gamma)$. We have $\dot{\boldsymbol{\Delta}}_t = \widetilde{A}_D(\gamma)\boldsymbol{\Delta}_t + \boldsymbol{\Delta}_t A_D^\top(\gamma)$. Since $\widetilde{A}(\gamma)$ is Hurwitz for all $\gamma \geqslant \gamma_0$ we have $\boldsymbol{\Delta}_t = e^{\widetilde{A}_D(\gamma)t}\boldsymbol{\Delta}_0 e^{\widetilde{A}_D^\top(\gamma)t} \to 0$ as $t \to +\infty$. □

Lemma 2 *The matrix $\widetilde{\mathbf{E}}_t$ satisfies*

$$
\lim_{\gamma\to+\infty} \widetilde{\mathbf{E}}_\infty(\gamma) = 0. \tag{23}
$$

Proof. The solution $\widetilde{\mathbf{E}}_\infty(\gamma)$ of (21) is unique and it can be parametrized in γ as follows. Let

$$
W_1 := \mathrm{row}_{i=2}^N \Pi_{1i} \quad W_2 := \mathrm{col}_{i=2}^N \Pi_{1i} \tag{24}
$$

$$
W_0 := \begin{pmatrix} \Pi_{2,2} & \cdots & \Pi_{2,N} \\ \vdots & \ddots & \vdots \\ \Pi_{N,2} & \cdots & \Pi_{N,N} \end{pmatrix}, \quad \Lambda := \begin{pmatrix} \widetilde{\Sigma}_{2,2} & \cdots & \widetilde{\Sigma}_{2,N} \\ \vdots & \ddots & \vdots \\ \widetilde{\Sigma}_{N,2} & \cdots & \widetilde{\Sigma}_{N,N} \end{pmatrix}. \tag{25}
$$

With this definitions the equation (21) reads out as

$$
\begin{aligned}
\begin{pmatrix} 0 & 0 \\ 0 & \Lambda \end{pmatrix} =& \begin{pmatrix} A_C & W_1 \\ W_2 & W_0 - \gamma D \otimes P_\infty \end{pmatrix} \widetilde{\mathbf{E}}_\infty(\gamma)\\
&+ \widetilde{\mathbf{E}}_\infty(\gamma) \begin{pmatrix} A_C & W_1 \\ W_2 & W_0 - \gamma D \otimes P_\infty \end{pmatrix}^\top,
\end{aligned}
\tag{26}
$$

where $D = \mathrm{diag}_{i=2}^N(\lambda_i)$. The solution $\widetilde{\mathbf{E}}_\infty(\gamma)$ is analytic in $\gamma > 0$ and can be written (using a Taylor expansion) as

$$
\widetilde{\mathbf{E}}_\infty(\gamma) = \frac{1}{\gamma} \begin{pmatrix} Y_{1,1} + O\left(\frac{1}{\gamma^2}\right) & \frac{1}{\gamma}Y_{2,1} + O\left(\frac{1}{\gamma^2}\right) \\ \frac{1}{\gamma}Y_{2,1}^\top + O\left(\frac{1}{\gamma^2}\right) & Y_{3,1} + \frac{1}{\gamma}Y_{3,2} + O\left(\frac{1}{\gamma^2}\right) \end{pmatrix} \tag{27}
$$

where $Y_{3,1}$ is the unique (since $\sigma(D \otimes P_\infty) \cap \sigma(-D \otimes P_\infty) = \varnothing$) solution of

$$
Y_{3,1}(D \otimes P_\infty) + (D \otimes P_\infty)Y_{3,1} = N^2\Lambda,
$$

$Y_{3,2}$ is the unique (since $\sigma(D \otimes P_\infty) \cap \sigma(-D \otimes P_\infty) = \varnothing$) solution of

$$
(D \otimes P_\infty)Y_{3,2} + Y_{3,2}(D \otimes P_\infty) = W_0 Y_{3,1} + Y_{3,1}W_0^\top,
$$

$Y_{2,1}$ is defined as $Y_{2,1} := W_1 Y_{3,1}(D \otimes P_\infty)^{-1}$, $Y_{1,1}$ is the unique (since $\sigma(A_C) \cap \sigma(-A_C^\top) = \varnothing$) solution of

$$
A_C Y_{1,1} + Y_{1,1}A_C^\top = -(W_1 Y_{2,1}^\top + Y_{2,1}W_1^\top).
$$

From (27) it follows that $\lim_{\gamma\to+\infty} \widetilde{E}_\infty(\gamma) = 0$. □

Clearly, from the above lemmas we conclude that

$$
\lim_{\gamma\to+\infty}\lim_{t\to+\infty} \widetilde{\mathbf{E}}_t = \lim_{\gamma\to+\infty}\lim_{t\to+\infty} \mathbf{E}_t = 0 \tag{28}
$$

which proves Theorem 3. An important consequence is that the error covariance of each filter at the sensor node tends (as $\gamma \to +\infty$ and $t \to +\infty$) to the optimal steady state P_∞ of the centralized filter.

Corollary 1 *For each $j = 1, \dots, N$, it holds*

$$
\lim_{\gamma\to+\infty}\lim_{t\to+\infty} [\mathbf{X}_t]_{j,j} = P_\infty. \tag{29}
$$

Remark 1 *In any discrete-time implementation γ cannot be chosen arbitrarily large, due to numerical issues. A larger γ requires a smaller integration step that constraints the communication lag among nodes. Consequently, any implmentation of the ADKF will suffer a certain performance degradation with respect to the CKBF, in accordance with what happens in the discrete time framework.*

Algorithm Broadcast Push-Sum

1: In all nodes set $s_{0,i} = C_i^\top R_i^{-1} C_i$ and $w_{0,i} = 0$, except for $w_{0,1} = 1$.
2: At time 0 each nodes sends $(s_{0,i}, w_{0,i})$ to itself.
3: At time t each node executes:
   1. Let $\{s_r, w_r\}$ be the pairs sent to i in round $t-1$.
   2. Let $s_{t,i} = \sum_r s_r$, $w_{t,i} = \sum_r w_r$.
   3. Send to all neighbors and to i (yourself):

$$\left(\frac{1}{\left|\mathcal{N}^{(i)}\right| + 1} s_{t,i},\ \frac{1}{\left|\mathcal{N}^{(i)}\right| + 1} w_{t,i} \right)$$

   4. $s_{t,i}/w_{t_i}$ is the estimate of G at step t (if $w_{t,i} = 0$ the estimate is not specified or 0).

Fig. 1. A modified version of the *Push-Sum* algorithm of [9] that makes possible the distributed computation G.

3.3 Distributed computation of P_∞

The ADKF is extremely simple to implement and the information exchange among nodes is reduced to a minimum. However each node i needs to compute (or to know) P_∞, thus the aim of this section is to show how the this computation can be implemented in a completely distributed manner. The matrix P_∞ can be computed by solving (5), a matrix equation with size n that does not depend on the graph structure and that even nodes with limited computational power can easily solve. Clearly, with many sensor nodes the size of C and R can be large, but $C^\top R^{-1} C$ is a $n \times n$ matrix. The solution of (5) requires $G = C^\top R^{-1} C$. When measurement noises are independent G is expressed as in (14), that is, the sum of the matrices $C_i^\top R_i^{-1} C_i$ over the graph. A distributed computation of G can be achieved by resorting to distributed algorithms to compute aggregate functions over graphs [9]. In Fig. 1 we report an algorithm derived from the *Protocol Push-Sum* of [9] to compute G in a distributed way. The speed of convergence of the local estimate to the true value of G can be analyzed in the light of the results of [9]. In the case of *static graphs*, the estimation of G can be executed off-line before the filtering phase. In presence of *time-varying graphs*, *sensor failure* or *insertion*, the *Push-Sum* algorithm can be kept running during the execution of the filter in order to adapt the value of G and therefore of P_∞. Finally, the value of N can be computed by the same distributed algorithm when it is not known at the nodes.

3.4 Distributed computation of γ

Theorem 2 states that the estimation error covariance is bounded for $\gamma > \gamma_0$, thus the lower bound γ_0 for γ needs to be known at the nodes. Consider the inequality

$$(1_N \otimes P_\infty^{-1}) \widetilde{A}_D(\gamma) + \widetilde{A}_D^\top(\gamma)(1_N \otimes P_\infty^{-1}) < 0 \qquad (30)$$

where $\widetilde{A}_D(\gamma)$ is as in (19). The left-hand part of (30) is

$$\begin{pmatrix} W_{1,1} & -W_{1,2} \\ -W_{2,1} & A_P - W_{2,2} \end{pmatrix} - 2\gamma \begin{pmatrix} 0 & 0 \\ 0^\top & \Lambda \otimes I_n \end{pmatrix} \qquad (31)$$

where, by using Riccati equation (5),

$$\begin{aligned} W_{1,1} :=& P_\infty^{-1} A_C + A_C^\top P_\infty^{-1} = -P_\infty^{-1} Q P_\infty^{-1} - G < 0 \\ W_{1,2} :=& W_{2,1}^\top := 2\sqrt{N} \left(\textstyle\sum_{l=1}^N t_{2,l} G_l \ \cdots \ \sum_{l=1}^N t_{N,l} G_l \right), \\ A_P :=& I_{N-1} \otimes (P_\infty^{-1} A + A^\top P_\infty^{-1}) \\ W_{2,2} :=& 2N \begin{pmatrix} \sum_{l=1}^N t_{2,l}^2 G_l & \cdots & \sum_{l=1}^N t_{2,l} t_{N,l} G_l \\ \vdots & \ddots & \vdots \\ \sum_{l=1}^N t_{N,l} t_{2,l} G_l & \cdots & \sum_{l=1}^N t_{N,l}^2 G_l \end{pmatrix} \end{aligned} \qquad (32)$$

Therefore, a necessary and sufficient condition for (30) is

$$A_P - W_{2,2} - 2\gamma \Lambda \otimes I_n - W_{2,1} W_{1,1}^{-1} W_{1,2} < 0 \qquad (33)$$

Using the fact that $\sum_{j=1}^N t_{i,j}^2 = 1$ for all $i = 1, \ldots, N$, $W_{2,2} \gamma 0$ and $W_{2,1} W_{1,2} \leqslant 4N^2 (I_{N-1} \otimes G)$, we obtain

$$\gamma > \frac{\|P_\infty^{-1} A + A^\top P_\infty^{-1}\| + \bar{\gamma}_0}{2(1 - \cos(\pi/N))} > \frac{\|P_\infty^{-1} A + A^\top P_\infty^{-1}\| + \bar{\gamma}_0}{\lambda_2(\mathcal{L})} \qquad (34)$$

$$\bar{\gamma}_0 = 4N^2 \mu \left((P_\infty^{-1} Q P_\infty^{-1} + G)^{-1} \right) \mu(G), \qquad (35)$$

where we have used the inequality $\lambda_2(\mathcal{L}) \geqslant 2(1 - \cos(\pi/N))$, that holds for a connected undirected graph with no multiple edges [5]. The lower bound depends on A, Q, P_∞, G and N, that are available at each node in view of the results of Section 3.3.

4 Additional results

4.1 Directed graphs

In this section we outline a generalization of ADKF to weighted directed graphs. The (i, j) entry of $\mathcal{A}$ is the weight associated to the edge. We assume that the graph is *simple* and strongly connected, *i.e.* there exists a directed path between each pair of nodes. The *Laplacian* $\mathcal{L} \in \mathbb{R}^{N \times N}$ is defined as $\mathcal{L} := \mathcal{M} - \mathcal{A}$, $\mathcal{M} = \mathrm{diag}_i(\sum_{j=1}^N \mathcal{A}_{i,j})$. $\mathcal{L}$ has a zero eigenvalue with eigenvector $\mathbf{1}_N$ and if the graph is strongly connected $\mu(-\mathcal{L}) = 0$. The equations for the ADKF at the i-th sensor node are obtained from (7) by replacing N with $\mathrm{tr}(D)/d_i$,

$$\begin{aligned} \dot{\widehat{\mathbf{x}}}_t^{(i)} =& A \widehat{\mathbf{x}}_t^{(i)} + \frac{\mathrm{tr}(D)}{d_i} K_i (\mathbf{y}_t^{(i)} - C_i \widehat{\mathbf{x}}_t^{(i)}) \\ &+ \gamma P_\infty \sum_{j \in \mathcal{N}^{(i)}} (\widehat{\mathbf{x}}_t^{(j)} - \widehat{\mathbf{x}}_t^{(i)}), \end{aligned} \qquad (36)$$

with $K_i = P_\infty {C_i}^\top {R_i}^{-1}$ and $\gamma > 0$ a parameter. Define $A_i := A - (\mathrm{tr}(D)/d_i) K_i C_i$ and A_D as in (9).

Theorem 4 *There exists $\gamma_0 > 0$ such that for all $\gamma > \gamma_0$ $A_D(\gamma)$ is Hurwitz stable.*

Theorem 5 *As $\gamma \to +\infty$, the covariance matrix of the estimation error of the ADKF* (36) *tends to the covariance matrix of the estimation error of the CKBF* (4) *when $t \to +\infty$.*

4.2 Suboptimal distributed Kalman filter (SDKF)

By replacing the gains K_i in ADKF with non optimal gains L_i we obtain an algorithm, named SDKF, that achieves consensus but it is no longer optimal for $\gamma \to \infty$.

$$\dot{\widehat{\mathbf{x}}}_t^{(i)} = A\widehat{\mathbf{x}}_t^{(i)} + NL_i(\mathbf{y}_t^{(i)} - C_i\widehat{\mathbf{x}}_t^{(i)}) + \gamma S \sum_{j\in\mathcal{N}^{(i)}} (\widehat{\mathbf{x}}_t^{(j)} - \widehat{\mathbf{x}}_t^{(i)}), \tag{37}$$

with $L_i \in \mathbb{R}^{n\times q_i}$, $S \in \mathcal{S}(n)$ and $\gamma > 0$ design parameters. Define $A_i := A - NL_iC_i$ and $A_D(\gamma) := \mathrm{diag}_i(A_i) - \gamma(\mathcal{L}\otimes S)$.

Theorem 6 *For any $L = \mathrm{col}_i(L_i)$ such that $A_C := A - LC$ is Hurwitz stable and any $S \in \mathcal{P}_+(n)$ such that $A_C S + SA_C^\top + Q + LRL^T = 0$ there exists $\gamma_0 > 0$, such that for all $\gamma > \gamma_0$ $A_D(\gamma)$ is Hurwitz stable.*

Theorem 7 *For all $\gamma > \gamma_0$ the estimation error covariance matrix $\mathbf{X}_t$ is uniformly bounded in time and $\lim_{t\to+\infty} \mathbf{X}_t = \mathbf{X}_\infty^C(\gamma)$ where $\mathbf{X}_\infty^C(\gamma)$ is the unique positive definite solution of $A_D(\gamma)\mathbf{X}_\infty^C(\gamma) + \mathbf{X}_\infty^C(\gamma)A_D^\top(\gamma) + \Psi_h = 0$ that satisfies*

$$\lim_{\gamma\to+\infty} \mathbf{X}_\infty^C(\gamma) = \lim_{\gamma\to+\infty}\lim_{t\to+\infty} \mathbf{X}_t = U_N \otimes S_L. \tag{38}$$

5 Comparison with related approaches

DKF. The following popular continuous-time distributed filter was proposed in [13]. as an enhancement of [12].

$$\begin{aligned}\dot{\widehat{\mathbf{x}}}_t^{(i)} =& A\widehat{\mathbf{x}}_t^{(i)} + P_iC_i^\top R_i^{-1}\left(\mathbf{y}^{(i)} - C_i\widehat{\mathbf{x}}_t^{(i)}\right) \\ &+ \gamma P_i \sum_{j\in\mathcal{N}^{(i)}} \left(\widehat{\mathbf{x}}_t^{(i)} - \widehat{\mathbf{x}}_t^{(i)}\right) \qquad (39)\\ \dot{P}_i =& AP_i + P_iA^\top + Q - P_iC_i^\top R_i^{-1}C_iP_i \qquad (40)\end{aligned}$$

with γ a positive parameter. Notice the similarity with (7) Clearly, P_i is bounded when (C_i, A) is an observable pair. A slight extension is to apply the algorithm by replacing $\mathbf{y}^{(i)}$ with $\mathrm{col}_{j\in\mathcal{N}^{(i)}}(\mathbf{y}^{(j)})$ and C_i with $\mathrm{col}_{j\in\mathcal{N}^{(i)}}(C_j)$. In any case, the DKF can be applied only when local observability conditions hold, and in particular $C_i \neq 0$, that is, all the sensor nodes (or their immediate neighbors) have measurements. In this condition the estimation error variance is bounded.

MKCF. The Multi-agent Consensus Filter (MKCF) was originally proposed in [17] and extended in [7]. It assumes that only the nodes for which $g_i = 1$ have measurements and in thi case $C_i = I$,

$$y_i(t) = g_i z_i = g_i(x + v_i). \tag{41}$$

The MKCF algorithm needs to exchange both the estimates and the matrices P_i among adjacent nodes. There are no theoretical stability results for the MKCF. MKCF can applied to a wider sets of network with respect to DKF but it is limited by the assumption $C_i = I$.

DKBF. The Distributed Kalman-Bucy Filter [19] builds over the approach of [10], where it was shown that the optimal centralized accuracy can be attained if the consensus gains were chosen by taking into account the estimation error covariance of each node. Since this is not feasible in a distributed setting and resulted in a conservative choice of the consensus gain, [19] proposed

$$\begin{aligned}\dot{\widehat{\mathbf{x}}}_t^{(i)} =& A\widehat{\mathbf{x}}_t^{(i)} + P_iC_i^\top R_i^{-1}(z_i - C_i\widehat{\mathbf{x}}_t^{(i)}) \\ &+ P_i \sum_{j\in\mathcal{N}^{(i)}} P_j^{-1}\left(\widehat{\mathbf{x}}_t^{(j)} - \widehat{\mathbf{x}}_t^{(i)}\right) \qquad (42)\\ \dot{P}_i =& AP_i + P_iA^\top + Q - P_iC_i^\top R_i^{-1}C_iP_i \\ &- P_i \sum_{j\in\mathcal{N}^{(i)}} \left(P_j^{-1} - P_i^{-1}\right) P_i. \qquad (43)\end{aligned}$$

For this filter [19] proves stability in mean of the estimation error under an *ad-hoc* assumption that is however difficult to verify about the convergence of the estimation error covariance matrices of the nodes.

IKCF. The Information-Weighted Kalman consensus filter (IKCF) [7] extends the MKCF to $C_i \neq I$. It also explicitly models communication noises. In absence of communication noises the algorithm is

$$\begin{aligned}\dot{\widehat{\mathbf{x}}}_{\mathbf{t}}^{(\mathbf{i})} =& A\widehat{\mathbf{x}}_t^{(i)} + g_iP_iC_i^\top R_i^{-1}(z_i - C_i\widehat{\mathbf{x}}) \\ &+ P_i \sum_{j\in\mathcal{N}^{(i)}} P_j^{-1}\left(\widehat{\mathbf{x}}_t^{(j)} - \widehat{\mathbf{x}}_t^{(i)}\right) \qquad (44)\\ \dot{P}_i =& (A + |\mathcal{N}^{(i)}|I)P_i + P_i(A + |\mathcal{N}^{(i)}|I)^\top + Q \\ &- g_i^2 P_iC_i^\top R_i^{-1}C_iP_i - P_i \sum_{j\in\mathcal{N}^{(i)}} \left(P_j^{-1} - P_i^{-1}\right) P_i. \qquad (45)\end{aligned}$$

The proof of the boundedness of the estimation error provided in [7] contains some technical issues [3].

6 Simulation results

We consider several scenarios. Let $N = 5$ with the nodes connected in a chain, i.e. the edges are $(i, i+1)$, and system (1) with

$$A = \begin{pmatrix} -0.1 & 0 & 0 & 0 \\ 0.5 & -0.5 & 0 & 0 \\ 1.5 & 0 & -0.2 & 0 \\ -1 & 0 & 1 & 0 \end{pmatrix} \tag{46}$$

Scenario 1: complete information. The system state is available to all nodes, *i.e.* $C_i = I_4$.

Scenario 2: local observability. The output is available to all nodes, $C_i = [1,\ 1,\ 1,\ 1]$. (C_i, A) is observable.

Scenario 3: collective observability. The system is available to all nodes, $C_1 = [1,\ 0,\ 1,\ 0]$, $C_2 = [0,\ 1,\ 0,\ 0]$, $C_3 = [1,\ 0,\ 0,\ 0]$, $C_4 = [1,\ 0,\ 1,\ 1]$, $C_5 = [0,\ 1,\ 1,\ 0]$. (C_i, A) is not observable, but (C, A) is observable.

Scenario 4: sparse state availability. The system state is available only to nodes in $\{1,\ 5\}$, *i.e.* $C_1 = C_5 = I_4$, $C_2 = C_3 = C_4 = 0$.

Scenario 5: sparse measurement availability. The output is available only to $\{1,\ 5\}$, $C_2 = C_3 = C_4 = 0$,

$$C_1 = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad C_5 = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \tag{47}$$

The comparison includes 100 simulations with $t \in [0, 50]$ and integration step $\mathrm{d}t = 2 \cdot 10^{-3}$. $Q = q_f I_4$ with $q_f = 0.3$ and $R_i = r_f I_i$, where $r_f = 0.6$ and I_i has the same number of rows as C_i. For ADKF, SDKF and DKF $\gamma = 100$. DKF can be used only when there is local observability. However, at nodes with no measurements K_i is not needed and P_i may assume any value. We used $P_i = I$ in the simulations of Scenario 4. MKCF needs $C_i = I$, thus it can be used in scenarios 1 and 4. However, when R_i is scalar the terms are still congruent, thus we used MKCF also for scenarios 2, 3. The results are summarized in Table 1. Notice that when it can be applied, the DKF attains a better consensus than ADKF. This fact however may depend on the value of γ. To investigate this issue we have computed the variance of the estimation error $\mathrm{tr}([\mathbf{X}_\infty]_{i,i})$ at each node as a function of γ in Scenario 3 (notice that $A_D(\gamma)\mathbf{X}_\infty + \mathbf{X}_\infty A_D^\top(\gamma) = -\Psi_h$ holds true). The results are shown in Fig. 2 (left). The plot confirms Theorem 3: when γ increases the variance of the estimation error of all the nodes converge to the optimal value $\mathrm{tr}(P_\infty)$. Fig. 2 (right) shows that the consensus is a linear function of γ. Fig. 3 illustrates the distributed computation of P_∞ with the *Push-Sum* algorithm on a 5×5 grid of 25 nodes and the output matrices of scenario 3. C_1–C_4 are used at the grid corners, the other nodes have no measurement. With $\mathrm{d}t = 2 \cdot 10^{-3}$ the nodes converge at P_∞ in less than 0.2 s. At $t = 0.5$ s node 8 starts with C_5 and P_∞ is updated in less than 0.1 s.

7 Conclusions

Our results do not extend immediately to the discrete-time case (Remark 1). Thus, it is interesting to derive a discrete-time implementation of ADKF and to characterize the loss of accuracy and consensus. Further extensions include communication delays and disturbances.

References

[1] S. Battilotti, F. Cacace, M. d'Angelo, and A. Germani. Asymptotically optimal consensus-based distributed filtering of continuous-time linear system. *IFAC World Congress, Berlin*, 2020.

[2] G. Battistelli, L. Chisci, G. Mugnai, A. Farina, and A. Graziano. Consensus-based linear and nonlinear filtering. *IEEE Trans. on Automatic Control*, 60(5):1410–1415, 2014.

[3] F. Cacace. Comments on "distributed information-weighted Kalman consensus filter for sensor networks". *Automatica*, 109:108552, 2019.

[4] C.-Y. Chong. Forty years of distributed estimation: A review of noteworthy developments. In *2017 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, pages 1–10. IEEE, 2017.

[5] R.A. Horn and C.R. Johnson. *Matrix analysis*. Cambridge University Press, 2012.

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

[7] H. Ji, F.L. Lewis, Z. Hou, and D. Mikulski. Distributed information-weighted Kalman consensus filter for sensor networks. *Automatica*, 77:18–30, 2017.

[8] A.T. Kamal, J.A. Farrell, and A.K. Roy-Chowdhury. Information weighted consensus filters and their application in distributed camera networks. *IEEE Trans. on Automatic Control*, 58(12):3112–3125, 2013.

[9] D. Kempe, A. Dobra, and J. Gehrke. Gossip-based computation of aggregate information. In *Proc. of the 44th Annual IEEE Symp. Foundations of Computer Science, 2003*, pages 482–491. IEEE, 2003.

[10] J. Kim, H. Shim, and J. Wu. On distributed optimal Kalman-Bucy filtering by averaging dynamics

		$\mathrm{tr}(P_\infty)$	CKBF	ADKF $\gamma = 10^2$	SDKF $\gamma = 10^2$	$\mathrm{tr}(S_L)$	DKF [13],$\gamma = 10^2$	MKCF [17]	DKBF [19]	IKCF [7]
Scenario 1	mse	0.319	0.310	0.382	0.502	0.405	0.408	7.374	0.553	1.544
	st.dev.			0.031	0.044		0.004	0.411	0.047	0.155
Scenario 2	mse	0.797	0.742	0.781	1.971	1.912	0.820	2.539	0.898	$> 10^3$
	st.dev.			0.014	0.040		0.003	0.078	0.025	$> 10^3$
Scenario 3	mse	0.553	0.537	0.591	1.341	1.302	N.A.	17.207	0.863	2.818
	st.dev.			0.028	0.072		N.A.	3.632	0.107	0.703
Scenario 4	mse	0.532	0.522	0.595	0.909	0.764	0.727	$> 10^2$	0.959	1.406
	st.dev.			0.064	0.220		$< 10^{-3}$	> 10	0.134	0.481
Scenario 5	mse	0.582	0.564	0.652	2.805	2.813	N.A.	N.A.	1.020	2.519
	st.dev.			0.050	0.267		N.A.	N.A.	0.203	0.519

Table 1
Comparison of **mse** and **consensus**. Consensus is measured by the mse standard dev. across nodes (smaller values are better).

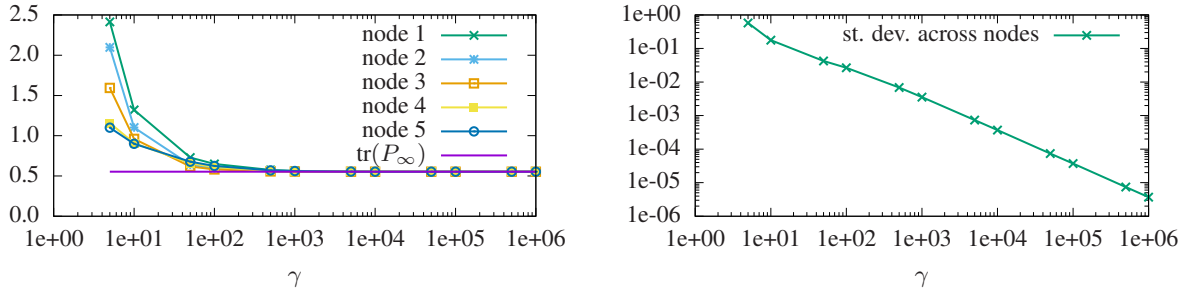

Fig. 2. ADKF: $\mathrm{tr}([\mathbf{X}_\infty]_{i,i})$ for the nodes of Scenario 3, $i \in \{1,2,3,4,5\}$, compared with the optimal variance $\mathrm{tr}(P_\infty)$ as a function of γ (left). Standard deviation across nodes of the estimation error variances shows that consensus increases with γ (right).

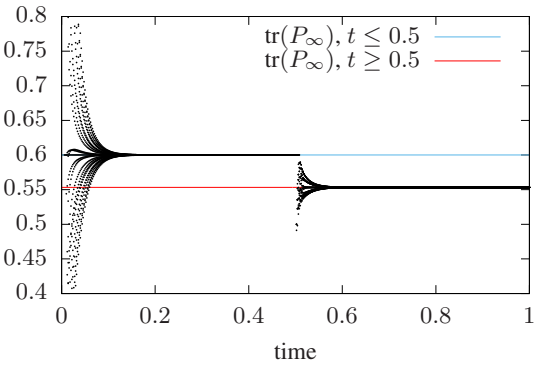

Fig. 3. Distributed computation of P_∞ for $N = 25$ with the *Push-Sum* algorithm. One node is activated at $t = 0.5$.

of heterogeneous agents. In *2016 IEEE 55th Conf. on Decision and Control*, pages 6309–6314. IEEE, 2016.

[11] F.L. Lewis, H. Zhang, K. Hengster-Movric, and A. Das. *Cooperative control of multi-agent systems: optimal and adaptive design approaches*. Springer Science & Business Media, 2013.

[12] R. Olfati-Saber. Distributed Kalman filter with embedded consensus filters. In *Proc. of the 44th IEEE Conf. on Decision and Control*, pages 8179–8184. IEEE, 2005.

[13] R. Olfati-Saber. Distributed Kalman filtering for sensor networks. In *Proc. of the 46th IEEE Conf. on Decision and Control*, pages 5492–5498. IEEE, 2007.

[14] R. Olfati-Saber, J.A. Fax, and R.M. Murray. Consensus and cooperation in networked multi-agent systems. *Proc. of the IEEE*, 95(1):215–233, 2007.

[15] Z. Qu. *Cooperative control of dynamical systems: applications to autonomous vehicles*. Springer Science & Business Media, 2009.

[16] W. Ren and U.M. Al-Saggaf. Distributed Kalman–Bucy filter with embedded dynamic averaging algorithm. *IEEE Systems Jour.*, 12(2):1722–1730, 2017.

[17] W. Ren, R.W. Beard, and D.B. Kingston. Multi-agent Kalman consensus with relative uncertainty. In *Proc. of the 2005, American Control Conf., 2005*, pages 1865–1870. IEEE, 2005.

[18] M. Taj and A. Cavallaro. Distributed and decentralized multicamera tracking. *IEEE Signal Processing Magazine*, 28(3):46–58, 2011.

[19] J. Wu, A. Elser, S. Zeng, and F. Allgöwer. Consensus-based distributed Kalman-Bucy filter for continuous-time systems. *IFAC-PapersOnLine*, 49(22):321–326, 2016.

[20] Y. Zhao, Y. Liu, G. Wen, W. Ren, and G. Chen. Designing distributed specified-time consensus protocols for linear multiagent systems over directed graphs. *IEEE Trans. on Automatic Control*, 64(7):2945–2952, 2018.

[21] Y. Zhao, Y. Liu, G. Wen, X. Yu, and G. Chen. Distributed average tracking for Lipschitz-type of nonlinear dynamical systems. *IEEE Trans. on Cybernetics*, 49(12):4140–4152, 2018.