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.

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{align}
\dot{x}(t) &= Ax(t), \\
y(t) &= Cx(t),
\end{align}

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

\[ \dot{w}(t) = (A + LC)w(t) - Ly(t), \]

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 centralised, 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 \( \mathcal{V} = \{1, \ldots, n\} \) 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 \( \mathcal{V} = \{\mathcal{V}_1, \ldots, \mathcal{V}_N\} \), and let \( \mathcal{G}_i = (\mathcal{V}_i, \mathcal{E}_i) \) be the \( i \)th 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

\[ A = \begin{bmatrix} A_1 & \cdots & A_{1N} \\ \vdots & \ddots & \vdots \\ A_{N1} & \cdots & A_N \end{bmatrix} = AD + AC, \]

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

We make the following assumptions:

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

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

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

\[ \dot{x}(t) = ADx(t) + ACx(t), \quad y(t) = Cx(t). \]

The dynamics of the \( i \)th subsystem then read as

\[ \dot{x}_i(t) = A_ix_i(t) + \sum_{j \in N_i^\text{in}} a_{ij}x_j(t), \]

\[ y_i(t) = C_ix_i(t), \quad i = 1, \ldots, N, \]

where \( N_i^\text{in} = \{ j \in \{1, \ldots, N\} \setminus \{i\} : A_{ij} \neq 0 \} \) denotes the in-neighbors of subsystem \( i \). We also define the set of out-neighbors as \( N_i^\text{out} = \{ j \in \{1, \ldots, N\} \setminus \{i\} : A_{ji} \neq 0 \} \).

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

(A3) **local knowledge:** the \( i \)th control center knows the local system and measurement matrices \( A_i \) and \( C_i \) as well as its neighboring matrices \( A_j, j \in N_i^\text{in} \); and 

(A4) **local communication:** the \( i \)th control center receives information from its in-neighboring control centers \( j \in N_i^\text{in} \), and it sends information to its out-neighboring control centers \( j \in N_i^\text{out} \) 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,

\[ \dot{w}(t) = (AD + LC)w(t) + ACw(t) - Ly(t), \]

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

\[ \dot{e}(t) = (AD + AC + LC)e(t). \]

Due to the observability assumption (A2), there exists \( L_i \) such that \( A_i + L_iC_i \) is Hurwitz, or equivalently, \( AD + LC \) is Hurwitz. Let \( \lambda_{\max} \in \mathbb{R} \) be the largest real part of the eigenvalues of \( AD + LC \). 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 \( \Gamma_D : \dot{e}(t) = (AD + AC)e(t) + v(t) \) and \( \Gamma_C : v(t) = ACe(t) \). Since both \( \Gamma_C \) and \( \Gamma_D \) are causal and
satisfies the spectral radius condition [24, Theorem 4.11]
\[ \rho((j\omega I - A_D - LC)^{-1}A_C) < 1 \] for all \( \omega \in \mathbb{R} \).

If condition (9) holds, then \( G(j\omega) \) 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 \( G(j\omega) \) is Schur stable for all \( \omega \in \mathbb{R} \). Then, for all \( x(0) \) and \( w(0) \), it holds \( \lim_{t \to \infty} (w(t) - x(t)) = 0 \).

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 Gauss-Jacobi waveform relaxation method applied to the filter (6) yields the waveform relaxation iteration
\[ \dot{w}^{(k)}(t) = (A_D + LC)w^{(k)}(t) + A_Cw^{(k-1)}(t) - Ly(t), \]
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 on \( k \) in the variable \( w \). On the other hand, for fixed \( k \), (10) is a continuous-time dynamical system in the variable \( w^{(k)} \) and \( A_Cw^{(k-1)} \) 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 \( \lim_{k \to \infty} \max_{t \in [0,T]} \| w^{(k)}(t) - w(t) \| = 0 \), 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 appropriately 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 \( G(\sigma + j\omega) \) 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 decentralized 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.

### 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 each 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)
\[ w^{(k)}_h(t) = (A_D + LC)w^{(k)}_h(t) + A_Cw^{(k-1)}_h(t) - Ly(t + hT), \]
where \( t \in [0,T], k \in \{1,\ldots,k_t\} \) for some \( k_t \in \mathbb{N}_{>0} \), and \( L = \text{blkdiag}(L_{1},\ldots,L_{N}) \). Third and finally, control centers exchange their estimated state and iterate over \( h \). The filter (11) is initialized with \( w^{(0)}_h(0) = w_0 = [w^{(0)}_{1,0}, \ldots, w^{(0)}_{N,0}]^T \) and \( w^{(k)}_h(0) = w^{(k)}_{h-1}(T) \) for \( h \geq 1 \) and \( k = 1,\ldots,k_t \). Additionally, \( w^{(0)}_h(t) \) is the initial profile vector with components \( w^{(0)}_{ij} \), where \( w^{(0)}_{ij}(t) \) is a guess of control center \( i \) for the state \( x_j(t) \) in the interval \( t \in [hT, (h+1)T]\) where 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 \( w_{ij,h} \) and the continuous-time measurements \( y_{ij,h}(t) = y_{ij}(t + hT) \) need to be stored by the \( i\)-th

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

<table>
<thead>
<tr>
<th>Input</th>
<th>Matrices ( A_i, C_i, A_{ij} ) with ( j \in \mathbb{N}_{&gt;0}, L_i ) (see Section II-A);</th>
</tr>
</thead>
<tbody>
<tr>
<td>Param.</td>
<td>Horizon ( T \in \mathbb{R}_{&gt;0} ), Number of iterations ( k_t ), State ( w_0 );</td>
</tr>
<tr>
<td>Require</td>
<td>Error contractivity with ( k_t ) and ( T ) (see Theorem 3.1);</td>
</tr>
<tr>
<td>1 for ( h = 0,1,\ldots ) do</td>
<td>Collect measurements ( y_{ij,h}(t) = y_j(t + hT) ) with ( t \in [0,T] );</td>
</tr>
<tr>
<td>2 Select initial profile ( w_{ij,0}^{(0)} ) with ( t \in [0,T], j \in \mathbb{N}_{&gt;0} );</td>
<td></td>
</tr>
<tr>
<td>3 for ( k = 1,\ldots,k_t ) do</td>
<td>Integrate the differential equation ( w^{(k)}<em>{ij,h}(t) = (A_i + L_iC_i)w^{(k)}</em>{ij,h}(t) + \sum_{j \in \mathbb{N}<em>{&gt;0}} A</em>{ij}w^{(k-1)}<em>{ij,h} - L</em>{ij}y_{ij,h} );</td>
</tr>
<tr>
<td>4 Transmit ( w^{(k)}_{ij,h} ) to control centers ( N_i^{\text{out}} );</td>
<td></td>
</tr>
<tr>
<td>5 Receive ( w^{(k)}_{ij,h} ) from control centers ( N_i^{\text{in}} );</td>
<td></td>
</tr>
<tr>
<td>6 Update initial state as ( w_{ij,0} = w^{(k_t)}_{ij}(T) );</td>
<td></td>
</tr>
</tbody>
</table>
control center. Finally, our method generates a state estimate with delay $T$, and control centers need to perform $k_t$ 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 $k \in \{1, \ldots, k_t\}$, $t \in [0, T]$, $h \in \mathbb{N}$. 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 $e_h^{(k)}(t)$ 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 $h \in \mathbb{N}$

$$\lim_{T \to \infty, k_t \to \infty} \|e_h^{(k)}(T)\|_\infty = 0;$$

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

$$\|e_h^{(k)}(T)\|_\infty \leq \varepsilon \|e_h^{(k)}(0)\|_\infty = \varepsilon \|e_h^{(k)}(T)\|_\infty;$$

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

$$\lim_{h \to \infty} \left(\|e_h^{(k)}(t)\|_\infty / \max_{\tau \in [0,t]} \|e_h^{(k)}(\tau)\|_\infty\right) \leq \eta^k.$$

In what follows we show that for sufficiently large values of $k_t$ and $T$, similar conditions as in Subsection II-B guarantee stage-wise convergence, $\varepsilon$-contractivity, and asymptotic convergence. We also derive explicit bounds on $k_t$ 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 $w_h^{(k)}(0) = w_h^{(k)}(T)$. Hence, we start by analyzing the error dynamics resulting from the distributed filter (11). Define the augmented filter error $e_h(t) = [e_h^{(1)T} \ldots e_h^{(k)T}]^T$, and consider the error system

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

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

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 $t \in [0, T]$, $k \in \{1, \ldots, k_t\}$, and $h \in \mathbb{N}$. Assume that each exponentially-scaled initial profile error $t \to \exp(-\mu t)(x(t + hT) - w_h^{(0)}(t))$ 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 $G(\sigma + j\omega)$ 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 $k_t \in \mathbb{N}$ and $T \in \mathbb{R}_{>0}$ such that the estimation error is $\varepsilon$-contractive.

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

$$e_h^{(k)}(t) = \mathcal{L}^{-1}(\mathcal{G}(s)^h) \ast e_h^{(0)}(t) + E_h \exp(A_h t) e_h(0),$$

(13)

where $\mathcal{L}^{-1}$ denotes the inverse Laplace transform, $\ast$ denotes the convolution operator, $G(s)$ is defined in equation (8), and $E_k = [0_{n \times (k-1)n}]$ $I_n$ 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 $e_h^{(k)}(T)$ converges as $T \to \infty$ and $k_t \to \infty$ if and only if the error system (12) with system matrix $A_e$ is stable and the input-output map $\mathcal{L}^{-1}(\mathcal{G}(s)^h)$ from $e_h^{(0)}(t)$ to $e_h^{(k)}(t)$ vanishes. Equivalently, the error system (12) is asymptotically stable if and only if $A_e + LC$ is Hurwitz (since $A_e$ is block-triangular), and the input-output map $\mathcal{L}^{-1}(\mathcal{G}(s)^h)$ 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 $\{s \in \mathbb{C} : \Re(s) \geq \sigma\}$, see also [23, Section 3 and Appendix] and [22, Section 2]. Hence, the input-output map $\mathcal{L}^{-1}(\mathcal{G}(s)^h)$ vanishes if and only if $\rho(G(s)) < 1$ for all $s \in \{s \in \mathbb{C} : \Re(s) \geq \sigma\}$.

To prove the equivalence of statements (i) and (ii).

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

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

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

where \( \|A\|_\infty \) is the induced \( \infty \)-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 \( y(t) = G(t) \ast u(t) \), that is,

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

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

\[
\|[e^j_h(t)]\|_\infty \leq \sum_{j=0}^{k-1} \gamma \cdot \exp(\lambda_{\max} t) j^\tau j! A_j^C z_h .
\]

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

\[
\|[e^j_h(t)]\|_\infty \leq \|f_h(t)\|_\infty + \|L^1 - (G(s))^k \ast (e^j(0) t)\|_\infty .
\]

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

\[
\|f_h(t)\|_\infty \leq \sum_{j=0}^{k-1} \|\exp((A_D + LC) t)\|_{\infty} \cdot \frac{t^j}{j!} A_j^C z_h .
\]

where \( \gamma > 0 \) is a constant independent of \( k \) and \( t \). Moreover, if the estimation error is \( \varepsilon \)-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:

\[
f(t) = E_k \exp(A_C t) e^j(0) = \sum_{j=0}^{k-1} \exp((A_D + LC) t)^j j! A_j^C z_h .
\]

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 \( \max_{j \in \mathbb{N}} A_j^C/j! \) always exists, we have that \( \max_{j \in \mathbb{N}} A_j^C/j! \) is necessarily bounded.

Since 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 \( k_t \) 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 \( \alpha = \|z_0\|_\infty = \|x(0) - w(0) u(0)\|_\infty \) for the initial error in the first stage, and we assume the initial profile errors to be uniformly \( \beta \)-bounded, that is, there is \( \beta > 0 \) such that \( \max_{t \in [0,T]} \|x(t) + h T - w^k(0) u(t)\|_\infty \leq \beta \) 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 \( \beta \)-bounded. For every \( \varepsilon \in (0,1) \), \( h \in \mathbb{N} \), and \( \nu \in [0,1] \), the estimation error is \( \varepsilon \)-contractive if

\[
k_t \geq \frac{\log(\beta) - \log(\nu \varepsilon)}{\log(\|G\|_{\mathcal{L}_1[0,\infty)})} ,
\]

(16a)

\[
\lambda_{\max}[T - \text{step}(T - 1)(k_t - 1) \log(T)] \geq 1 ,
\]

(16b)

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

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

\[
\|[e^j_h(t)]\|_\infty \leq \gamma_k \cdot \exp(\lambda_{\max}) \max\{1, t^{k-1}\} \ast \|G\|_{\mathcal{L}_1[0,t]}^k ,
\]

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

\[
\gamma_k t_k \cdot \exp(\lambda_{\max}) \max\{1, T^{k-1}\} \ast \|G\|_{\mathcal{L}_1[0,t]}^k \leq \varepsilon \alpha
\]

implies \( \|[e^j_h(t)]\|_\infty \leq \varepsilon \|[e^j(0)]\|_\infty \). For (17) to hold, we split the error between the free and forced responses as

\[
\gamma_k t_k \cdot \exp(\lambda_{\max}) \max\{1, T^{k-1}\} \ast \|G\|_{\mathcal{L}_1[0,t]}^k \leq (1 - \nu) \varepsilon \alpha \]

and

\[
\|G\|_{\mathcal{L}_1[0,t]}^k \leq (1 - \nu) \varepsilon \alpha ,
\]

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

\[
\|G\|_{\mathcal{L}_1[0,T]}^k \leq (1 - \nu) \varepsilon \alpha ,
\]

Since \( \|G\|_{\mathcal{L}_1[0,\infty)} < 1 \), the latter condition is equivalent to condition (16a). Consequently, given \( k_t \) as determined in (16a), the contractivity condition (17) holds if \( \gamma_k t_k \cdot \exp(\lambda_{\max}) \max\{1, T^{k-1}\} \ast \|G\|_{\mathcal{L}_1[0,t]}^k \leq (1 - \nu) \varepsilon \alpha \), 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 \( \|z_h\|_\infty \leq \alpha \).
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 $A_i + L_iC_i$ is Hurwitz. Additionally, to satisfy the Schur stability condition $\rho(G(\sigma + j\omega)) < 1$, it is sufficient for each control center $i$ to verify the following small gain condition for each $\omega \in \mathbb{R}$:

$$
\| (\sigma + j\omega)I_{\mathcal{V}_i} - A_i - L_iC_i \|_{\infty}^{-1} \sum_{j=1, j \neq i}^{N} A_{ij} \|_{\infty} < 1. \tag{18}
$$

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 $\sigma$ (if $A$ is known to be unstable, then $\| A \|_{\infty}$ is an estimate of its divergence rate since $\rho(A) < \| A \|_{\infty}$).

Local verification of contractivity conditions: Notice that

$$
\| G \|_{\mathcal{L}_1[0,t]} = \max_{i \in \{1, \ldots, N\}} \sum_{j=1}^{N} \int_{0}^{t} \| \mathcal{L}^{-1} (G_{ij}(s)) \| \, ds, \tag{19}
$$

where $G_{ij}(s) = (sI_{\mathcal{V}_i} - A_i - L_iC_i)^{-1}A_{ij}$. 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 $A_i + L_iC_i$, all control centers can agree on an upper bound for $\lambda_{\text{max}}$. Regarding $\gamma$, by using the notation in Theorem 3.2, we have $\gamma = c_1c_2$. Notice that $c_1$ is computed by using the block-diagonal matrix $A_D + LC$, and it can be computed distributively. For $c_2 = \max_{i \in \{1, \ldots, N\}} \| A_{C,i} \|_{\infty} / j_i$, we have that

$$
\| A_{C,i} \|_{\infty} / j_i \leq \| AC \|_{\infty} / j_i. \tag{20}
$$

An upper bound for $c_2$, and hence on $\gamma$, can then be computed via distributed computation by using the infinity norm of $AC$. Finally, if bounds on the constants $\alpha$ and $\beta$ are given, each control center can calculate bounds on the convergence rate (15) and also lower bounds for the values $k_t$ 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 $k_t$ and $T$ [25]).

Discretization and error estimation: In Algorithm 1, the continuous time signal $w_i^{(k)}(t)$ should be transmitted at every iteration $k$. In practice, only an approximation $\tilde{w}_i^{(k)}(t)$ of $w_i^{(k)}(t)$ can be communicated. For instance, $\tilde{w}_i^{(k)}(t)$ 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 $d_i^{(k)}(t) = \tilde{w}_i^{(k)}(t) - w_i^{(k)}(t)$ enters the filter (11). Following the same reasoning as in Theorem 3.2, the resulting discretization error $\epsilon_{\text{dis}}(t)$ is bounded by

$$
\| \epsilon_{\text{dis}}^{(k)}(t) \|_{\infty} \leq \sum_{j=1}^{k} \| G \|_{\mathcal{L}_1[0,t]} \| \mathcal{L}^{-1} (d_j^{(k)}(t)) \|_{\infty}. \tag{20}
$$

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 $\| \epsilon_{\text{dis}}^{(k)}(t) \|_{\infty}$ can be made arbitrarily small by reducing the discretization errors $\max_{\tau \in [0,t]} \| d^{(k)}(\tau) \|_{\infty}$. 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.

**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 $\lambda_{\text{max}}$, 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 $k_t$ in (16) (which follow directly from the error decay rate (15)), and the discretization error bound (20). In order to achieve these goals, each output injection $L_i$ should be designed to minimize the local convergence rate $\lambda_{\text{max}}$ and the induced $L_{\infty}$-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 $\lambda_{\text{max}}$, notice that $\lambda_{\text{max}} = \max_{i \in \{1, \ldots, N\}} \lambda_{\text{max},i}$, where $\lambda_{\text{max},i}$ is the largest real part of the eigenvalues of $A_i + L_iC_i$. Due to the observability assumption (A2), each $\lambda_{\text{max},i}$ can be assigned arbitrarily by choosing $L_i$. Likewise, let $G_i(s)$ denote the $i$th row of $G(s)$, then $\| G \|_{\mathcal{L}_1[0,t]}$ is the maximum of $\| G_i \|_{\mathcal{L}_1[0,t]}$:

$$
\sum_{j=1}^{n} \int_{0}^{t} \| \mathcal{L}^{-1} (G_{ij}(s)) \| \, ds \leq \{ i \in \{1, \ldots, N\} \}. \tag{20}
$$

Hence, by using locally available information, $L_i$ can be chosen to minimize the induced $L_{\infty}$-norm $\| G_i \|_{\mathcal{L}_1[0,t]}$; see [26], [27] for two computationally efficient solutions to the $L_1$ disturbance attenuation problem. Finally, we remark that, even though both $\lambda_{\text{max}}$ and $\| G \|_{\mathcal{L}_1[0,t]}$ can be independently minimized by locally choosing $L_i$, the multi-objective optimization problem of simultaneously minimizing $\lambda_{\text{max}}$ 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 be chosen to either minimize $\lambda_{\text{max}}$ or $\| G \|_{\mathcal{L}_1[0,t]}$. Whereas the former objective minimizes the number of iterations $k_t$ in (16a), the latter minimizes the time delay $T$ in (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

$$L_{\text{network}} = \begin{bmatrix}
3 & -1 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\
-1 & (3+\delta) & -1 & 0 & -1 & -1 & -1 & 0 & 0 & 0 \\
0 & -1 & (3+\delta) & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\
-1 & 0 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & (3+\delta) & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & -1 & -1 & -1 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & -1 & -1 & -1 & 4
\end{bmatrix}$$

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 $L_{\text{network}}$ (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, $y_5 = [x_1 x_4]^T$ and $y_{10} = [x_7 x_8]^T$.

For our simulation study, we let $\alpha = 0.1$, $\beta = 10$, $\varepsilon = 0.9$, 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 $\delta$. 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 $k_f \geq 4$ ($\nu = 0.8$). To 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

$$L_1 = \begin{bmatrix}
-4.3717 & 0.1417 \\
-2.9767 & 3.2002 \\
-1.7779 & -4.4024 \\
0.4706 & -7.6283 \\
3.9937 & -2.9448
\end{bmatrix}, \quad L_2 = \begin{bmatrix}
-0.3080 & -2.5078 \\
-7.9606 & -1.8253 \\
-2.0158 & -4.0394 \\
3.4468 & 1.4142 \\
-7.5935 & -0.1931
\end{bmatrix}$$

Fig. 4 shows the result of our estimation algorithm for $T = 0.5s$ and $k_f = 5$.

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 $\theta$, $\dot{\theta}$ 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.
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

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.”