Continuous-Time Distributed Observers with Discrete Communication

Florian Dörfler*, Fabio Pasqualetti, and Francesco Bullo

***Abstract*—This work presents a distributed algorithm for observer design for linear continuous-time systems. We assume the dynamical system to be partitioned into disjoint areas, and we let each area be equipped with a control center. Each control center knows local dynamics, collects local observations, performs local computation, and communicates with neighboring control centers at discrete times. For our continuous-discrete estimation algorithm we prove convergence, we characterize its convergence rate, and we show robustness against discretization and communication errors. Our technical approach is inspired by waveform relaxation methods and combines tools from estimation theory, decentralized control theory, and parallel computation. We illustrate the effectiveness of our algorithm with illustrative examples in sensor networks and electric power systems.**

I. Introduction

In response to recent technological advances, distributed estimation is receiving a tremendous scientific interest. In distributed estimation, the goal is for a team of coordinated units to estimate the state of a large-scale dynamical system via local measurements, local knowledge of the system structure and parameters, and distributed computation. Distributed estimation problems arise, for instance, in sensor networks, electric power grids, and industrial control systems.

In this paper we address the problem of distributed estimation by designing distributed observer for a set of coordinated control centers. Ever since the pioneering work by Aoki and Li [1] on observability with partial measurements, the problem of distributed observer design has received a tremendous amount of attention. A variety of solution strategies have been proposed under different assumptions on observability, information sets, and communication constraints. We refer the reader to [2]–[9] for a set of seminal references, and to [10]–[14] for a sample of more recent approaches. Early approaches in discrete [1]–[3] and continuous-time [4]–[9] settings primarily investigate the challenges of decentralized stabilization of the observer error dynamics, reduction of computational complexity, and information fusion in the presence of noise, possibly with the initial observer synthesis relying on global system knowledge and the ultimate data processing being centralized. Instead, recent approaches [10]–[14] focus on distributed observer design with a distributed observer synthesis based on local information, local computation and data processing, and communication constraints.

This material is based upon work supported by NSF grants IIS-0904501 and CPS-1135819.

Florian Dörfler, Fabio Pasqualetti, and Francesco Bullo are with the Center for Control, Dynamical Systems and Computation, University of California, Santa Barbara, {dorfler,fabiopas,bullo}@engineering.ucsb.edu

Distributed observers are usually designed for discrete-time systems, since communication and computation are performed with digital data, and rely upon distributed mechanisms to merge local computations, such as intermediate data fusion or averaging steps. The application of these observer to continuous-time systems, when possible [4]–[9], [14], requires real-time communication of continuous-time signals, such as measurements or state variables of local filters. An exception is presented in [15, Section 5], where continuous-time observers are coupled with discrete-time estimates. In this case, however, the observers design is centralized, and it relies on global system knowledge. To the best of our knowledge, there is no practically applicable solution to continuous-time distributed observer problem with discrete-time communication.

The motivation for continuous-time estimation over an a-priori discretization of the plant is not only the theoretical challenge to fill a gap in the estimation literature, but also the absence of a-priori errors due to sampling and discretization schemes, the applicability to plants with fast dynamics equipped with low bandwidth communication systems (the discretized model requires high sampling rate and high bandwidth communication to accurately capture fast dynamics), and the flexibility of adapting the time resolution (sampling rate) in an event-based fashion [16].

In this paper, we provide an asymptotically convergent and practically-applicable solution to the distributed continuous-time observer problem with discrete-time communication. We consider a large-scale, linear, and continuous-time system in a deterministic and noiseless setting. We assume the system to be partitioned into disjoint areas, and we let each area be equipped with a control center capable of estimating the state of its own (isolated) area and of communicating with neighboring control centers. In contrast to classic approaches in decentralized estimation [5]–[9] sharing a similar problem setup, our method does not rely on a centralized observer design and real-time communication of continuous-time signals. Rather, we propose a fully distributed estimation algorithm which combines local continuous-time estimation and communication with neighboring units at discrete time instants (the communicated data can either be a continuous-time signal or an approximate representation), and allows each control center to estimate the state of its own area with a pre-specified time-delay. Hence, our continuous-discrete algorithm is suited for applications that can tolerate a certain delay in state estimation, such as smoothing, off-line state estimation, attack and fault detection in a distributed setting, and output control for systems with slow dynamics.

Our continuous-discrete algorithm is inspired by our earlier investigation of waveform relaxation methods [17]–[19], and it combines decentralized control techniques [6], [20] with waveform relaxation methods developed for parallel numerical integration [21]–[23]. Compared to our earlier work and the waveform relaxation literature, our current approach is based on state-space analysis. Thanks to this novel approach, we can derive a necessary and sufficient convergence condition reminiscent of the small-gain and spectral radius criteria found in the literature on decentralized estimation [6], [20] and waveform relaxation [21]–[23]. We further derive convergence rate estimates as well as sufficient conditions for contractivity of the estimation error for a finite number of communication rounds and measurements collected in a finite-time horizon. Finally, we discuss the design of our estimation algorithm via local computation, and we quantify the effect of discretization and communication errors. We validate the performance and the applicability of our estimation algorithm with illustrative examples from sensor networks and power systems.

The remainder of this article is organized as follows: Section II presents our distributed estimation setup, and an intuitive approach based on decentralized control theory and parallel numerical integration. Section III presents our continuous-discrete estimation algorithm, an analysis of its convergence properties, and a comprehensive discussion of its implementation issues. Section IV validates our algorithm with two illustrative examples. Finally, Section V concludes the paper.

II. Problem Setup and Preliminary Concepts

Consider the linear continuous-time dynamical system

$$\begin{aligned} \dot{x}(t) &= Ax(t), \\ y(t) &= Cx(t), \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$ is the state, $y \in \mathbb{R}^p$ is the measurement, and $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{p\times n}$. Generally, a continuous-time estimation filter for system (1) takes the form

$$\dot{w}(t) = (A + LC)w(t) - Ly(t), \tag{2}$$

where $L \in \mathbb{R}^{n\times p}$ is an output injection matrix. If the pair (C, A) is observable, then L can be chosen so that the matrix $A + LC$ is Hurwitz and the error $w(t) - x(t)$ converges with time. The filter (2) is inherently centralized, and it typically cannot be implemented for large-scale systems because (i) the state $w(t)$ is of high dimension causing high computational effort, (ii) the sensors collecting the measurements $y(t)$ are spatially distributed, and (iii) real-time continuous communication of all measurements to a central processor is required. In this work we propose a distributed estimation strategy for large-scale continuous-time systems.

A. Problem setup

We assume the system (1) to be sparse. In particular, let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the directed graph associated with the matrix $A = [a_{ij}]$, that is, the nodal set is $\mathcal{V} = \{1, \dots, n\}$ and there exists a directed edge $(i, j) \in \mathcal{E} \subseteq \mathcal{V}\times\mathcal{V}$ if and only if $a_{ij} \neq 0$. Let $\mathcal{V}$ be partitioned into N disjoint sets $\mathcal{V} = \{\mathcal{V}_1, \dots, \mathcal{V}_N\}$, and let $\mathcal{G}_i = (\mathcal{V}_i, \mathcal{E}_i)$ be the ith subgraph of $\mathcal{G}$ with vertices $\mathcal{V}_i$ and edges $\mathcal{E}_i = \mathcal{E} \cap (\mathcal{V}_i \times \mathcal{V}_i)$. According to this partition, and possibly after permuting the states, the matrix A reads as

$$A = \begin{bmatrix} A_1 & \cdots & A_{1N} \\ \vdots & \vdots & \vdots \\ A_{N1} & \cdots & A_N \end{bmatrix} = A_D + A_C, \tag{3}$$

where $A_i \in \mathbb{R}^{n_i\times n_i}$, $A_{ij} \in \mathbb{R}^{n_i\times n_j}$, A_D is block-diagonal, and $A_C = A - A_D$. Notice that, if $A_D = \text{blkdiag}(A_1, \dots, A_N)$, then A_D represents the decoupled diagonal subsystems and A_C describes their interconnections. Additionally, if A is sparse, then several blocks in A_C are zero.

We make the following assumptions:

(A1) **local measurements:** C is block-diagonal, that is, $C = \text{blkdiag}(C_1, \dots, C_N)$, $C_i \in \mathbb{R}^{p_i\times n_i}$;

(A2) **observability:** each pair (A_i, C_i) is observable.

Given the above structure, system (1) results from the interconnection of N subsystems as

$$\dot{x}(t) = A_D x(t) + A_C x(t), \quad y(t) = Cx(t). \tag{4}$$

The dynamics of the ith subsystem then read as

$$\begin{aligned} \dot{x}_i(t) &= A_i x_i(t) + \sum\nolimits_{j\in\mathcal{N}_i^{\text{in}}} A_{ij} x_j(t), \\ y_i(t) &= C_i x_i(t), \quad i = 1, \dots, N, \end{aligned} \tag{5}$$

where $\mathcal{N}_i^{\text{in}} = \{j \in \{1, \dots, N\} \setminus \{i\} \;:\; A_{ij} \neq 0\}$ denotes the in-neighbors of subsystem i. We also define the set of out-neighbors as $\mathcal{N}_i^{\text{out}} = \{j \in \{1, \dots, N\} \setminus \{i\} \;:\; A_{ji} \neq 0\}$.

We assume the presence of a *control center* in each subsystem with the following capabilities:

(A3) **local knowledge:** the ith control center knows the local system and measurement matrices A_i and C_i as well as its neighboring matrices A_{ij} for $j \in \mathcal{N}_i^{\text{in}}$; and

(A4) **local communication:** the ith control center receives information from its in-neighboring control centers $j \in \mathcal{N}_i^{\text{in}}$, and it sends information to its out-neighboring control centers $j \in \mathcal{N}_i^{\text{out}}$ at discrete times.

B. Decentralized estimation and waveform relaxation

In this section we recall some preliminary results and definitions to motivate our approach in Section III. Consider the filter (2) with a block-diagonal output injection, that is,

$$\dot{w}(t) = (A_D + LC)w(t) + A_C w(t) - Ly(t), \tag{6}$$

with $L = \text{blkdiag}(L_1, \dots, L_N)$, and $L_i \in \mathbb{R}^{n_i\times p_i}$. The error $e(t) = x(t) - w(t)$ obeys the dynamics

$$\dot{e}(t) = (A_D + A_C + LC)e(t)\,. \tag{7}$$

Due to the observability assumption (A2), there exists L_i such that $A_i + L_i C_i$ is Hurwitz, or equivalently, $A_D + LC$ is Hurwitz. Let $\lambda_{\text{max}} \in \mathbb{R}$ be the largest real part of the eigenvalues of $A_D + LC$. We employ a small-gain approach to large-scale interconnected systems [6], [20] and rewrite the error dynamics (7) as the closed-loop interconnection of the two subsystems $\Gamma_D :\;\; \dot{e}(t) = (A_D + LC)e(t) + v(t)$ and $\Gamma_C : v(t) = A_C e(t)$. Since both Γ_C and Γ_D are causal and internally stable, the error dynamics (7) are stable if the loop transfer matrix

$$G(s) = (sI - A_D - LC)^{-1} A_C \tag{8}$$

satisfies the spectral radius condition [24, Theorem 4.11]

$$\rho\left((\mathrm{j}\,\omega I - A_D - LC)^{-1} A_C\right) < 1 \text{ for all } \omega \in \mathbb{R}. \tag{9}$$

If condition (9) holds, then $G(\mathrm{j}\omega)$ is *Schur stable* for all $\omega \in \mathbb{R}$.

Lemma 2.1: ***(Decentrally-stabilized state estimation)*** Consider system (4) and the filter (6). Let $A_D + LC$ be Hurwitz, and assume that $G(\mathrm{j}\,\omega)$ is Schur stable for all $\omega \in \mathbb{R}$. Then, for all $x(0)$ and $w(0)$, it holds $\lim_{t\to\infty}\big(w(t) - x(t)\big) = 0$.

Observe that an implementation of the decentralized filter (6) requires control centers to continuously exchange their local estimation $w_i(t)$. This continuous communication obstacle can be overcome by means of a parallel integration of the filter (6). Among the available methods developed for parallel computation, the classical *waveform relaxation* is compatible with assumptions (A1)-(A4) [21]–[23]. The Gauss-Jacobi waveform relaxation method applied to the filter (6) yields the *waveform relaxation iteration*

$$\dot{w}^{(k)}(t) = (A_D + LC) w^{(k)}(t) + A_C w^{(k-1)}(t) - Ly(t)\,, \tag{10}$$

where $k \in \mathbb{N}$ is a discrete iteration index, $t \in [0,T]$ is the continuous-time variable in the interval $[0,T]$, $w^{(k)} : [0,T] \to \mathbb{R}^n$ is a trajectory with initial condition $w^{(k)}(0) = w_0$ for each $k \in \mathbb{N}$. The waveform relaxation iteration (10) is initialized with a profile $w^{(0)} : [0,T] \to \mathbb{R}^n$. Notice that (10) is an iteration in k on the variable w. On the other hand, for fixed k, (10) is a continuous-time dynamical system in the variable $w^{(k)}$ and $A_C w^{(k-1)}$ is a known input, since the value of w at iteration $k-1$ is used. The iteration (10) is said to be *uniformly convergent* if $\lim_{k\to\infty} \max_{t\in[0,T]} \big\|w^{(k)}(t) - w(t)\big\|_\infty = 0$, where w is the solution of the non-iterative filter (6).

An elegant analysis of the waveform relaxation iteration (10) can be carried out in the Laplace domain [23], where the map from $w^{(k-1)}$ to $w^{(k)}$ is $(sI - A_D - LC)^{-1} A_C$. Similar to the usual Gauss-Jacobi iteration, convergence of the iteration (10) relies on the contractivity of the iteration operator in an appropriate exponentially-scaled function space. The following result and two different methods of proof can be found in [22, Theorems 2.2 and 2.3] and [23, Theorem 5.2].

Lemma 2.2: ***(Waveform relaxation)*** Consider the waveform relaxation iteration (10) with input $y : [0,T] \to \mathbb{R}^p$. Assume the existence of $\mu \in \mathbb{R}$ such that the exponentially-scaled measurement $t \to y(t)\exp(-\mu t)$ and all its derivatives exist and are bounded. Let $\sigma = \max\{\mu, \lambda_{\max}\}$. The iteration (10) is uniformly convergent if $G(\sigma + \mathrm{j}\,\omega)$ is Schur stable for all $\omega \in \mathbb{R}$.

In case of smooth and bounded measurements $y(t)$, with $t \in [0,T]$, and stable filter dynamics, we have that $\sigma = 0$, and the convergence condition for the waveform relaxation iteration (10) equals the Schur stability condition (9) for decentralized stabilization of the filter error dynamics. Hence, by means of the waveform relaxation iteration (10), the decentrally stabilized estimation filter (6) can be implemented in a distributed fashion. On the other hand, as a disadvantage of this analysis, the measurements $y(t)$ are required to be smooth and bounded, and the continuous signal $w^{(k-1)}(t)$ needs to be communicated at every iteration. Moreover, the proofs [22], [23] leading to Lemma 2.2 are based on an infinite horizon analysis for both time t and iteration k, the convergence rate for finite t and k are not characterized, and the robustness properties of the waveform relaxation iteration (10) are not investigated. In the next section, we present a more refined, less restrictive, and novel analysis to combine decentralized estimation with waveform relaxation, leading to a continuous-time estimation algorithm with discrete communication.

Algorithm 1: *Continuous-Discrete Estimation (ith center)*

Input : Matrices A_i, C_i, A_{ij} with $j \in \mathcal{N}_i^{\text{in}}$, L_i (see Section II-A);
Param. : Horizon $T \in \mathbb{R}_{>0}$, Number of iterations k_{f}, State w_0;
Require : Error contractivity with k_{f} and T (see Theorem 3.1);

for $h = 0, 1, \dots$ **do**
  Collect measurements $y_{i,h}(t) = y_i(t + hT)$ with $t \in [0,T]$;
  Select initial profile $w_{j,h}^{(0)}(t)$ with $t \in [0,T]$, $j \in \mathcal{N}_i^{\text{in}}$;
  for $k = 1, \dots, k_{\text{f}}$ **do**
    Integrate the differential equation
    $\dot{w}_{i,h}^{(k)} = (A_i + L_i C_i) w_{i,h}^{(k)} + \sum_{j\in\mathcal{N}_i^{\text{in}}} A_{ij} w_{j,h}^{(k-1)} - L_i y_{i,h}$,
    with $t \in [0,T]$ and initial state $w_{i,h}^{(k)}(0) = w_{i,0}$;
    Transmit $w_{i,h}^{(k)}$ to control centers $\mathcal{N}_i^{\text{out}}$;
    Receive $w_{j,h}^{(k)}$ from control centers $\mathcal{N}_i^{\text{in}}$;
  Update initial state as $w_{i0} = w_{i,h}^{(k_{\text{f}})}(T)$;

III. Continuous-Distributed Estimation

In this section we present our distributed estimation algorithm. Our algorithm estimates the state of system (1) in a *moving time-window* fashion, where time is divided into intervals $[hT, (h{+}1)T]$, for some uniform horizon $T \in \mathbb{R}_{>0}$ and $h \in \mathbb{N}$. Our estimation algorithm is formally presented in Algorithm 1 and it consists of the following three steps for every stage $h \in \mathbb{N}$. First, control centers collect local measurements in the time interval $[hT, (h+1)T]$. Second, control centers estimate their local state in the interval $[hT, (h+1)T]$ through the distributed and iterative filter (in vector form)

$$\dot{w}_h^{(k)}(t) = \big(A_D + LC\big) w_h^{(k)}(t) + A_C w_h^{(k-1)}(t) - Ly(t + hT), \tag{11}$$

where $t \in [0,T]$, $k \in \{1,\dots,k_{\text{f}}\}$ for some $k_{\text{f}} \in \mathbb{N}_{>0}$, and $L = \text{blkdiag}(L_1,\dots,L_N)$. Third and finally, control centers exchange their estimated local state and iterate over h. The filter (11) is initialized with $w_0^{(k)}(0) = w_0 = \begin{bmatrix} w_{1,0}^{\mathsf{T}} & \dots & w_{N,0}^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}}$ and $w_h^{(k)}(0) = w_{h-1}^{(k_{\text{f}})}(T)$ for $h \geq 1$ and $k = 1,\dots,k_{\text{f}}$. Additionally, $w_h^{(0)}(t)$ is the initial profile vector with components $w_{j,h}^{(0)}(t)$, where $w_{j,h}^{(0)}(t)$ is a guess of control center i for the state $x_j(t)$ in the interval $t \in [hT, (h+1)T]$ (we let different control centers use identical initial guesses about common neighboring states).

Algorithm 1 requires integration capabilities at each control center, local measurements, and synchronous communication between neighboring control centers at discrete time instants. Additionally, the state $w_{i,h}$ and the continuous-time measurements $y_{i,h}(t) = y_i(t + hT)$ need to be stored by the i-th control center. Finally, our method generates a state estimate with delay T, and control centers need to perform k_{f} iterations (integration and communication) within each time interval of length T. Consequently, our method is not suited for applications requiring real-time state estimates, such as control of critical systems with fast dynamics, and it can be employed whenever a delay in state estimation can be tolerated, for instance for offline attack detection [19].

In order to state the convergence notions for Algorithm 1, we define the *estimation error*

$$e_h^{(k)}(t) = x(t+hT) - w_h^{(k)}(t)$$

for $k \in \{1,\dots,k_{\mathrm{f}}\}$, $t \in [0,T]$, $h \in \mathbb{N}$. Three notions of convergence are considered: convergence of the error $e_h^{(k)}(t)$ in k and t within each stage h, contractivity of the terminal error $e_h^{(k_{\mathrm{f}})}(t)$ over multiple stages h, and asymptotic convergence of the error $e_h^{(k)}(t)$ in k and over the stages h.

*Definition 1 (***Convergence notions***):* For Algorithm 1, the estimation error $e_h^{(k)}(t)$ is

(i) *stage-wise convergent* if for all $h \in \mathbb{N}$

$$\lim_{T\to\infty,k_{\mathrm{f}}\to\infty} \big\|e_h^{(k_{\mathrm{f}})}(T)\big\|_\infty = 0\,;$$

(ii) ε*-contractive* if for all $h \in \mathbb{N}$ and for some $\varepsilon \in (0,1)$

$$\big\|e_{h+1}^{(k_{\mathrm{f}})}(T)\big\|_\infty \leq \varepsilon\big\|e_{h+1}^{(k_{\mathrm{f}})}(0)\big\|_\infty = \varepsilon\big\|e_h^{(k_{\mathrm{f}})}(T)\big\|_\infty\,;$$

(iii) *asymptotically convergent* if for some $\eta \in [0,1)$ and all $t \in [hT,(h+1)T]$

$$\lim_{h\to\infty}\Big(\|e_h^{(k_{\mathrm{f}})}(t)\|_\infty \,/\, \max_{\tau\in[0,t]}\|e_h^{(0)}(\tau)\|_\infty\Big) \leq \eta^{k_{\mathrm{f}}}\,.$$

In what follows we show that for sufficiently large values of k_{f} and T, similar conditions as in Subsection II-B guarantee stage-wise convergence, ε-contractivity, and asymptotic convergence. We also derive explicit bounds on k_{f} and T, and performance and robustness characterizations.

A. Convergence and performance of Algorithm 1

Compared to the classic functional analysis approaches to waveform relaxation methods, our analysis relies on a control-theoretic state-space approach. Notice that for each stage $h \in \mathbb{N}$, Algorithm 1 is fully characterized by the distributed filter (11) with state $w_h^{(k)}(t)$, and that, over multiple stages, the filter (11) is updated according to $w_h^{(k)}(0) = w_{h-1}^{(k_{\mathrm{f}})}(T)$. Hence, we start by analyzing the error dynamics resulting from the distributed filter (11). Define the augmented filter error $e_h(t) = \big[e_h^{(1)\mathsf{T}} \,\dots\, e_h^{(k_{\mathrm{f}})\mathsf{T}}\big]^{\mathsf{T}}$, and consider the error system

$$\dot{e}_h(t) = \underbrace{\begin{bmatrix} A_D+LC & 0 & \cdots & 0 \\ A_C & A_D+LC & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & A_C & A_D+LC \end{bmatrix}}_{A_e} e_h(t)$$

$$+ \begin{bmatrix} A_C \\ 0 \\ \vdots \\ 0 \end{bmatrix} e_h^{(0)}(t), \quad e_h(0) = \begin{bmatrix} z_h \\ \vdots \\ z_h \end{bmatrix}, \tag{12}$$

where $e_h^{(0)}(t) = x(t+hT) - w_h^{(0)}(t)$ is the initial profile error and $z_h = x(hT) - w_h(0)$ the initial condition error at stage h.

Theorem 3.1: ***(Convergence of distributed filter*** (11)***)*** Consider system (1) and the distributed filter (11) with $t \in [0,T]$, $k \in \{1,\dots,k_{\mathrm{f}}\}$, and $h \in \mathbb{N}$. Assume that each exponentially-scaled initial profile error $t \to \exp(-\mu t)\big(x(t+hT) - w_h^{(0)}(t)\big)$ is integrable for some $\mu \in \mathbb{R}$ and for all $h \in \mathbb{N}$. The following statements are equivalent:

(i) For every $z_h \in \mathbb{R}^n$ and $e_h^{(0)} : [0,\infty) \to \mathbb{R}^n$, the estimation error is stage-wise convergent;

(ii) The matrix $A_D + LC$ is Hurwitz with spectrum to the left of $\lambda_{\max} < 0$, and the transfer matrix $G(\sigma + \mathrm{j}\,\omega)$ is Schur stable for $\sigma = \max\{\mu, \lambda_{\max}\}$ and for all $\omega \in \mathbb{R}$.

If the equivalent statements (i) and (ii) hold, then for any $\varepsilon \in (0,1)$ and $h \in \mathbb{N}$, there exist sufficiently large $k_{\mathrm{f}} \in \mathbb{N}$ and $T \in \mathbb{R}_{>0}$ such that the estimation error is ε-contractive.

Proof: Consider the error system (12). By superposition of the homogeneous and the particular solution, we obtain

$$e_h^{(k_{\mathrm{f}})}(t) = \mathcal{L}^{-1}\left((G(s))^{k_{\mathrm{f}}}\right) \star e_h^{(0)}(t) + E_{k_{\mathrm{f}}}\exp(A_e t)e_h(0)\,, \tag{13}$$

where $\mathcal{L}^{-1}$ denotes the inverse Laplace transform, $\star$ denotes the convolution operator, $G(s)$ is defined in equation (8), and $E_k = \begin{bmatrix} 0_{n\times(k-1)n} & I_n & 0_{n\times(k_f-k)n}\end{bmatrix}$ with I_n being the n-dimensional identity matrix and $0_{n\times p}$ being the $(n \times p)$-dimensional matrix of zero entries.

Since the initial profile error $e_h^{(0)}(t)$ and the initial condition error z_h are arbitrary and independent, the final error $e_h^{(k_{\mathrm{f}})}(T)$ converges as $T \to \infty$ and $k_{\mathrm{f}} \to \infty$ if and only if the error system (12) with system matrix A_e is stable and the input-output map $\mathcal{L}^{-1}\big((G(s))^{k_{\mathrm{f}}}\big)$ from $e_h^{(0)}(t)$ to $e_h^{(k_{\mathrm{f}})}(t)$ vanishes. Equivalently, the error system (12) is asymptotically stable if and only if $A_i + L_iC_i$ is Hurwitz (since A_e is block-triangular), and the input-output map $\mathcal{L}^{-1}\big((G(s))^{k_{\mathrm{f}}}\big)$ vanishes if and only if $\rho(G(s)) < 1$ for all s in the domain of the Laplace transform.

Since each subsystem $G(s)$ and the overall error system (12) are asymptotically stable with spectrum in the left-half plane bounded by $\lambda_{\max} < 0$, and since the exponentially scaled input $\exp(-\mu t)e_h^{(0)}(t)$ is integrable, the domain of the Laplace transform is given by $\{s \in \mathbb{C} : \mathrm{Real}(s) \geq \sigma\}$, see also [23, Section 3 and Appendix] and [22, Section 2]. Hence, the input-output map $\mathcal{L}^{-1}\big((G(s))^{k_{\mathrm{f}}}\big)$ vanishes if and only if

$$\sup_{\{s\in\mathbb{C}\,:\,\mathrm{Real}(s)\geq\sigma\}} \rho(G(s)) = \sup_{\{s\in\mathbb{C}\,:\,\mathrm{Real}(s)=\sigma\}} \rho(G(s)) < 1\,,$$

where the equality follows since the spectral radius satisfies the maximum principle [22], [23]. The latter condition is equivalent to $G(\sigma + \mathrm{j}\,\omega)$ being Schur stable for all $\omega \in \mathbb{R}$. This proves the equivalence of statements (i) and (ii).

Due to stage-wise convergence of the error $\|e_h^{(k_{\mathrm{f}})}(T)\|_\infty$, it follows that for every $h \in \mathbb{N}$ and for any $\varepsilon \in (0,1)$ there are k_{f} and T sufficiently large such that the error is ε-contractive. ■

According to Theorem 3.1 the estimation error is stage-wise convergent, and for sufficiently large values of k_{f} and T it is also ε-contractive. In what follows, we provide sufficient and explicit bounds quantifying how large $k_\mathrm{f} \in \mathbb{N}_{>0}$ and $T \in \mathbb{R}_{>0}$ have to be.

For a dynamical system with impulse response $G(t) = \mathcal{L}^{-1}(G(s))$, the truncated $\mathcal{L}_1$-norm is defined as

$$\|G\|_{\mathcal{L}_1[0,t]} = \int_0^t \|G(\tau)\|_\infty \, d\tau \,,$$

where $\|A\|_\infty$ is the induced ∞-norm of the matrix A. It can be shown that $\|G\|_{\mathcal{L}_1[0,t]}$ is the induced $\mathcal{L}_\infty$-gain [24] of the linear dynamical system $y(t) = G(t) \star u(t)$, that is,

$$\max_{\tau\in[0,t]} \|y(\tau)\|_\infty \leq \|G\|_{\mathcal{L}_1[0,t]} \cdot \max_{\tau\in[0,t]} \|u(\tau)\|_\infty \,, \tag{14}$$

and the inequality (14) is tight over all nontrivial $u : [0,t] \to \mathbb{R}$, $u \in \mathcal{L}_\infty$.

Theorem 3.2: ***(Performance of distributed filter)*** Consider system (1), the distributed filter (11), and the filter error dynamics (12). The estimation error satisfies

$$\begin{aligned} \big\|e_h^{(k)}(t)\big\|_\infty \leq & \sum\nolimits_{j=0}^{k-1} \gamma \cdot \exp(\lambda_{\max} t) t^j \, \|z_h\|_\infty \\ & + \|G\|^k_{\mathcal{L}_1[0,t]} \max_{\tau\in[0,t]} \|e_h^{(0)}(\tau)\|_\infty \,, \end{aligned} \tag{15}$$

where $\gamma > 0$ is a constant independent of k and t. Moreover, if the estimation error is ε-contractive and $\|G\|_{\mathcal{L}_1[0,T]} < 1$, then the estimation error is asymptotically convergent.

Proof: Consider (12) and (13), and introduce the shorthand $f_h(t)$ for the free response:

$$f_h(t) = E_k \exp(A_e t) e_h(0) = \sum\nolimits_{j=0}^{k-1} \exp((A_D + LC)t) \frac{t^j}{j!} A_C^j z_h \,.$$

By the triangle inequality, the error (13) is bounded as

$$\big\|e_h^{(k)}(t)\big\|_\infty \leq \|f_h(t)\|_\infty + \big\|\mathcal{L}^{-1}(G(s)^k) \star e_h^{(0)}(t)\big\|_\infty \,.$$

For each $t \in [0,T]$, the worst-case free response is bounded as

$$\begin{aligned} \|f_h(t)\|_\infty & \leq \sum\nolimits_{j=0}^{k-1} \underbrace{\|\exp((A_D + LC)t)\|_\infty}_{\leq c_1 e^{\lambda_{\max} t}} \underbrace{\left\|\frac{t^j}{j!} A_C^j\right\|_\infty}_{\leq c_2 t^j} \|z_h\|_\infty \\ & \leq \sum\nolimits_{j=0}^{k-1} c_1 c_2 \cdot \exp(\lambda_{\max} t) t^j \, \|z_h\|_\infty \,, \end{aligned}$$

where $c_2 = \max_{j\in\mathbb{N}} \|A_C^j/j!\|_\infty > 0$ is a finite constant.[1]

Since the error system (1) is a cascade of k subsystems, each with transfer function $G(s)$, a stage-by-stage $\mathcal{L}_\infty$-norm bounding of the forced response yields

$$\begin{aligned} & \max_{\tau\in[0,t]} \left\|\mathcal{L}^{-1}\left(G(s)^k\right) \star e_h^{(0)}(\tau)\right\|_\infty \\ & \qquad \leq \|G\|^k_{\mathcal{L}_1[0,t]} \max_{\tau\in[0,t]} \left\|e_h^{(0)}(\tau)\right\|_\infty \,. \end{aligned}$$

The bounds on the free and the forced response give the convergence rate estimate (15). Finally, notice that z_h, and hence $f_h(t)$, vanishes with h due to error contractivity. Then, if $\|G\|_{\mathcal{L}_1[0,T]} < 1$, asymptotic convergence follows then directly from (15). ■

Theorem 3.2 shows that the filter error consists of two terms. The first contribution is due to a mismatch in the system and filter initial states, and it vanishes with time within each stage. The second contribution arises because the state of neighboring subsystems is a priori unknown, and it vanishes with the number of iterations. Hence, Algorithm 1 is asymptotically convergent for sufficiently large T and k_f if $\|G\|_{\mathcal{L}_1[0,T]} < 1$, and (15) characterizes its convergence rate.

To state the following result in a compact way, we introduce $\alpha = \|z_0\|_\infty = \|x(0) - w_0^{(0)}(0)\|_\infty$ for the initial error in the first stage, and we assume the initial profile errors to be *uniformly β-bounded*, that is, there is $\beta > 0$ such that $\max_{t\in[0,T]} \|x(t + hT) - w_h^{(0)}(t)\|_\infty \leq \beta$ for all $h \in \mathbb{N}$.

Theorem 3.3: ***(Constants for error contractivity)*** For the system (1) and the distributed filter (11), assume that $A_D + LC$ is Hurwitz, $\|G\|_{\mathcal{L}_1[0,\infty)} < 1$, and the initial profile errors are uniformly β-bounded. For every $\varepsilon \in (0,1)$, $h \in \mathbb{N}$, and $\nu \in [0,1]$, the estimation error is *ε-contractive* if

$$k_\mathrm{f} \geq \frac{\log(\beta) - \log(\nu\alpha\varepsilon)}{\left|\log(\|G\|_{\mathcal{L}_1[0,\infty)})\right|} \,, \text{ and} \tag{16a}$$

$$\frac{|\lambda_{\max}| T - \mathrm{step}(T-1)(k_\mathrm{f} - 1)\log(T)}{\log(\gamma k_\mathrm{f}) - \log((1-\nu)\varepsilon)} \geq 1 \,, \tag{16b}$$

where $\mathrm{step} : \mathbb{R} \to \{0,1\}$, $\mathrm{step}(x) = 0$ if $x \leq 0$, and $\mathrm{step}(x) = 1$ otherwise for $x \in \mathbb{R}$.

Proof: Let $h = 0$. A sequential application of triangle inequalities to equation (15) yields

$$\big\|e_0^{(k)}(t)\big\|_\infty \leq \gamma\alpha k \cdot \exp(\lambda_{\max} t) \max\{1, t^{k-1}\} + \beta \, \|G\|^k_{\mathcal{L}_1[0,t]} \,,$$

where $k \in \mathbb{N}$ and $t \in [0.T]$. Then the condition

$$\gamma\alpha k_\mathrm{f} \cdot \exp(\lambda_{\max} T) \max\{1, T^{k_\mathrm{f}-1}\} + \beta \, \|G\|^{k_\mathrm{f}}_{\mathcal{L}_1[0,T]} \leq \varepsilon\alpha \tag{17}$$

implies $\|e_0^{(k_\mathrm{f})}(T)\|_\infty \leq \varepsilon \|e_0^{(k_\mathrm{f})}(0)\|_\infty$. For (17) to hold, we split the error between the free and forced responses as

$$\begin{aligned} & \gamma\alpha k_\mathrm{f} \cdot \exp(\lambda_{\max} T) \max\{1, T^{k_\mathrm{f}-1}\} \leq (1-\nu)\varepsilon\alpha \quad \text{and} \\ & \beta \, \|G\|^{k_\mathrm{f}}_{\mathcal{L}_1[0,T]} \leq \nu\varepsilon\alpha \,, \end{aligned}$$

where $\nu \in [0,1]$. To render the forced response error independent of the interval length T, we impose the stronger condition

$$\beta \, \|G\|^{k_\mathrm{f}}_{\mathcal{L}_1[0,T]} \leq \beta \, \|G\|^{k_\mathrm{f}}_{\mathcal{L}_1[0,\infty)} \leq \nu\alpha\varepsilon \,.$$

Since $\|G\|_{\mathcal{L}_1[0,\infty)} < 1$, the latter condition is equivalent to condition (16a). Consequently, given k_f as determined in (16a), the contractivity condition (17) holds if $\gamma\alpha k_\mathrm{f} \cdot \exp(\lambda_{\max} T) \max\{1, T^{k_\mathrm{f}-1}\} \leq (1-\nu)\varepsilon\alpha$, or equivalently, if condition (16b) holds. To conclude the proof, notice that the same reasoning can be sequentially applied to each stage $h \in \mathbb{N}$ with $\|z_h\|_\infty \leq \alpha\varepsilon^h$. ■

Due to the contractivity of the filtering error, if k_f and T are chosen to satisfy conditions (16a) and (16b) in Theorem 3.3, then Algorithm 1 is asymptotically convergent.

[1]Since the matrix exponential $\exp(A_C) = \sum_{j=0}^\infty A_C^j/j!$ always exists, we have that $\max_{j\in\mathbb{N}} A_C^j/j!$ is necessarily bounded.

B. Implementation issues, robustness, and local computation

Local verification of convergence conditions: It is possible to verify the convergence conditions in Theorem 3.1 with local information only. In particular, notice that each control center knows the pair (A_i, C_i) and selects L_i such that $A_i + L_i C_i$ is Hurwitz. Additionally, to satisfy the Schur stability condition $\rho(G(\sigma + \mathrm{j}\,\omega)) < 1$, it is sufficient[2] for each control center i to verify the following *small gain condition* for each $\omega \in \mathbb{R}$:

$$\left\| (\sigma + \mathrm{j}\,\omega I_{|\mathcal{V}_i|} - A_i - L_i C_i)^{-1} \sum\nolimits_{j=1, j\neq i}^{N} A_{ij} \right\|_\infty < 1. \quad (18)$$

If the process (1) is stable and each control center initializes $w^{(0)}(t) = 0$, then $\sigma = 0$, and (18) can be checked with local information. For an unstable process, each control center needs an estimate of the divergence rate of A to calculate σ (if A is known to be unstable, then $\|A\|_\infty$ is an estimate of its divergence rate since $\rho(A) \leq \|A\|_\infty$).

Local verification of contractivity conditions: Notice that

$$\|G\|_{\mathcal{L}_1[0,t]} = \max_{i\in\{1,\dots,N\}} \sum\nolimits_{j=1}^{N} \int_0^t \left|\mathcal{L}^{-1}\left(G_{ij}(s)\right)\right| \, d\tau, \quad (19)$$

where $G_{ij}(s) = (sI_{|\mathcal{V}_i|} - A_i - L_i C_i)^{-1} A_{ij}$. Hence, each control center can calculate a lower bound on $\|G\|_{\mathcal{L}_1[0,t]}$. Likewise, since each control center specifies the eigenvalues of $A_i + L_i C_i$, all control centers can agree on an upper bound for λ_{max}. Regarding γ, by using the notation in Theorem 3.2, we have $\gamma = c_1 c_2$. Notice that c_1 is computed by using the block-diagonal matrix $A_D + LC$, and it can be computed distributively. For $c_2 = \max_{j\in\mathbb{N}} \|A_C^j / j!\|_\infty$, we have that

$$\left\| A_C^j / j! \right\|_\infty \leq \|A_C\|_\infty^j / j! \,.$$

An upper bound for c_2, and hence on γ, can then be computed via distributed computation by using the infinity norm of A_C. Finally, if bounds on the constants α and β are given, each control center can calculate bounds on the convergence rate (15) and also lower bounds for the values k_{f} and T specified in (16a)-(16b), which guarantee error contractivity (a standard distributed leader election algorithm should be used by the agents to agree on common values for k_{f} and T [25]).

Discretization and communication errors: In Algorithm 1, the continuous time signal $w_i^{(k)}(t)$ should be transmitted at every iteration k. In practice, only an approximation $\hat{w}_i^{(k)}(t)$ of $w_i^{(k)}(t)$ can be communicated. For instance, $\hat{w}_i^{(k)}(t)$ could be a finite basis representation or a sampled approximation of $w_i^{(k)}(t)$. Furthermore, the communication channel is typically subject to noise affecting the signal $w_i^{(k)}(t)$. In either of these situations, an additional input $d^{(k)}(t) = \hat{w}_i^{(k)}(t) - w_i^{(k)}(t)$ enters the filter (11). Following the same reasoning as in Theorem 3.2, the resulting discretization error $e_{\text{dt}}^{(k)}(t)$ is bounded by

$$\left\| e_{\text{dt}}^{(k)}(t) \right\|_\infty \leq \sum\nolimits_{j=1}^{k} \|G\|_{\mathcal{L}_1[0,t]}^j \max_{\tau\in[0,t]} \left\| d^{(j)}(\tau) \right\|_\infty . \quad (20)$$

Under the convergence condition $\|G\|_{\mathcal{L}_1[0,\infty)} < 1$ (see Theorem 3.3), Algorithm 1 does not amplify errors induced by noise or discretization. Moreover, in the noise-free case, the estimation error $\|e_{\text{dt}}^{(k)}(t)\|_\infty$ can be made arbitrarily small by reducing the discretization errors $\max_{\tau\in[0,t]} \left\| d^{(k)}(\tau) \right\|_\infty$. The latter can be achieved by increasing the number of sample points or the number of basis functions; see Section IV-B for an example. An analogous analysis holds for the case of discretized measurements.

[2] Condition (18) follows since $\rho(G) \leq \|G\|_\infty$ for every $G \in \mathbb{C}^{n\times n}$.

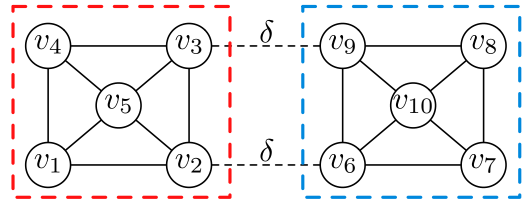


Fig. 1. A sensor network composed of two weakly-coupled subnetworks. The parameter δ specifies the coupling strength among the two parts.

Optimal observer design via distributed computation: We have considered a set of local observers (11) with output injections L_i. The output injections matrices L_i play a central role in our analysis, since they determine the local convergence rate λ_{max}, and they affect the input-output gain $\|G\|_{\mathcal{L}_1[0,t)}$ of the system $G(s) = (sI - A_D - LC)^{-1} A_C$. Moreover, the convergence estimates (15) and (16), and the discretization error bound (20) depend upon the choice of L_i.

Consider now the inverse problem of designing the local output injections L_i in order to minimize the convergence rate estimate (15) of the estimation error, the required time horizon T and number of iterations k_{f} in (16) (which follow directly from the error decay rate (15)), and the discretization error (20). In order to achieve these goals, each output injection L_i should be designed to minimize the local convergence rate λ_{max} and the induced $\mathcal{L}_\infty$-norm $\|G\|_{\mathcal{L}_1[0,t]}$ (or its upper bound $\|G\|_{\mathcal{L}_1[0,\infty)}$). In the following, we show that each of these two optimization problems can be solved by using only locally available information.

Regarding the minimization of λ_{max}, notice that $\lambda_{\text{max}} = \max_{i\in\{1,\dots,N\}} \lambda_{\text{max},\,i}$, where $\lambda_{\text{max},\,i}$ is the largest real part of the eigenvalues of $A_i + L_i C_i$. Due to the observability assumption (A2), each $\lambda_{\text{max},\,i}$ can be assigned arbitrary by choosing L_i. Likewise, let $G_i(s)$ denote the ith row of $G(s)$, then, $\|G\|_{\mathcal{L}_1[0,t]}$ is the maximum of $\|G_i\|_{\mathcal{L}_1[0,t]} = \sum_{j=1}^{n} \int_0^t \left|\mathcal{L}^{-1}\left(G_{ij}(s)\right)\right| \, d\tau$ for $i \in \{1,\dots,N\}$. Hence, by using locally available information, L_i can be chosen to minimize the induced $\mathcal{L}_\infty$-norm $\|G_i\|_{\mathcal{L}_1[0,t]}$; see [26], [27] for two computationally efficient solutions to the $\mathcal{L}_1$ disturbance attenuation problem. Finally, we remark that, even though both λ_{max} and $\|G\|_{\mathcal{L}_1[0,t]}$ can be independently minimized by locally choosing L_i, the multi-objective optimization problem of simultaneously minimizing λ_{max} and $\|G\|_{\mathcal{L}_1[0,t]}$ requires the knowledge of the global system matrices and no convex formulation is known [27]. Hence, for practical applications L_i should to be chosen to either minimize $\|G\|_{\mathcal{L}_1[0,\infty)}$ or λ_{max} (for a given $\|G\|_{\mathcal{L}_1[0,\infty)}$). Whereas the former objective minimizes the number of iterations k_{f} in (16a), the latter minimizes the time delay T in (16b).

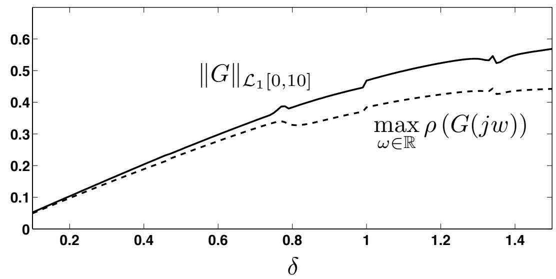


Fig. 2. Let $G(s)$ be the transfer matrix of the network in Fig.1 as defined in (8). Fig. 2 shows the worst-case spectral radius and the truncated $\mathcal{L}_1$-norm of $G(s)$ as a function of the coupling strength δ.

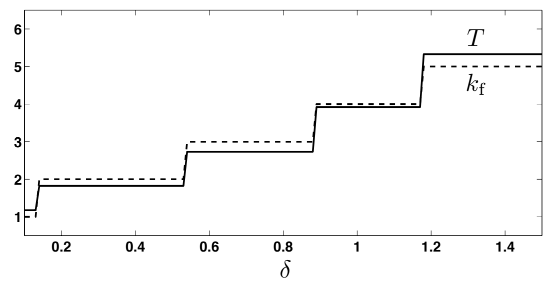


Fig. 3. For the network in Fig. 1, Fig. 3 shows the smallest values of the parameters k_{f} and T derived from equations (16a) and (16b).

IV. Illustrative Examples

We now present two numerical examples. In Section IV-A we illustrate Theorem 3.3 on a sensor network composed of two weakly coupled subnetworks. In Section IV-B we use Algorithm 1 to estimate the state of a power network.

A. Distributed state estimation for weakly-coupled networks

Consider the sensor network in Fig. 1. Let x be the vector of agent states with dynamics given by $\dot{x} = -L_{\text{network}}\, x$, where

$$L_{\text{network}} = \left[\begin{array}{ccccc|ccccc} 3 & -1 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & (3+\delta) & -1 & 0 & -1 & -\delta & 0 & 0 & 0 & 0 \\ 0 & -1 & (3+\delta) & -1 & -1 & 0 & 0 & 0 & -\delta & 0 \\ -1 & 0 & -1 & 3 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & -1 & 4 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & -\delta & 0 & 0 & 0 & (3+\delta) & -1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 3 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 3 & -1 & -1 \\ 0 & 0 & -\delta & 0 & 0 & -1 & 0 & -1 & (3+\delta) & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & 4 \end{array}\right]$$

is the network Laplacian matrix. Agents v_5 and v_{10} aim to estimate the state of agents $\{v_1, v_2, v_3, v_4, v_5\}$ and $\{v_6, v_7, v_8, v_9, v_{10}\}$, respectively. Assume that v_5 (resp. v_{10}) knows the upper (resp. lower) blocks of L_{network} (assumption (A3)), and that agent v_5 (resp. v_{10}) measures the state of agents v_1, v_4 (resp. v_7, v_8), that is, $y_5 = [x_1\, x_4]^{\mathsf{T}}$ and $y_{10} = [x_7\, x_8]^{\mathsf{T}}$.

For our simulation study, we let $\alpha = 1$, $\beta = 10$, $\varepsilon = 0.9$, and we select each injection matrix L_i so that the eigenvalues of $(A_i + L_i C_i)$ equal $\{-2, -4, -6, -8, -10\}$. The performance of our convergence and contractivity estimates are reported in Fig. 2 and in Fig. 3 as a function of δ. For instance, for $\delta = 1$ we have $\|G\|_{\mathcal{L}_1[0,10]} = 0.4686$, $\rho(G(j\omega)) = 0.3843$, $c_1 = 1.8993$, $c_2 = 1$, $T \geq 3.9207$, and $k_{\mathrm{f}} \geq 4$ ($\nu = 0.8$). To

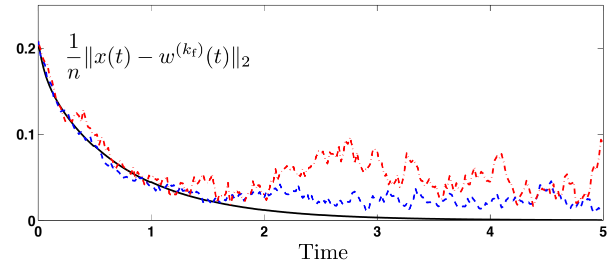


Fig. 4. This figure shows the normalized estimation error as a function of time. The black solid line corresponds to the error in the absence of communication error between control centers. The dashed blue (resp. dashed-dot red) line is obtained when the signals exchanged between control centers are corrupted by normally distributed noise with zero mean and standard deviation equal to 5 (resp. 10). For the simulation we use $T = 0.5$s (this value is smaller than its conservative estimate from equation (16b)) and $k_{\mathrm{f}} = 5$, and we set the initial guess of the neighboring state to zero.

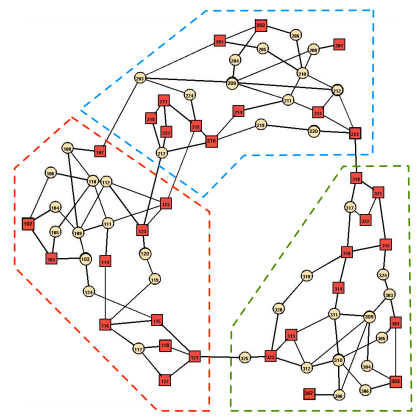


Fig. 5. This figure shows a single-line diagram of the IEEE RTS96 power network, which is composed of three weakly-coupled areas.

conclude this example, we implement our estimation algorithm for the case $\delta = 1$, $\alpha = 1$, $\beta = 10$, and $\varepsilon = 0.9$. The output injection matrices are chosen as

$$L_1 = \begin{bmatrix} -4.3717 & 0.1417 \\ -2.9767 & 3.2002 \\ -1.7779 & -4.4024 \\ 0.4706 & -7.6283 \\ 3.9937 & -2.9448 \end{bmatrix}, \; L_2 = \begin{bmatrix} -0.3080 & -2.5078 \\ -7.9606 & -1.8253 \\ -2.0158 & -4.0394 \\ 3.4468 & 1.4142 \\ -7.5953 & -0.1931 \end{bmatrix}.$$

Fig. 4 shows the result of our estimation algorithm for $T = 0.5$s and $k_{\mathrm{f}} = 5$.

B. Distributed state estimation for power networks

Consider the RTS96 power network illustrated in Fig. 5. The network data is reported in [28]. By Jacobian linearization and elimination of the algebraic equations the power network dynamics are obtained as the linear time-invariant swing dynamics

$$M\ddot{\theta} + D\dot{\theta} + Y\theta = 0\,,$$

where θ, $\dot{\theta}$ are the vectors of generator rotor angles and frequencies, M and D are the diagonal matrices of generator inertia and damping coefficients, and Y is the Kron-reduced admittance matrix [29] weighted by the linearized power flows.

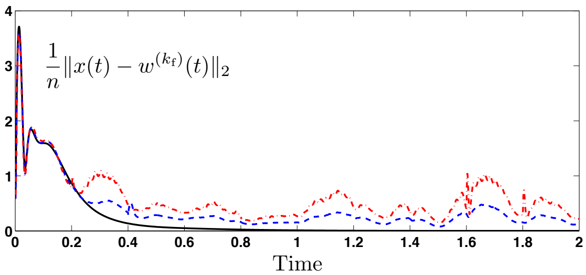


Fig. 6. This figure shows the normalized estimation error as a function of time. The black solid line corresponds to the error in the absence of discretization error. The dashed blue (resp. dashed-dot red) line is obtained when the signals exchanged between control centers are sampled with period 0.01s (resp. 0.02s). For the simulation we use $T = 0.2$s, $k_\mathrm{f} = 20$, and $h = 10$.

Assume that each area is monitored by a control center, and that control centers implement Algorithm 1 under assumptions (A1)-(A4). For assumption (A2) to be guaranteed, we let each control center measure the rotor angle of a subset of generators in its area. The results of our simulation study are reported in Fig. 6. This example demonstrates the applicability of our continuous-discrete estimation Algorithm 1 to a large-scale system and shows its effectiveness in the presence of severe discretization errors.

V. Conclusions

We have presented a continuous-discrete estimation algorithm, which allows a team of control centers, or agents, to estimate the state of a continuous system via decentralized computation and discrete-time communication. Our continuous-discrete estimation algorithm requires only local knowledge of the system and parameters, and, under a reasonable set of assumptions, it can be fully designed via decentralized computation. We have characterize stability and convergence of our continuous-discrete estimation algorithm, we have proved its robustness against discretization and communication errors, and we validated our results with two illustrative examples.

References

[1] M. Aoki and M. Li, "Partial reconstruction of state vectors in decentralized dynamic systems," *IEEE Transactions on Automatic Control*, vol. 18, no. 3, pp. 289–292, 1973.

[2] J. Speyer, "Computation and transmission requirements for a decentralized linear-quadratic-Gaussian control problem," *IEEE Transactions on Automatic Control*, vol. 24, no. 2, pp. 266–269, 1979.

[3] H. R. Hashemipour, S. Roy, and A. J. Laub, "Decentralized structures for parallel Kalman filtering," *IEEE Transactions on Automatic Control*, vol. 33, no. 1, pp. 88–94, 1988.

[4] A. Willsky, M. Bello, D. Castanon, B. Levy, and G. Verghese, "Combining and updating of local estimates and regional maps along sets of one-dimensional tracks," *IEEE Transactions on Automatic Control*, vol. 27, no. 4, pp. 799–813, 1982.

[5] D. D. Šiljak and M. Vukcevic, "Decentralization, stabilization, and estimation of large-scale linear systems," *IEEE Transactions on Automatic Control*, vol. 21, no. 3, pp. 363–366, 1976.

[6] D. D. Šiljak, *Decentralized Control of Complex Systems*. Dover Publications, 2012.

[7] M. Sundareshan and R. Elbanna, "Design of decentralized observation schemes for large-scale interconnected systems: some new results," *Automatica*, vol. 26, no. 4, pp. 789–796, 1990.

[8] M. Sundareshan, "Decentralized observation in large-scale systems," in *IEEE Conf. on Decision and Control including the 15th Symposium on Adaptive Processes*, vol. 15, 1976, pp. 659–664.

[9] M. Sundareshan and P. Huang, "On the design of a decentralized observation scheme for large-scale systems," *IEEE Transactions on Automatic Control*, vol. 29, no. 3, pp. 274–276, 1984.

[10] F. S. Cattivelli and A. H. Sayed, "Diffusion strategies for distributed Kalman filtering and smoothing," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2069–2084, 2010.

[11] J. Cortés, "Distributed Kriged Kalman filter for spatial estimation," *IEEE Transactions on Automatic Control*, vol. 54, no. 12, pp. 2816–2827, 2009.

[12] U. A. Khan and J. M. F. Moura, "Distributing the Kalman filter for large-scale systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4919–4935, 2008.

[13] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri, "Distributed Kalman filtering based on consensus strategies," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, pp. 622–633, 2008.

[14] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *IEEE Conf. on Decision and Control*, New Orleans, LA, USA, Dec. 2007, pp. 5492–5498.

[15] M. G. Bello, "Centralized and decentralized map updating and terrain masking analysis," Ph.D. dissertation, MIT, Dept. of Electrical Engineering and Computer Science, 1981.

[16] G. P. Rao and N. K. Sinha, "Continuous-time models and approaches," in *Identification of Continuous-Time Systems*, N. K. Sinha and G. P. Rao, Eds. Kluwer Academic Publishers, 1991.

[17] F. Dörfler, F. Pasqualetti, and F. Bullo, "Distributed detection of cyber-physical attacks in power networks: A waveform relaxation approach," in *Allerton Conf. on Communications, Control and Computing*, Allerton, IL, USA, Sep. 2011, pp. 1486–1491.

[18] F. Pasqualetti, F. Dörfler, and F. Bullo, "Cyber-physical security via geometric control: Distributed monitoring and malicious attacks," in *IEEE Conf. on Decision and Control*, Maui, HI, USA, Dec. 2012, pp. 3418–3425.

[19] ——, "Attack detection and identification in cyber-physical systems," *IEEE Transactions on Automatic Control*, Aug. 2012, conditionally accepted.

[20] M. Vidyasagar, *Input-Output Analysis of Large-Scale Interconnected Systems: Decomposition, Well-Posedness and Stability*. Springer, 1981.

[21] E. Lelarasmee, A. E. Ruehli, and A. L. Sangiovanni-Vincentelli, "The waveform relaxation method for time-domain analysis of large scale integrated circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 1, no. 3, pp. 131–145, 1982.

[22] U. Miekkala and O. Nevanlinna, "Convergence of dynamic iteration methods for initial value problems," *SIAM Journal on Scientific and Statistical Computing*, vol. 8, no. 4, pp. 459–482, 1987.

[23] Z. Z. Bai and X. Yang, "On convergence conditions of waveform relaxation methods for linear differential-algebraic equations," *Journal of Computational and Applied Mathematics*, vol. 235, no. 8, pp. 2790–2804, 2011.

[24] S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control Analysis and Design*, 2nd ed. Wiley, 2005.

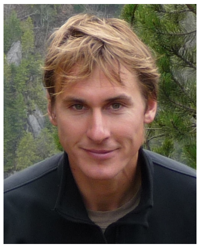
[25] N. A. Lynch, *Distributed Algorithms*. Morgan Kaufmann, 1997.

[26] F. Blanchini and M. Sznaier, "Rational L_1 suboptimal compensators for continuous-time systems," *IEEE Transactions on Automatic Control*, vol. 39, no. 7, pp. 1487–1492, 1994.

[27] C. Scherer, P. Gahinet, and M. Chilali, "Multiobjective output-feedback control via LMI optimization," *IEEE Transactions on Automatic Control*, vol. 42, no. 7, pp. 896–911, 1997.

[28] C. Grigg, P. Wong, P. Albrecht, R. Allan, M. Bhavaraju, R. Billinton, Q. Chen, C. Fong, S. Haddad, S. Kuruganty, W. Li, R. Mukerji, D. Patton, N. Rau, D. Reppen, A. Schneider, M. Shahidehpour, and C. Singh, "The IEEE Reliability Test System - 1996. A report prepared by the Reliability Test System Task Force of the Application of Probability Methods Subcommittee," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 1010–1020, 1999.

[29] F. Dörfler and F. Bullo, "Kron reduction of graphs with applications to electrical networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150–163, 2013.



Florian Dörfler (S'09) is a Doctoral Candidate in Mechanical Engineering at the University of California at Santa Barbara, and he is affiliated with the Center for Nonlinear Studies at the Los Alamos National Laboratories. He received his Diplom degree in Engineering Cybernetics from the University of Stuttgart, Germany, in 2008. He is recipient of the 2009 Regents Special International Fellowship, the 2011 Peter J. Frenkel Foundation Fellowship, the 2010 American Control Conference Student Best Paper Award, and the 2011 AACC O. Hugo Schuck Best Paper Award. His main research interest is networked control theory with applications in energy systems and robotic coordination.

Fabio Pasqualetti (S'07) is a Postdoctoral Researcher in Mechanical Engineering at the University of California, Santa Barbara. He received a Doctor of Philosophy degree in Mechanical Engineering from the University of California, Santa Barbara, in 2012, a Laurea Magistrale degree "summa cum laude" in Automation Engineering from the University of Pisa, Pisa, Italy, in 2007, and a Laurea degree "summa cum laude" in Computer Engineering from the University of Pisa, Pisa, Italy, in 2004.

His main research interest is in secure control systems, with application to multi-agent networks, distributed computing and power networks. Other interests include vehicle routing and combinatorial optimization, with application to distributed patrolling and camera surveillance.

Francesco Bullo (S'95–M'99–SM'03–F'10) is a Professor with the Mechanical Engineering Department at the University of California, Santa Barbara. He received the Laurea degree "summa cum laude" in Electrical Engineering from the University of Padova, Italy, in 1994, and the Ph.D. degree in Control and Dynamical Systems from the California Institute of Technology in 1999. From 1998 to 2004, he was an Assistant Professor with the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign.

His main research interest is multi-agent networks with application to robotic coordination, distributed computing and power networks. Other interests include vehicle routing, geometric control, and motion planning problems. He has published more than 200 papers in international journals, books and refereed conferences. He is the coauthor, with Andrew D. Lewis, of the book "Geometric Control of Mechanical Systems" (Springer, 2004) and, with Jorge Cortés and Sonia Martínez, of the book "Distributed Control of Robotic Networks" (Princeton, 2009). His students' papers were finalists for the Best Student Paper Award at the IEEE Conference on Decision and Control (2002, 2005, 2007), and the American Control Conference (2005, 2006, 2010). He is an IEEE Fellow and has served or is serving on the Editorial Boards of the "IEEE Transactions on Automatic Control," the "ESAIM: Control, Optimization, and the Calculus of Variations" and the "SIAM Journal of Control and Optimization."