Distributed bounded-error state estimation based on practical robust positive invariance

Stefano Riverso, Daria Rubini and Giancarlo Ferrari-Trecate

United Technologies Research Center Ireland, 4th Floor, Penrose Business Center, Penrose Wharf, Cork, Ireland.

Dipartimento di Ingegneria Industriale e dell’Informazione, Università degli Studi di Pavia, Pavia, Italy

(Received 00 Month 20XX; accepted 00 Month 20XX)

We propose a state estimator for linear discrete-time systems composed by coupled subsystems affected by bounded disturbances. The architecture is distributed in the sense that each subsystem is equipped with a local state estimator that exploits suitable pieces of information from parent subsystems. Furthermore, each local estimator reconstructs the state of the corresponding subsystem only. Differently from methods based on moving horizon estimation, our approach does not require the on-line solution to optimization problems. Our state-estimation scheme, which is based on the notion of practical robust positive invariance developed in (Raković, Kern, & Friedeisen, 2011), also guarantees satisfaction of constraints on local estimation errors and it can be updated with a limited computational effort when subsystems are added or removed.

Keywords: distributed state estimation, robust positive invariant sets, bounded error estimation.

1. Introduction

In modern engineering there are several examples of applications composed by a large number of subsystems and for which centralized operations can be very expensive. For instance, the use of centralized controllers and state estimators can be hampered by the complexity of the design stage or by demanding computational and communication requirements for on-line operations. An alternative approach is to decompose the plant into physically coupled subsystems and design local controllers and state estimators associated to each subsystem. In these cases, local devices can operate in parallel using computational resources available at subsystem locations. Approaches with these features have been studied since the 1970’s under the banner of decentralized and distributed control.

Available distributed state estimation schemes can be classified according to different criteria. First, the goal of a local state estimator can be either to reconstruct the state of the overall plant (Alriksson & Rantzer, 2006; Kamgarpour & Tomlin, 2008; Carli, Chiuso, Schenato, & Zampieri, 2008; Farina, Ferrari-Trecate, & Scattolini, 2010a; Battistelli, Chisci, Morrocchi, & Papi, 2011; Xu, Gupta, & Fischione, 2013; Millán et al., 2013) or a subset of it (Vadigepalli & Doyle III, 2003; Khan & Moura, 2008; Stankovic, Stankovic, & Stipanovic, 2009; Farina, Ferrari-Trecate, & Scattolini, 2010b; Farina & Scattolini, 2011; Haber & Verhaegen, 2013; Riverso, Farina, Parisini, & Polycarpou, 2013). In particular, several contributions on distributed Kalman filtering address the former problem. In the latter

Corresponding author. Email: riverss@utrc.utc.com

The research leading to these results has received funding from the European Union Seventh Framework Programme [FP7/2007-2013] under grant agreement n° 257462 HYCON2 Network of excellence.
case, estimators are termed partition-based if subsystems have non-overlapping states and a local estimator reconstructs the state of the corresponding subsystem only. Second, the topology of the communication network connecting local estimators can be different, ranging from all-to-all communication (Vadigepalli & Doyle III, 2003) to transmission of information only from each subsystem to its children, i.e. subsystems influenced by it (Khan & Moura, 2008; Stankovic et al., 2009; Farina et al., 2010b; Farina & Scattolini, 2011; Millán et al., 2013; Riverso, Farina, et al., 2013; Zhang, 2010; Boem et al., 2011). Third, local estimators can be based on unconstrained models (Vadigepalli & Doyle III, 2003; Khan & Moura, 2008; Stankovic et al., 2009; Zhang, 2010; Boem et al., 2011) or can cope with constraints on system variables such as disturbances, states (Farina et al., 2010b) and estimation errors (Farina & Scattolini, 2011).

In this paper we propose a novel partition-based state estimator for linear discrete-time subsystems affected by bounded disturbances. Research on partitioned-based architectures is motivated by several applications ranging from chemical processes (Schneider, Scheu, & Marquardt, 2013) to electric systems (Bolognani, Carli, & Todescato, 2014) and automated highways (Zhang, 2010). More in general, partition-based architectures allow one to generalize decentralized or distributed state-feedback controllers to the output-feedback case without introducing centralized computations due to the state-estimation task. Similarly to the method proposed in (Farina et al., 2010b; Farina & Scattolini, 2011) and (Riverso, Farina, et al., 2013), our scheme is distributed in the sense that computation of local state estimates can be performed in parallel but only after each estimator has received suitable pieces of information from parent subsystems. Moreover, as in (Farina & Scattolini, 2011), state estimators account for constraints on subsystem disturbances and guarantee the fulfillment of a priori specified constraints on local estimation errors. Differently from the scheme in (Farina et al., 2010b), that is based on moving horizon estimation, and similarly to (Farina & Scattolini, 2011; Riverso, Farina, et al., 2013), local estimators have a Luenberger structure and therefore do not require the on-line solution to optimization problems. Furthermore, most operations needed for the design of a local estimator can be performed using computational resources collocated with the corresponding subsystem and the only centralized task requires the analysis of a system whose order is equal to the number of subsystems.

In order to guarantee convergence of state estimates in absence of disturbances and fulfillment of prescribed constraints on the estimation error, we rely on the notion of practical robust positive invariance developed in (Raković et al., 2011) that is applied to the error dynamics. We also highlight that most of the appealing computational features of our method directly follow from results provided in (Raković et al., 2011) for the case of polytopic constraints. Since practical robust positive invariance implies worst-case robustness against the propagation of errors between subsystems, our design method involves some degree of conservatism. In the attempt of maximizing chances of successful design, we provide guidelines on the choice of local estimator parameters. We also show that when subsystems are added or removed, the state estimation scheme can be updated with limited efforts. More in detail, we prove that, in order to preserve convergence and fulfillment of constraints on estimation errors, (i) the plug-in of a subsystem requires the decentralized design of local estimators for the subsystem and its children only, besides the re-execution of the centralized analysis task mentioned above; (ii) the unplugging of a subsystem does not require any update. Compared to the distributed state estimator proposed in (Farina & Scattolini, 2011), our scheme has several distinctive features. First, the use of the notion of practical robust positive invariance instead of the more standard concept of robust positive invariance, allows us to achieve a less conservative design procedure (see (Raković, Kern, & Findeisen, 2010) for a discussion on the degree of conservativeness of various invariance concepts). Second, our local estimators can take advantage of the knowledge of parents’ outputs and this can be fundamental for successful estimator design, as demonstrated in Section 6 through examples. Third, the method in (Farina & Scattolini, 2011) requires to analyze in a centralized fashion the stability of a system whose order is equal to the sum of the orders of all subsystems.

The paper is structured as follows. Local state estimators are described in Section 2. In Section 3 we introduce practical robust decentralized invariance and show how it can be applied for guar-
anticipe convergence of estimators and constraint satisfaction. In Section 4 we detail the design of local estimators. Section 5 describes how to retune the estimators when subsystems are added or removed from the network. Furthermore, a comparison with the plug-and-play design approach in (Rivero, Farina, et al., 2013) and distributed estimators in (Zhang, 2010) and (Boem et al., 2011) is conducted. In Section 6 we illustrate the use of the distributed state estimator for reconstructing the states of a string of masses and of a power network system. In addition, we compare our method with the state estimation scheme in (Farina & Scattolini, 2011). Section 7 is devoted to conclusions. A preliminary version of this work has been presented at the 12th European Control Conference (Rivero, Rubini, & Ferrari-Trecate, 2013).

Notation. We use $a : b$ for the set of integers $\{a, a + 1, \ldots, b\}$. The symbol $R^n_+$ stands for the vectors in $R^n$ with nonnegative elements. The column vector with $s$ components $v_1, \ldots, v_s$ is $v = (v_1, \ldots, v_s)$. The symbol $\oplus$ denotes the Minkowski sum, i.e. $A = B \oplus C$ if and only if $A = \{a : a = b + c, \ b \in B, \ c \in C\}$. Moreover, $\bigoplus_{i=1}^{s} G_i = G_1 \oplus \cdots \oplus G_s$. The symbol $1_{\alpha}$ (resp. $0_{\alpha}$) denotes a matrix or a column vector with all $\alpha$ elements equal to 1 (resp. 0). Given a matrix $A \in R^{n \times n}$, with entries $a_{ij}$ its entry-wise 1-norm is $||A||_1 = \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|$ and its Frobenius norm is $||A||_F^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2$. Given a vector $x \in R^n$ and a set $S \subseteq R^n$, $\text{dist}(x, S) = \inf_{s \in S} ||x - s||$. The pseudo-inverse of a matrix $A \in R^{m \times n}$ is denoted with $A^\dagger$.

The set $X \subseteq R^n$ is Robust Positively Invariant (RPI) for $x_{t+1} = f(x(t), w(t))$, $w(t) \in W \subseteq R^m$ if $x(t) \in X \Rightarrow f(x(t), w(t)) \in X, \forall w(t) \in W$. The RPI set $X$ is maximal if it includes every other RPI set. The set $X \subseteq R^n$ is positively invariant for $x_{t+1} = f(x(t))$ if $x(t) \in X \Rightarrow f(x(t)) \in X$. The set $X \subseteq R^n$ is a $\lambda$-contractive RPI set, with $\lambda \in [0, 1)$ for $x_{t+1} = f(x(t))$ if $x(t) \in X \Rightarrow f(x(t)) \in \lambda X$. A $C$-set is a set that is compact, convex and contains the origin.

2. Distributed State Estimator (DSE)

We consider a discrete-time Linear Time Invariant (LTI) system

$$
\begin{align*}
x^+ &= Ax + Bu + Dw \\
y &= Cx
\end{align*}
$$

where $x \in R^n$, $u \in R^m$, $y \in R^p$ and $w \in R^r$ are the state, the input, the output and the disturbance, respectively, at time $t$ and $x^+$ stands for $x$ at time $t + 1$. The state is partitioned into $M$ state vectors $x[i] \in R^{n_i}$, $i \in M = 1 : M$ such that $x = (x[1], \ldots, x[M])$ and $n = \sum_{i \in M} n_i$. Similarly, the input, the output and the disturbance are partitioned into $M$ vectors $u[i] \in R^{m_i}$, $y[i] \in R^{p_i}$, $w[i] \in R^{r_i}$, $i \in M$ such that $u = (u[1], \ldots, u[M])$, $m = \sum_{i \in M} m_i$, $y = (y[1], \ldots, y[M])$, $p = \sum_{i \in M} p_i$, $w = (w[1], \ldots, w[M])$ and $r = \sum_{i \in M} r_i$.

We assume the dynamics of the $i$-th subsystem is given by

$$
\begin{align*}
\Sigma[i]: \quad x^+[i] &= A_{ii}x[i] + B_{ii}u[i] + \sum_{j \in N_i} A_{ij}x[j] + D_{ij}w[i] \\
y[i] &= C_{ii}x[i]
\end{align*}
$$

where $A_{ii} \in R^{n_i \times n_i}$, $i, j \in M$, $B_{ii} \in R^{n_i \times m_i}$, $D_{ii} \in R^{n_i \times r_i}$, $C_{ii} \in R^{p_i \times n_i}$ and $N_i$ is the set of parents of subsystem $i$ defined as $N_i = \{j \in M : A_{ij} \neq 0, i \neq j\}$. Since $y[i]$ depends on the local state $x[i]$ only, subsystems $\Sigma[i]$ are output-decoupled and then $C = \text{diag}(C_1, \ldots, C_M)$. Similarly, subsystems $\Sigma[i]$ are input- and disturbance-decoupled, i.e. $B = \text{diag}(B_1, \ldots, B_M)$ and $D = \text{diag}(D_1, \ldots, D_M)$. We also assume

$$
w[i] \in W_i \subset R^{r_i}
$$
In this section we propose a DSE for (1). We define for $i \in \mathcal{M}$ the local state estimator

$$
\tilde{\Sigma}_{[i]} : \quad \tilde{x}_{[i]} = A_{ii}\tilde{x}_{[i]} + B_i u_{[i]} - L_{ii}(y_{[i]} - C_i \tilde{x}_{[i]}) + \sum_{j \in \mathcal{N}_i} A_{ij} \tilde{x}_{[j]} - \sum_{j \in \mathcal{N}_i} \delta_{ij} L_{ij}(y_{[j]} - C_j \tilde{x}_{[j]})
$$

(4)

where $\tilde{x}_{[i]} \in \mathbb{R}^{n_i}$ is the state estimate, $L_{ij} \in \mathbb{R}^{n_i \times p_j}$ are gain matrices and $\delta_{ij} \in \{0,1\}$. Hereafter we assume $\delta_{ij} = 0$ and $L_{ij} = 0$ if $j \not\in \mathcal{N}_i$. This implies that $\tilde{\Sigma}_{[i]}$ depends only on local variables ($\tilde{x}_{[i]}$, $u_{[i]}$ and $y_{[i]}$) and parents’ variables ($\tilde{x}_{[j]}$ and $y_{[j]}$, $j \in \mathcal{N}_i$). Binary parameters $\delta_{ij}$, $j \in \mathcal{N}_i$ can be chosen to take advantage of the knowledge of parents’ outputs ($\delta_{ij} = 1$) or to reduce the amount of information received from parents ($\delta_{ij} = 0$).

Defining the state estimation error as

$$
e_{[i]} = x_{[i]} - \tilde{x}_{[i]},
$$

(5)

from (2), (4) and (5), we obtain the local error dynamics

$$
e_{[i]}^+ = \bar{A}_{ii} e_{[i]} + \sum_{j \in \mathcal{N}_i} \bar{A}_{ij} e_{[j]} + D_i w_{[i]}
$$

(6)

where $\bar{A}_{ii} = A_{ii} + L_{ii} C_i$ and $\bar{A}_{ij} = A_{ij} + \delta_{ij} L_{ij} C_j$, $i \neq j$. Our main goal is to solve the following problem.

**Problem 1:** Design local state estimators $\tilde{\Sigma}_{[i]}$, $i \in \mathcal{M}$ that

(a) are nominally convergent, i.e. when $W = \{0\}$ it holds

$$
||e_{[i]}(t)|| \rightarrow 0 \text{ as } t \rightarrow \infty
$$

(7)

(b) guarantee

$$
e_{[i]}(t) \in E_i, \forall t \geq 0
$$

(8)

where $E_i \subseteq \mathbb{R}^{n_i}$ are prescribed sets containing the origin in their interior. □

Defining the collective variable $e = (e_{[1]}, \ldots, e_{[M]}) \in \mathbb{R}^n$, from (26) one obtains the collective dynamics of the estimation error

$$
e^+ = \bar{A} e + Dw
$$

(9)

where the matrix $\bar{A}$ is composed by blocks $\bar{A}_{ij}$, $i, j \in \mathcal{M}$.

We equip system (9) with constraints $e \in E = \prod_{i \in \mathcal{M}} E_i$ and $w \in W = \prod_{i \in \mathcal{M}} W_i$. In Section 3 we address Problem 1 under the following assumptions

**Assumption 1:** The matrices $\bar{A}_{ii}$, $i \in \mathcal{M}$ are Schur.

**Assumption 2:** The sets $E_i$ and $W_i$, $i \in \mathcal{M}$ are $C$-sets.

Assumption 1 is equivalent to the detectability of pairs $(A_{ii}, C_i)$ and it is needed for guaranteeing stability of the error dynamics. We highlight that if gains $L_{ij}$ are such that $\bar{A}$ is Schur, then property (7) holds. In view of Assumption 2, only bounded disturbances will be considered in the sequel. This is however a common setting for estimation schemes based on worst-case disturbance analysis (Farina et al., 2010b; Farina & Scattolini, 2011; Zhang, 2010; Boem et al., 2011). Furthermore, Assumption 2 requires boundedness of estimation errors, which is needed in output-feedback
controllers based on, e.g., model predictive control (Mayne, Raković, Findeisen, & Allgöwer, 2006). In particular, sets $E_i$ are design parameters fixed by the user and chosen on the basis of the desired accuracy of state estimates, which depends on the specific application. If, in addition to Assumption 1, Assumption 2 holds, then there is an RPI set $\Omega \subseteq E$ for the constrained system (9) (see (Kolmanovsky & Gilbert, 1998)) and $e(0) \in \Omega$ guarantees property (8). Remarkably, when sets $E_i$ and $W_i$ are polytopes, an RPI set $\Omega$ can be found solving a Linear Programming (LP) problem (Raković & Baric, 2010). However the LP problem includes the collective model (1) in the constraints and computations become prohibitive for large $n$.

In absence of coupling between subsystems (i.e. $A_{ij} = 0, \ i \neq j$) the estimator dynamics (4) and error dynamics (26) are decoupled as well. Therefore, under Assumptions 1 and 2, properties (7) and (8) can be guaranteed computing RPI sets $\Omega_i \subseteq E_i$ for each local error dynamics and requiring $e_i(0) \in \Omega_i$. Furthermore, if $E_i$ and $W_i$ are polytopes, the computation of sets $\Omega_i, i \in M$ amounts to the solution of $M$ LP problems that can be solved in parallel using computational resources collocated with subsystems. In order to propose a partially decentralized design procedure in presence of coupling between subsystems, one has to take into account how coupling propagates errors between subsystems. As we will show in the next section, the notion of practical robust positive invariance, proposed in (Raković et al., 2011) allows one to study precisely this issue and offers a computationally feasible, yet conservative, procedure for solving Problem 1.

3. Practical robust positive invariance for state estimation

In this section, we show how the main results of (Raković et al., 2011), applied to the error dynamics (26) equipped with constraints (3) and (8), allow one to guarantee properties (7) and (8).

Given a collection of sets $S = \{S_i, i \in M\}, S_i \subseteq \mathbb{R}^{n_i}$ and a set $\Theta \subseteq \mathbb{R}^M_+$, we define a parameterized family of sets $S(S, \Theta) = \{(\theta_1 S_1, \ldots, \theta_M S_M) : \theta \in \Theta\}$, where $\theta = (\theta_1, \ldots, \theta_M)$. Intuitively, scalars $\theta_i$ can be interpreted as scaling factors.

**Definition 1:** The family of sets $S(S, \Theta)$ is practical Robust Positive Invariant (pRPI) for the constrained local error dynamics given by (26), (3) and (8), if, for all $i \in M$ and all $(\theta_1 S_1, \ldots, \theta_M S_M) \in S(S, \Theta)$, one has

$$\theta_i S_i \subseteq E_i$$

$$\bar{A}_{ii} \theta_i S_i \oplus \bigoplus_{j \in N_i} \bar{A}_{ij} \theta_j S_j \oplus D_i W_i \subseteq \theta^+_i S_i$$

$$(\theta^+_1 S_1, \ldots, \theta^+_M S_M) \in S(S, \Theta)$$

**Assumption 3:** The sets $S_i, i \in M$ are $C$-sets containing the origin in their interior.

The main issue we will address in the sequel is the following: given $S$, is there a nonempty set $\Theta \subseteq \mathbb{R}^M_+$ such that the family $S(S, \Theta)$ is pRPI? Under Assumption 3, an answer is provided in (Raković et al., 2011), where it is proposed to first derive the dynamics of the scaling factors $\theta_i$. More precisely, for all $i, j \in M$ we set

$$\mu_{ij} = \begin{cases} \min_{\mu \geq 0} \{ \mu : \bar{A}_{ij} S_j \subseteq \mu S_i \} & \text{if } i = j \text{ or } j \in N_i \\ 0 & \text{otherwise} \end{cases}$$

$$\alpha_i = \min_{\beta \geq 0} \{ \beta : D_i W_i \subseteq \beta S_i \}.$$
and define the collective dynamics of the scaling factors

\[ \theta^+ = T\theta + \alpha \]  

(13)

where the entries of \( T \in \mathbb{R}^{M \times M} \) are \( T_{ij} = \mu_{ij} \) and \( \alpha = (\alpha_1, \ldots, \alpha_M) \). It is easy to show that (13) guarantees

\[ e_{[i]} \in \theta_i S_i \Rightarrow e_{[i]}^+ \in \theta_i^+ S_i. \]  

(14)

For fulfilling (10a), let us define

\[ \Theta_0 = \{ \theta \in \mathbb{R}^M : \forall i \in \mathcal{M}, \ \theta_i S_i \subseteq E_i \} \]  

(15)

The key assumption used in (Raković et al., 2011) for providing a set \( \Theta \) that makes \( S(S, \Theta) \) a pRPI family is the following one.

**Assumption 4:**  
(i) \( T \) is Schur.  
(ii) The unique equilibrium point \( \bar{\theta} \) of system (13) is such that \( \bar{\theta} \in \Theta_0 \).  
(iii) The set \( \Theta \) is an invariant set for system (13) and constraint set \( \Theta_0 \), i.e. \( \forall \theta \in \Theta \subseteq \Theta_0 \), \( \theta^+ \in \Theta \).

Roughly speaking, Assumption 4 avoids scaling factors (and hence estimation errors) to become unbounded because of coupling between local estimators. Further comments on the meaning of Assumption 4 can be found in (Raković et al., 2011). Next, we provide the main result on invariance of estimation errors.

**Lemma 1** ((Raković et al., 2011)): *Let Assumptions 1-4 hold. Then,*

(i) there is a non-trivial convex and compact positively invariant set \( \Theta \) for system (13) equipped with constraints \( \theta \in \Theta_0 \);

(ii) \( S(S, \Theta) \) is pRPI for (26) with constraints (3) and (8). \( \square \)

Lemma 1 guarantees that

\[ \theta(0) \in \Theta \text{ and } e_{[i]}(0) \in \theta_i(0) S_i, \ \forall i \in \mathcal{M} \Rightarrow e_{[i]}(t) \in \theta_i(t) S_i, \ \forall i \in \mathcal{M}, \ \forall t \geq 0 \]  

(16)

Furthermore, as shown in (Raković et al., 2011), \( \text{dist}(e_{[i]}(t), \bar{\theta}_i S_i) \to 0 \) as \( t \to \infty \). In the nominal case, i.e. \( W = \{0\} \), one has \( \alpha = 0 \) in (13). Then \( \bar{\theta} = 0 \) and property (7) is guaranteed. Also (8) holds since, from (16) and (10a) one has \( e_{[i]}(t) \in \theta_i(t) S_i \subseteq E_i \). Therefore, Problem 1 is solved if we can design local state estimators fulfilling the assumptions of Lemma 1. A design procedure to achieve this goal is proposed in Section 4.

**Remark 1:** Note that, according to (16), the initialization of the local estimators requires to find a suitable initial state \( \theta(0) \in \Theta \) for system (13) and this is a centralized operation. In order to allow each estimator to locally compute its initial state, one can build offline an inner box approximation \( \bar{\Theta} = \prod_{i=1}^M [\theta_i, \bar{\theta}_i] \) contained in \( \Theta \) and choose \( \bar{x}_{[i]}(0) \) such that \( x_{[i]}(0) - \bar{x}_{[i]}(0) \in [\bar{\theta}_i, \theta_i] \).

4. Design of local estimators

In this section, we propose a method to design the distributed state estimator presented in Sections 2 and 3. The key issue is how to compute suitable gains \( L_{ij} \) and binary variables \( \delta_{ij} \) such that Assumption 4 holds. From now on we consider polytopic sets \( E_i, W_i \) and \( S_i, i \in \mathcal{M} \) verifying
Assumptions 2 and 3. Without loss of generality we can write

$$E_i = \{ h^{T}_{i,\tau} e_{[i]} \leq 1, \forall \tau \in 1 : \tau_{i} \} = \{ H_i e_{[i]} \leq 1_{\tau} \}$$  \hspace{1cm} (17a)

$$W_i = \{ f^{T}_{i,\upsilon} w_{[i]} \leq 1, \forall \upsilon \in 1 : \upsilon_{i} \} = \{ F_i w_{[i]} \leq 1_{\upsilon} \}$$  \hspace{1cm} (17b)

$$S_i = \{ g^{T}_{i,\psi} s_{[i]} \leq 1, \forall \psi \in 1 : \psi_{i} \} = \{ G_i s_{[i]} \leq 1_{\psi} \}$$  \hspace{1cm} (17c)

where $H_i = (h^{T}_{i,1}, \ldots, h^{T}_{i,\tau_{i}}) \in \mathbb{R}^{\tau_{i} \times n_i}$, $F_i = (f^{T}_{i,1}, \ldots, f^{T}_{i,\upsilon_{i}}) \in \mathbb{R}^{\upsilon_{i} \times r_i}$ and $G_i = (g^{T}_{i,1}, \ldots, g^{T}_{i,\psi_{i}}) \in \mathbb{R}^{\psi_{i} \times n_i}$. The design procedure is summarized in Algorithm 1 that is composed by three parts. Operations in part (A) can be executed in parallel using computational resources associated with

Algorithm 1

Input: polytopic sets $E_i$, $W_i$, $i \in \mathcal{M}$ verifying Assumption 2.
Output: A pRPI family of sets $S(S, \Theta)$.

(A) Decentralized steps. For all $i \in \mathcal{M}$,
(I) compute the matrix $L_{ii}$ such that $\bar{A}_{ii}$ is Schur and has as many zero eigenvalues as possible;
(II) compute a $\lambda_i$-contractive set $S_i$ for

$$e_{[i]}^+ = \bar{A}_{ii} e_{[i]}$$  \hspace{1cm} (18)

verifying $S_i \subseteq E_i$ and set $\mu_{ii} = \lambda_i$;
(III) compute $\alpha_i$ as in (12).

(B) Distributed steps. For all $i \in \mathcal{M}$,
(I) if $\delta_{ij} = 1$, compute the matrix $L_{ij}, \forall j \in \mathcal{N}_i$ solving

$$\min_{L_{ij}} ||G_i \bar{A}_{ij} G_j^0||_p$$  \hspace{1cm} (19)

where either $p = 1$ or $p = F$.
(II) compute $\mu_{ij}$ as in (11).

(C) Centralized steps
(I) if matrix $T$ is not Schur stop;
(II) compute set $\Theta_0$ as in (15) and the equilibrium point $\bar{\theta}$ of system (13). If $\bar{\theta} \notin \Theta_0$ stop;
(III) compute the maximal invariant set $\Theta_{\infty}$ of system (13) equipped with constraint $\Theta_0$;
(IV) compute an inner box approximation $\bar{\Theta}$ of $\Theta_{\infty}$.

subsystems, i.e. in a decentralized fashion. Steps in part (B) have a distributed nature, meaning that computations are decentralized but they can be performed only after each system has received suitable pieces of information from its parents. Finally, design steps in part (C) require centralized computations involving only the $M$-th order system (13). Next, we comment each step of Algorithm 1 in details.

4.1 Part (A)

Step (AI) is the easiest one and it can be performed only if pairs $(A_{ii}, C_i), i \in \mathcal{M}$ are detectable. The requirement of placing eigenvalues of $\bar{A}_{ii}$ in zero is motivated by step (AII).

The computation of sets $S_i$ as in step (AII) has been suggested in (Raković et al., 2011) and it is based on the argument that if sets $S_i$ are $\lambda_i$-contractive for some $\lambda_i \in (0, 1)$, then sets $(1 - \lambda_i)S_i$
can be used for compensating coupling terms in the error dynamics. Remarkably, using the efficient procedures proposed in (Raković & Baric, 2010), the computation of a set $S_i$ amounts to solving the optimization problem

$$
\mathcal{P}_i(S_i^0, k_i) : \min_{\gamma_i, \beta_i, \{S_s^i\}_{s=1}^{k_i-1}} \gamma_i
$$

$$
\gamma_i \in [0, 1), \quad S_i^{k_i} \subseteq \gamma_i S_i^0 \quad (20b)
$$

$$
\beta_i \in \mathbb{R}_+, \quad \bigoplus_{s=0}^{k_i-1} S_s^i \subseteq \beta_i E_i \quad (20c)
$$

$$
S_s^i = \bar{A}_{ii}^s S_i^0, \quad \forall s = 1, \ldots, k_i \quad (20d)
$$

where $k_i \in \mathbb{N}$ and the set $S_i^0 \subset \mathbb{R}^{n_i}$ are provided as inputs. In particular, $(20)$ is an LP problem and the set $S_i$ can be obtained as $S_i = \beta_i^{-1} \bigoplus_{s=0}^{k_i-1} S_s^i$. Furthermore, the contractivity parameter is $\lambda_i = \frac{\delta_i + \gamma_i^*}{\delta_i^*}$, where $\gamma_i^*$ is a solution to $(20)$ and $\delta_i = \min \{ \bar{\delta} : \bigoplus_{s=0}^{k_i-1} S_s^i \subseteq \bar{\delta} S_i^0, \bar{\delta} \geq 1 \}$. Note that also $\delta_i$ can be computed solving an LP problem. As shown in (Raković & Baric, 2010), since the matrix $\bar{A}_{ii}$ is Schur, then, given a $\mathcal{C}$-set $S_i^0$, there exists a sufficiently large $k_i$ such that problem $(20)$ is feasible. Moreover, if all eigenvalues of $\bar{A}_{ii}$ are zero, feasibility of $(20)$ can be guaranteed setting $k_i = n_i$. Indeed since $\bar{A}_{ii}^{n_i} = 0_{n_i \times n_i}$, we have $S_i^{n_i} = \{0\}$ and hence, irrespectively of $S_i^0$, constraints $(20b)$ hold with $\alpha_i = 0$. Moreover, since from $(20d)$ sets $\{S_s^i\}_{s=1}^{k_i}$ are polytopes containing the origin, then there exists $\beta_i$ such that constraints $(20c)$ hold. We highlight that the scalar $\mu_{ii}$ computed as in $(11)$ is equal to the contractivity parameter $\lambda_i$.

Step (AIII) focuses on the computation of scalars $\alpha_i$. From $(12)$ and $(17b)$, using procedures proposed in (Kolmanovsky & Gilbert, 1998), we have $\alpha_i = \max_{\psi \in \psi_i \{z_i\}} \{z_i\}$ where

$$
z_i = \max_{w_i} g_i, \psi D_i w_i \quad (21)
$$

$$
\mathcal{F}_i w_i \leq 1_{\psi_i}
$$

Therefore, step (AIII) requires the solution to the $\psi_i$ LP problems $(21)$.

4.2 Part (B)

For the computation of matrices $L_{ij}$ and parameters $\mu_{ij}$, each system $\Sigma_{[i]}$ needs to receive the matrix $C_j$ and the set $S_j$ from parents $j \in \mathcal{N}_i$ such that $\delta_{ij} = 1$.

In step (BI), if $\delta_{ij} = 1$, the computation of matrices $L_{ij}$, $j \in \mathcal{N}_i$ is required. Since the choice of $L_{ij}$ affects the coupling term $A_{ij}$ and hence the Schurness of matrix T, we propose to reduce the magnitude of coupling by minimizing the magnitude of $A_{ij}$ in $(19)$, where $G_i$ and $G_j^0$ allow us to take into account the size of sets $S_i$ and $S_j$, respectively. Indeed, the term $\|G_i A_{ij} G_j^0\|_p$ is a measure of how much the coupling term $A_{ij}^s s_{[j]}$, $j \in \mathcal{N}_i$, affects the fulfillment of the constraint $s_{[i]} \in S_i$. More precisely, proceeding as in Appendix 8.7.3 of (Rivero, 2014), one can show that the greater $\|G_i A_{ij} G_j^0\|_p$, the more difficult is to guarantee $e_{[i]} \in E_i$.

We highlight that the minimization of $\|G_i A_{ij} G_j^0\|_1$ in $(19)$ amounts to an LP problem and the minimization of $\|G_i A_{ij} G_j^0\|_F$ can be recast into a Quadratic Programming (QP) problem. Hence the optimization problem $(19)$ can be solved in polynomial time. In particular, there is always a solution to the optimization problem $(19)$ because $L_{ij} = 0$ is always feasible. So far the parameters $\delta_{ij}$ have been considered fixed. However, if in step (BI) one obtains $L_{ij} = 0$ for some $j \in \mathcal{N}_i$, it is impossible to reduce the magnitude of the coupling term $A_{ij}$ and, from $(4)$, the knowledge of $y_{[j]}$ is useless. This suggests to revise the choice of $\delta_{ij}$ and set $\delta_{ij} = 0$. In step (BII), since $S_i$ are
polytopes, using procedures proposed in (Kolmanovsky & Gilbert, 1998) we can compute scalars \( \mu_{ij} \) as

\[
\mu_{ij} = \max_{\psi \in [1: \bar{\psi}_i]} \{ \max_{s[i]} g_i, \bar{\psi}_i, A_{ij} s[j] : G_j s[j] \leq 1 \psi_i \}.
\]

that requires the solution of \( \bar{\psi}_i \) LP problems.

### 4.3 Part (C)

In step (CI) we check the Schurness of matrix \( T \). If the test fails, Assumption 4-(i) cannot be fulfilled and the only possibility is to restart the algorithm after increasing the number of variables \( \delta_{ij} \) that are equal to one.

In step (CII), since the sets \( S_i \) and \( E_i \) are polytopes, using results from (Kolmanovsky & Gilbert, 1998) the computation of the set \( \Theta_0 \) can be done as follows

\[
\Theta_0 = \prod_{i=1}^{M} [0, \bar{\theta}_i]
\]

\[
\bar{\theta}_i = \left( \max_{\tilde{\theta}_i \in \Theta_0} \left\{ \sup_{\tau \in [1: \bar{\tau}_i]} s[i] : G_i s[i] \leq 1 \psi_i \right\} \right)^{-1}.
\]

Moreover, in step (CII) we compute the equilibrium point \( \bar{\theta} \) of system (13). If \( \bar{\theta} \notin \Theta_0 \) we cannot guarantee property (8) and therefore the algorithm stops. Note that if \( W_i = \{0\} \), \( \forall i \in M \), the equilibrium point \( \bar{\theta} \) is the origin and hence \( \bar{\theta} \in \Theta_0 \) by construction.

According to Assumption 4-iii , the set \( \Theta \) of all feasible contractions \( \theta \) is computed as an RPI set for system (13) and constraints \( \theta \in \Theta_0 \). In particular, since \( T \) is Schur and \( \Theta_0 \) is a polytope, using results from (Gilbert & Tan, 1991) we can compute the maximal RPI set \( \Theta_\infty \) by solving a suitable LP problem.

As discussed in Remark 1, a decentralized initialization of state estimators is possible computing an hyperrectangle \( \bar{\Theta} \) contained in \( \Theta_\infty \). This is done in step (CIV). More precisely, using results from (Bemporad, Filippi, & Torrisi, 2004), we can set \( \bar{\Theta} = \prod_{i=1}^{M} [0, \bar{\theta}_i] \) where

\[
\bar{\theta}_i = \max_{\tilde{\theta} \in \Theta_\infty} \gamma^T \tilde{\theta},
\]

\[
\gamma = (\gamma_1, \ldots, \gamma_M)
\]

\[
\gamma_i = (\max_{\theta} \theta_i : \theta \in \Theta_\infty)^{-1}.
\]

As described in (Bemporad et al., 2004), the vector \( \gamma \) is used for maximizing the volume of \( \bar{\Theta} \). From (24) and (25) the computation of the hyper-rectangle \( \bar{\Theta} \) requires the solution of \( M + 1 \) LP optimization problems.

**Remark 2:** We can account for subsystems coupled through inputs and disturbances by replacing the dynamics (2) with

\[
x^+[i] = A_{ii} x[i] + B_i u[i] + \sum_{j \in N_i} (A_{ij} x[j] + B_{ij} u[j] + D_{ij} w[j]) + D_{ii} w[i]
\]

\[
y[i] = C_i x[i]
\]

where \( D_{ii} = D_i \). The overall system is still described by (1) but matrices \( B \) and \( D \) are no longer block-diagonal. For dealing with input coupling, it is sufficient to replace the local state estimator...
In (10) and (12), the role played by the set $D$ in (4) with

$$
\tilde{x}^+_i = A_i \tilde{x}_i + B_i u_i - L_i (y_i - C_i \tilde{x}_i) + \sum_{j \in N_i} (A_{ij} \tilde{x}_j + B_{ij} u_j) - \sum_{j \in N_i} \delta_{ij} L_{ij} (y_j - C_j \tilde{x}_j).
$$

Note that we now require that each estimator $\tilde{\Sigma}_{[j]}$ receives the inputs $u_{[j]}$ from parent subsystems $\Sigma_{[j], j \in N_i}$. The local error dynamics (26) becomes

$$
e^+_i = \bar{A}_i e_i + \sum_{j \in N_i} \bar{A}_{ij} e_{[j]} + D_{ii} w_i + \sum_{j \in N_i} D_{ij} w_{[j]} \quad (26)
$$

In (10) and (12), the role played by the set $D_i W_i$ is now played by the larger set $D_i W_i \bigoplus_{j \in N_i} D_{ij} W_j$. As for Algorithm 1, no change is needed. Just note that for computing scalars $\alpha_i$ through LPs in step (AIII), the optimization problem (21) must be replaced with

$$
z_i = \max_{w_i} g_{i, \psi} \left( D_{ii} w_i + \sum_{j \in N_i} D_{ij} w_{[j]} \right) \quad (27)
$$

$$
\mathcal{F}_k w_k \leq 1_{\hat{\nu}_k}, \quad k \in \{i\} \cup N_i
$$

5. Large-scale systems with variable number of subsystems

In this section, we discuss the retuning of the DSE when a subsystem is added or removed. We highlight that plug-in and unplugging of subsystems are here considered as offline operations. In particular, we will show how to preserve properties (7) and (8) without performing all computations required by Algorithm 1. As a starting point, we consider system (1) equipped with a DSE designed using Algorithm 1.

5.1 Plug-in operation

Assume the new subsystem $\Sigma_{[M+1]}$ is plugged in and set $\tilde{M} = M \cup \{M+1\}$. Since the overall system has changed, in principle one has to design the DSE from scratch running Algorithm 1. Note however that Part (A) of Algorithm 1 is decentralized and therefore it has to be executed for the new subsystem only. Part (B) of Algorithm 1 involves only the new subsystem, its parents and its children $C_{M+1} = \{j \in M : A_{M+1,j} \neq 0, j \neq M + 1\}$. In fact, subsystem $\Sigma_{[M+1]}$ needs sets $S_j$ from its parents for computing parameters $\mu_{M+1,j}, j \in N_{M+1}$. Moreover since children of $\Sigma_{[M+1]}$ have a new parent, they need to know $S_{M+1}$ in order to update parameters $\mu_{k,M+1}, k \in C_{M+1}$.

If Step (CI) or Step (CII) fail, we declare that system $\Sigma_{[M+1]}$ can not be added, because the family of sets $S(S, \Theta)$ is not a pRPI. In Algorithm 2 we summarize the computations for updating the DSE that are triggered by the addition of $\Sigma_{[M+1]}$. We highlight that Algorithm 2 can be easily generalized to the case of subsystems coupled through inputs and disturbances (see Remark 2). The only change is to assume that sets $W_{ij}, j \in N_{M+1}$ are passed as additional input arguments, since they are needed in steps (B) and (C).

5.2 Unplugging operation

Assume subsystem $\Sigma_{[q]}, q \in M$ is removed. We will show that no update of the DSE is required in order to guarantee (7) and (8). In the following, vectors, matrices and sets with a hat are quantities
Algorithm 2

**Input:** new subsystem $\Sigma_{[M+1]}$ with sets $E_{M+1}$ and $W_{M+1}$.

**Output:** an updated pRPI family of sets $\mathcal{S}(S,\Theta)$.

(A) **Decentralized steps**

For $i = M+1$ execute Steps (AI)-(AIII) of Algorithm 1;

(B) **Distributed steps**

- For subsystem $\Sigma_{[M+1]}$, if $\delta_{M+1,j} = 1$, compute the matrix $L_{M+1,j}$, $\forall j \in N_{M+1}$ solving $\min_{L_{M+1,j}} \|G_{M+1}A_{M+1,j}G_j\|_p$, $p = 1$ or $p = F$, and then compute $\mu_{M+1,j}$;
- For subsystems $\Sigma_{[k]}$, if $\delta_{k,M+1} = 1$, compute the matrix $L_{k,M+1}$, $\forall k \in C_{M+1}$ solving $\min_{L_{k,M+1}} \|G_kA_{k,M+1}G_{M+1}\|_p$, $p = 1$ or $p = F$, and then compute $\mu_{k,M+1}$;

(C) **Centralized steps**

Execute steps (CI)-(CIV) of Algorithm 1.

The matrix $\hat{\Theta}$ of system after subsystem $q$ has been removed. As an example, the matrix

$$
\hat{T} = \begin{bmatrix}
\mu_{11} & \cdots & \mu_{1,q-1} & \mu_{1,q+1} & \cdots & \mu_{1,M} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mu_{q-1,1} & \cdots & \mu_{q-1,q-1} & \mu_{q-1,q+1} & \cdots & \mu_{q-1,M} \\
\mu_{q+1,1} & \cdots & \mu_{q+1,q-1} & \mu_{q+1,q+1} & \cdots & \mu_{q+1,M} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mu_{M,1} & \cdots & \mu_{M,q-1} & \mu_{M,q+1} & \cdots & \mu_{M,M}
\end{bmatrix} \in \mathbb{R}^{M-1 \times M-1}
$$

is obtained from matrix $T$, by eliminating the $q$-th row and column. Next, we show Assumptions 4-(i), 4-(ii) and 4-(iii) are still verified after the removal of $\Sigma_{[q]}$.

Proposition 1: *If the matrix $T \in \mathbb{R}^{M \times M}$ in (13) is Schur, then also the matrix $\hat{T}$ is Schur.*

The proof of Proposition 1 (as well as the proofs of Propositions 2 and 3 below) can be found in Appendix.

The next result guarantees Assumption 4-(ii) still holds after the removal of subsystem $q$.

Proposition 2: For $q \in M$, let $\hat{\theta} = (\theta_1, \ldots, \theta_{q-1}, \theta_{q+1}, \ldots, \theta_{M-1})$, $\hat{\alpha} = (\alpha_1, \ldots, \alpha_{q-1}, \alpha_{q+1}, \ldots, \alpha_{M-1})$ and

$$
\hat{\Theta}_0 = \{\xi \in \mathbb{R}^{M-1} : (\xi_1, \ldots, \xi_{q-1}, 0, \xi_{q}, \ldots, \xi_{M-1}) \in \Theta_0\}
$$

(28)

If Assumption 4-(ii) holds, the unique equilibrium $\hat{\theta}$ of system

$$
\hat{\theta}^+ = \hat{T}\hat{\theta} + \hat{\alpha}
$$

(29)

is such that $\hat{\theta} \in \hat{\Theta}_0$. 

11
Finally, the following proposition concerns Assumption 4-(iii).

**Proposition 3:** For \( q \in \mathcal{M} \), the set
\[
\hat{\Theta} = \{ \hat{\theta} \in \mathbb{R}^{M-1} : (\hat{\theta}_1, \ldots, \hat{\theta}_{q-1}, 0, \hat{\theta}_q, \ldots, \hat{\theta}_{M-1}) \in \Theta_\infty \}
\]  

is an RPI set for system (29).

From Proposition 3 we have that the projection of set \( \Theta \) on the coordinates \( \hat{\theta} \) is still an RPI set for (1) after the removal of subsystem \( q \). We also note that the set \( \hat{\Theta} \) is not the maximal RPI, i.e. with a new execution of Step (CIII) of Algorithm 1 we could obtain an RPI set \( \hat{\Theta}_\infty \) verifying \( \hat{\Theta} \subseteq \hat{\Theta}_\infty \). We also note that the projection \( \hat{\Theta} \) of \( \Theta \) on the coordinates \( \hat{\theta} \) is a box verifying \( \hat{\Theta} \subseteq \Theta \). However, with a new execution of Step (CIV) of Algorithm 1 we could obtain a bigger inner box approximation.

**Remark 3:** Our approach to the design of DSEs has several features in common with the method proposed in (Riverso, Farina, et al., 2013). In both cases, a bounded-error setting is assumed and the key idea is to exploit robust invariance for confining the error in sets which are contained, by construction, in the set \( \Theta \). Moreover, both design procedures, if successful, guarantee convergence of errors to zero in absence of disturbances. There is however a major difference between the two design strategies. In (Riverso, Farina, et al., 2013) the synthesis algorithm is parent-based, meaning that the design of a local state estimator for subsystem \( \Sigma_{[i]} \) requires, at most, pieces of information from subsystems \( \Sigma_{[j]}, j \in \mathcal{N}_i \). In other words, centralized computations, as those required in step (C) of Algorithm 1, are completely avoided. This implies, the synthesis method in (Riverso, Farina, et al., 2013) is fully scalable as the number of subsystems increases. However, this feature comes at the price of requiring the errors are kept in invariant sets that cannot change over time. Instead, in the present paper, we relax this condition since errors \( e_{[i]} \) are confined in sets \( \theta_i S_i \) that can grow and shrink due to the dynamics of scaling factors.

DSEs have been also proposed in the context of fault detection, assuming nonlinear dynamics and partition-based (Zhang, 2010) or even more general (Boem et al., 2011) structures of the estimation scheme. However, in these papers the focus is on providing fault-detectability conditions based on suitably defined time-varying thresholds. Hence, the goal is to derive *a posteriori* bounds on the errors magnitude rather than guaranteeing the fulfillment of prescribed (and polytopic) constraints on the errors. In addition, it is assumed that disturbances are bounded in time-varying boxes, while, in our setting, disturbances are constrained in more general (albeit time-invariant) polytopic sets. We highlight that we exploit the precise knowledge of these sets for reducing conservativity of the design procedure as much as possible.

6. **Examples**

6.1 **String of masses**

In this example, we consider a large-scale system composed of \( M \) masses connected as in Figure 1. Our aim is to show how to design a DSE using Algorithm 1 and how to perform the plug-in of a new mass using Algorithm 2. Each mass \( i \in \mathcal{M} = 1 : M \), is a subsystem with input \( u_{[i]} \) and state variables \( x_{[i]} = (x_{[i,1]}, x_{[i,2]}) \), where \( x_{[i,1]} \) is the displacement with respect to a given equilibrium position, \( x_{[i,2]} \) is the horizontal velocity and \( 100u_{[i]} \) is an external force in the horizontal direction. The only output is \( y_{[i]} = x_{[i,1]} \). Moreover each state is affected by a local disturbance \( w_{[i]} = (w_{[i,1]}, w_{[i,2]}) \). The values of masses \( m_i \) have been randomly extracted in the interval [8, 12] while spring and damping coefficients are identical and equal to 0.5. Subsystems \( \Sigma_{[i]} \) have been obtained by discretizing continuous-time models with 0.2 sec sampling time, using zero-order-hold discretization for the local dynamics and treating \( x_{[j]}, j \in \mathcal{N}_i \) as exogenous signals. We consider
the following bounds on the state estimation error

\[ E_i = \{ e_{[i]} \in \mathbb{R}^2 : |e_{[i,1]}| \leq 3, |e_{[i,2]}| \leq 2 \} \]

and disturbance bounds \( W_i = 0.01E_i \).

We used a number of masses \( M \) between 2 and 100 and checked that the design of a DSE can be always performed running Algorithm 1. We highlight that increasing the number of masses, only the computational complexity of centralized steps in Algorithm 1 increases. In Figure 2 we plot sets \( E_5 \) and \( S_5 \) obtained designing the DSE for an array of 100 masses. In Figure 3, for \( M = 100 \)

we show the state estimation errors for all subsystems. We highlight that thanks to the execution of Step (AI) in Algorithm 1, we are able to reduce the error in few time instants. However, due to the presence of bounded disturbances, we can only guarantee boundedness of the state estimation errors \( e_{[i]}(t) \), for \( t \to \infty \). In Figure 4, for \( t \geq 0.6 \) we show the maximum state estimation error defined as

\[ \bar{e}_{[j]}(t) = \max_{i \in M} |x_{[i,j]}(t) - \tilde{x}_{[i,j]}(t)| \] (31)

where \( x_{[i,j]} \) and \( \tilde{x}_{[i,j]} \) are, respectively, the real and estimated state trajectory of the \( j \)-th state of the \( i \)-th subsystem.

Next we discuss the plug-in of a new mass. We assume the new mass is appended to the string and hence physically connected with mass \( M \) only. Using the same parameters described above, the plug-in operation was always allowed for different mass values. However, as the value of the new mass decreases (e.g. \( m_{M+1} \in (0, 5) \)), the coupling term \( A_{M+1} \) increases and the plug-in operation can be denied. In this case, however, we are still able to design the DSE by setting \( \delta_{M+1} = 1 \) and hence by receiving the output from mass \( M \). Indeed, the optimization problem (19) gives a gain \( L_{M+1} \) yielding \( A_{M+1} = 0 \). This example shows how to use the magnitude of the coupling terms for guiding the choice of parameters \( \delta_{ij} \).
6.2 Power network system

In this section, we apply the proposed distributed state estimator to a power network system composed by several power generation areas coupled through tie-lines. The dynamics of an area equipped with primary control and linearized around equilibrium value for all variables can be described by the following continuous-time LTI model (Saadat, 2002)

\[
\Sigma^{C}_{[i]} : \quad \dot{x}_{[i]} = A_{ii} x_{[i]} + \bar{B}_i \bar{u}_{[i]} + \sum_{j \in \mathcal{N}_i} A_{ij} x_{[j]} + w_{[i]} \quad (32)
\]

where \( x_{[i]} = (\Delta \theta_i, \Delta \omega_i, \Delta P_{m_i}, \Delta P_{v_i}) \) is the state, \( \bar{u}_{[i]} = (\Delta P_{r_{ef,i}}, \Delta P_{r_{fu}}) \) is composed by the control input of each area and the local power load and \( \mathcal{N}_i \) is the sets of parent areas, i.e. areas directly connected to \( \Sigma^{C}_{[i]} \) through tie-lines. In (32), \( w_{[i]} \in \mathbb{R}^{n_i} \) is the disturbance term for the \( i \)-th area and
it is bounded in the polytopic set $W_i \subset \mathbb{R}^{n_i}$. The matrices of system (32) are defined as

$$A_{ii}(\{P_{ij}\}_{j \in \mathcal{N}_i}) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\sum_{j \in \mathcal{N}_i} P_{ij} & -\frac{D_i}{2H_i} & \frac{1}{2H_i} & 0 \\ 0 & 0 & -\frac{1}{R_i} & 0 \\ -\frac{1}{T_{gi}} & \frac{1}{T_{gi}} & 0 & 0 \end{bmatrix}.$$ 

$$\bar{B}_i = \begin{bmatrix} 0 & 0 \\ 0 & -\frac{1}{2H_i} \\ 0 & 0 \\ \frac{1}{T_{si}} & 0 \end{bmatrix}.$$ 

The state estimation error

$$E_i = \{e_{[i]} \in \mathbb{R}^{n_i} : ||e_{[i,1]}||_{\infty} \leq 0.005, ||e_{[i,k]}||_{\infty} \leq 0.01, k \in 2:4\}. \quad (33)$$

For parameter values and the meaning of constants and variables, we defer the reader to Section 1 of (Riverso & Ferrari-Trecate, 2012a). We obtain models $\Sigma_{[i]}$ by discretizing models $\Sigma_{[i]}^C$ with 1 sec sampling time, using exact discretization and treating $\bar{u}_{[i]}, x_{[j]}, j \in \mathcal{N}_i$ and $w_{[i]}$ as exogenous signals. We note that using the proposed discretization scheme the set of parents $\mathcal{N}_i$ does not change. In the following we propose different distributed state estimators for a power network composed by four areas as in Figure 5. In Example 1 and 2, for each area, we consider the following bounds on the state estimation error

$$E_{\delta_i} = \{e_{\delta_i} \in \mathbb{R}^{n_i} : ||e_{\delta_i,1}||_{\infty} \leq 0.005, ||e_{\delta_i,k}||_{\infty} \leq 0.01, k \in 2:4\}.$$ 

We highlight that constraints (33) correspond to tolerate state estimation errors less then 10% of the maximum value assumed by the state variables. In Example 3, we consider constraints on the error equal to $2E_i$, $\forall i \in \mathcal{M}$.

### 6.2.1 Example 1

As first example, we consider $\delta_{ij} = 1$, $\forall j \in \mathcal{N}_i$, $W_i = \{0\}$, $\forall i \in \mathcal{M}$ (i.e. no disturbances act on the system) and assume to measure only the angular speed deviation $\Delta \omega_{[i]}$ of each area. Therefore, outputs of subsystem $i$ are given by

$$y_{[i]} = C_i x_{[i]}, \quad C_i = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}.$$ 

In this case, Algorithm 1 stops in Step (CI) because the computed sets $S_i$ are such that $T$ is not Schur. We highlight that from the results of Step (BI), one obtains the same results if parameters

---

1For the simulations, we use the load power steps given in Section 1.1 of (Riverso & Ferrari-Trecate, 2012a) and the control inputs computed using MPC controllers as in Section 2 of (Riverso & Ferrari-Trecate, 2012a).
6.2.2 Example 2

We consider $W_i = \{0\}$, $\forall i \in \mathcal{M}$, i.e. no disturbances act on the system, and we assume to measure both $\Delta \theta_i$ and $\Delta \omega_i$ of each area. Therefore the outputs are given by

$$y_i = C_i x_i, \quad C_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (35)$$

First we consider $\delta_{ij} = 0$, $\forall i \in \mathcal{M}$, $\forall j \in \mathcal{N}_i$. In this case, as in the first example, since we cannot take advantage of the knowledge of parents’ outputs, Algorithm 1 stops before its conclusion. Indeed, it is impossible to find sets $S_i$ such that $T$ is Schur. This example shows that if we also consider more output variables for each subsystem, Algorithm 1 can stop in Step (CI) due to the magnitude of the coupling terms $A_{ij}$. Now we consider $\delta_{ij} = 1$, $\forall i \in \mathcal{M}$, $\forall j \in \mathcal{N}_i$. In this case we can reduce the magnitude of the coupling terms. Solving optimization problems (19), we can compute matrices $L_{ij}$ such that $\bar{A}_{ij} = 0_{n_i \times n_j}$, hence the Schurness of matrix $T$ is guaranteed since sets $S_i$ are $\lambda_i$-contractive. In this case, $T = \text{diag}(0.923, 0.796, 0.575, 0.874)$ and $\bar{\Theta} = \prod_{i=1}^{4}[0,1]$. We note that if matrix $T$ is diagonal, Step (CIV) of Algorithm 1 can be skipped since $\Theta_\infty = \Theta$.

We performed an estimation experiment initializing the local state estimators $\tilde{\Sigma}_i$, $i \in \mathcal{M}$ with $\tilde{x}_i(0) = x_i(0) - e_i(0)$, where $e_i(0)$ is a vertex of the set $S_i$. In Figure 6 we show the maximum state estimation error defined as in (31). From Figure 6 we note that, since no disturbances act on the system, the state estimation error $\tilde{e}_i(t)$ converges to zero as $t \to \infty$, i.e. (7) is verified.

Next, we describe a plug-in operation. Assume one wants to connect a new area 5 to areas $j \in \mathcal{N}_5 = \{2,4\}$. Parameter values for area 5 can be found in (Riverso & Ferrari-Trecate, 2012a). Setting $\delta_{5j} = 1$ and $\delta_{j5} = 1$, in the distributed steps of Algorithm 2 we can compute gains $L_{5j}$ and $L_{j5}$ zeroing the terms $A_{5j}$ and $A_{j5}$ in the error dynamics. Hence, the new matrix $T$ is still diagonal and Schur (because $S_5$ is $\lambda_i$-contractive, with $\lambda_5 = 0.951$). Hence, the set $\bar{\Theta}$ is the hypercube $\prod_{i=1}^{5}[0,1]$. This example highlights that the plug-in of new generation areas is never denied, independently of the interconnection topology. Furthermore, there is no need to execute centralized steps in Algorithm 2 as their results can be anticipated. For the power network with 5 areas we performed an experiment initializing the local state estimators $\tilde{\Sigma}_i$, $i \in \mathcal{M}$ with $\tilde{x}_i(0) = x_i(0) - e_i(0)$, where $e_i(0)$ is a vertex of the set $S_i$. In Figure 7 we show the maximum state estimation error defined as in (31).
6.2.3 Example 3

We consider the power network in Figure 5 with output variables given in (35) and sets \( W_i = \{ w_{i[k]} \in \mathbb{R}^{n_i} : ||w_{i[k]}||_\infty \leq 10^{-5}, \ k = 1 : 4 \}, \ i \in \mathcal{M} = 1 : 4 \). As in Example 2, by considering \( \delta_{ij} = 1, \ \forall i \in \mathcal{M}, \ \forall j \in \mathcal{N}_i \), Algorithm 1 does not stop at any intermediate step. We have performed a similar experiment as in Example 2, but generating statistically independent random samples \( w_{i[t]}(t) \) from the uniform distribution on \( W_i \). In Figure 8 and 9, we show the maximum estimation error at the beginning of the experiment (Figure 8) and for \( t \geq 10 \) (Figure 9). In particular, even in presence of disturbances on the system the state estimation error \( e_{i[t]}(t) \) lies in the set \( E_i, \ \forall i \in \mathcal{M} \).
6.3 Comparison with different approaches

In the previous examples we considered polytopic sets $E_i$, $i \in \mathcal{M}$ defined in (33) that are also zonotopes, i.e. centrally symmetric polytopes. This is required by the DSE in (Farina & Scattolini, 2011). Local state estimators in (Farina & Scattolini, 2011) depend on the state of parent systems, but not on their outputs. This corresponds to setting $\delta_{ij} = 0$, $i,j \in \mathcal{M}$ in our scheme. Using the DSE in (Farina & Scattolini, 2011), we cannot compute the observers for the power network system in Examples 1 and 2. In fact, since all parameters $\delta_{ij}$ are zero it is impossible to reduce the magnitude of the coupling by using parents’ outputs, as we do in Example 2.

Moreover in (Farina & Scattolini, 2011), the authors look for a family of sets $S$ composed by mRPI sets $S_i$ verifying

$$\bar{A}_{ii} S_i \oplus \bigoplus_{j \in N_i} A_{ij} S_j \subseteq S_i \subseteq E_i. \quad (36)$$

We note that (36) is a special case of the parameterized RPI family in (10) when $\delta_{ij} = 0$ (i.e. $A_{ij} = A_{ji}$) and $W_i = \{0\}$. Since from (36) one has $\theta^+ = \theta$, the matrix $T$ is not Schur, and for guaranteeing convergence of the estimates one has to check Schurness of the overall matrix $A + LC$. In our distributed estimator, convergence of the estimates can be checked by testing the Schurness of $T$, i.e. no assumptions on the overall matrices are needed.

7. Conclusions

In this paper, we proposed a novel partition-based DSE for linear discrete-time subsystems affected by bounded disturbances. Our method guarantees convergence of the state estimates and the fulfillment of given bounds on state estimation errors. Similarly to (Farina & Scattolini, 2011) our DSE can be directly used together with state-feedback distributed control schemes such as (Riverso & Ferrari-Trecate, 2012b) although further research is needed for assessing the stability properties of the closed-loop system. A performance analysis of our DSE in presence of stochastic disturbances will be also a topic for future research.

Appendix A. Proof of Proposition 1

The proof of Proposition 1 hinges on Perron-Frobenius theory for nonnegative matrices. Next, we provide relevant definition, deferring the reader to (Meyer, 2000) for further details.

**Definition 2:** The graph, $\Gamma(Q) = (V, E)$ of $Q \in \mathbb{R}^{M \times M}$ is the directed graph with nodes $V = 1 : M$ and edges $E = \{(i,j) : q_{ij} \neq 0\}$ where $q_{ij}$ is the $ij$-th element of the matrix $Q$.

**Definition 3:** A directed graph $\Gamma$ is strongly connected if for any pair of nodes $(N_i, N_j)$ there exists a sequence of edges which leads from $N_i$ to $N_j$.

**Definition 4:** A matrix $Q \in \mathbb{R}^{M \times M}$ is irreducible if there is no permutation matrix $P$ such that

$$Z = P Q P^T = \begin{bmatrix} Q_{11} & Q_{12} \\ 0 & Q_{22} \end{bmatrix},$$

where $Q_{11} \in \mathbb{R}^{q \times q}$, $Q_{22} \in \mathbb{R}^{M-q, M-q}$ and $Q_{12} \in \mathbb{R}^{q, M-r}$, $0 < q < M$.

**Proof of Proposition 1.** The matrix $T$ in (13) is nonnegative, i.e. $\mu_{ij} \geq 0$, $\forall i,j \in \mathcal{M}$. Moreover,
Therefore, we can conclude that the projection of set $\Theta_\infty$ is strongly connected, $T$ is irreducible (Meyer, 2000, p. 671). Let

$$\mathcal{T} = \begin{bmatrix}
\mu_{11} & \cdots & \mu_{1,q-1} & 0 & \mu_{1,q+1} & \cdots & \mu_{1,M} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mu_{q-1,1} & \cdots & \mu_{q-1,q-1} & 0 & \mu_{q-1,q+1} & \cdots & \mu_{q-1,M} \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\mu_{q+1,1} & \cdots & \mu_{q+1,q-1} & 0 & \mu_{q+1,q+1} & \cdots & \mu_{q+1,M} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mu_{M,1} & \cdots & \mu_{M,q-1} & 0 & \mu_{M,q+1} & \cdots & \mu_{M,M} 
\end{bmatrix} \in \mathbb{R}^{M \times M} \quad (A1)$$

From Weilandt’s Theorem (Meyer, 2000, p. 675), one has $\rho(T) \leq \rho(T)$. Moreover, up to a permutation matrix, one has $\mathcal{T} = \begin{bmatrix} 0 & 0 \\ 0 & \tilde{T} \end{bmatrix}$ and hence $\rho(T) \leq \rho(T)$. Therefore $\rho(T) \leq \rho(T)$ and the proof is concluded recalling that, by assumption, $\rho(T) < 1$. \hfill \Box

Appendix B. Proof of Proposition 2

First we define matrix $\mathcal{T}$ as in (A1). Since $\mu_{ij} \geq 0$, $\forall i, j \in \mathcal{M}$, the elements of matrices $\mathcal{T}_k^k$ and $T_k^k$ are nonnegative $\forall k \geq 0$. Moreover we can show that the $ij$-th element of $\mathcal{T}_k^k$ (with abuse of notation, $\mathcal{T}_{ij}^k$) is smaller than the $ij$-th element of $T_k^k$ (with abuse of notation, $T_{ij}^k$), i.e. $\mathcal{T}_{ij}^k \leq T_{ij}^k$, $\forall i, j \in \mathcal{M}$ and $\forall k \geq 0$. Let $\hat{\tau} = (\hat{\theta}_1, \ldots, \hat{\theta}_{q-1}, 0, \hat{\theta}_q, \ldots, \hat{\theta}_{M-1}) \in \mathbb{R}^M$, where $\hat{\theta}_i$ is the $i$-th component of vector $\hat{\theta}$ and

$$\hat{\alpha} = \alpha \in \mathbb{R}_+^M, \hat{\alpha}_q = 0. \quad (B1)$$

The unique equilibrium point of system (13) can be written as $\hat{\theta} = \sum_{k=0}^\infty T_k^k \hat{\alpha}$. Moreover from (29) and the definitions of $\hat{\tau}$ and $\mathcal{T}$, we have that $\hat{\tau} = \sum_{k=0}^\infty \mathcal{T}_k^k \hat{\alpha}$. Since $\mathcal{T}_{ij}^k \leq T_{ij}^k$, then $\sum_{k=0}^\infty \mathcal{T}_{ij}^k \leq \sum_{k=0}^\infty T_{ij}^k$ and hence $\hat{\tau} \leq \hat{\theta}$ element-wise. From (23) and Assumption 4-ii, one has $\hat{\tau} \in \Theta_0$, Therefore, we can conclude that $\hat{\theta} \in \hat{\Theta}_0$. \hfill \Box

Appendix C. Proof of Proposition 3

Proof of Proposition 3. After subsystem $q$ has been removed, the dynamics of contraction factors $\hat{\theta}$ is given by (29). In the following we show that $\hat{\Theta}$ defined in (30) is an RPI set for (29).

From the invariance of set $\Theta_\infty$ we have that $T \theta + \alpha \in \Theta_\infty$, $\forall \theta \in \Theta_\infty$. Moreover, since $0 \in \Theta_\infty$, we have $T \theta + \hat{\alpha} \in \Theta_\infty$, $\forall \theta \in \Theta_\infty$ (where $T$ and $\hat{\alpha}$ are defined in (A1) and (B1)) i.e. the set $\Theta_\infty$ is also invariant for the LTI system $\theta^+ = T \theta + \hat{\alpha}$ and the $q$-th component of $\theta$ is always zero. Therefore, we can conclude that the projection of set $\Theta_\infty$ defined in (30) is an RPI set for system $\hat{\theta}^+ = \hat{T} \hat{\theta} + \hat{\alpha}$. \hfill \Box

References


