

# Distributed Finite-Element Kalman Filter for Field Estimation

Giorgio Battistelli, Luigi Chisci, *Senior Member, IEEE*, Nicola Forti, Giuseppe Pelosi, *Fellow, IEEE*, and Stefano Selleri, *Senior Member, IEEE*

**Abstract**—The paper deals with decentralized state estimation for spatially distributed systems described by linear partial differential equations from discrete in-space-and-time noisy measurements provided by sensors deployed over the spatial domain of interest. A fully scalable approach is pursued by decomposing the domain into possibly overlapping subdomains assigned to different processing nodes interconnected to form a network. Each node runs a local finite-dimensional discrete-time Kalman filter which exploits the finite element approach for spatial discretization, a backward Euler method for time-discretization and the parallel Schwarz method to iteratively enforce continuity of the field predictions over the boundaries of adjacent subdomains. Numerical stability of the adopted approximation scheme and stability of the proposed distributed finite element Kalman filter are mathematically proved. The effectiveness of the proposed approach is then demonstrated via simulation experiments concerning the estimation of a bi-dimensional temperature field.

**Index Terms**—Distributed-parameter systems, finite element method, Kalman filtering, networked state estimation.

## I. INTRODUCTION

THE RECENT breakthrough of wireless sensor network technology has made possible to cost-effectively monitor spatially distributed systems via deployment of multiple sensors over the area of interest. This clearly paves the way for several important practical monitoring applications concerning, e.g., weather forecasting [1], water flow regulation [2], fire detection, diffusion of pollutants [3], smart grids [4], vehicular traffic [5]. The problem of fusing data from different sensors can be accomplished either in a *centralized* way, i.e. when there is a single fusion center collecting data from all sensors and taking care of the overall spatial domain of interest, or in *distributed* (decentralized) fashion with multiple intercommunicating fusion centers (nodes) each of which can only access part of the sensor data and take care of a sub-region of the overall domain. The decentralized approach is preferable in terms of scalability

Manuscript received September 16, 2016; revised September 19, 2016; accepted November 9, 2016. Date of publication December 7, 2016; date of current version June 26, 2017. Recommended by Associate Editor C. Seatzu.

The authors are with the Dipartimento di Ingegneria dell'Informazione (DINFO), Università di Firenze, Firenze, Italy (e-mail: giorgio.battistelli@unifi.it; luigi.chisci@unifi.it; nicola.forti@unifi.it; giuseppe.pelosi@unifi.it; stefano.selleri@unifi.it).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TAC.2016.2636659

of computation with the problem size and will be, therefore, undertaken in this paper.

Since spatially distributed processes are usually modeled as infinite-dimensional systems, governed by *partial differential equations* (PDEs), distributed state estimation for such systems turns out to be a key issue to be addressed. While a lot of work has dealt with distributed filters for finite-dimensional, both linear [6]–[9] and nonlinear [10], systems as well as for multitarget tracking [11], considerably less attention has been devoted to the more difficult case of distributed-parameter systems.

Recent work [12]–[16] has addressed the design of distributed state estimators/observers for large-scale systems formed by the sparse interconnection of many subsystems (compartments). Such systems are possibly (but not necessarily) originated from spatial discretization of PDEs. In particular, [12] presents a fully scalable distributed Kalman filter based on a suitable spatial decomposition of a complex large-scale system as well as on appropriate observation fusion techniques among the local Kalman filters. In [13], non-scalable consensus-based multi-agent estimators are proposed wherein each agent aims to estimate the state of the whole large-scale system. In [14], a moving-horizon partition-based approach is followed in order to estimate the state of a large-scale interconnected system and decentralization is achieved via suitable approximations of covariances. Further, [15] deals with dynamic field estimation by wireless sensor networks with special emphasis on sensor scheduling for trading off communication/energy efficiency versus estimation performance. In [16], design of distributed continuous-time observers for partitioned linear systems is addressed.

As for the specific case of distributed-parameter systems, interesting contributions have been provided in [17], [18] which present consensus filters wherein each node of the network aims to estimate the system state on the whole spatial domain of interest.

In the present paper, as compared to [17], [18], a different strategy is adopted in which each node is only responsible for estimating the state over a sub-domain of the overall domain. This setup allows for a solution which is *scalable* with respect to the spatial domain (i.e., the computational complexity in each node does not depend on the size of the whole spatial domain but only of its region of competence). In this context, the contributions of the present paper are summarized as follows:

- 1) We develop *scalable* distributed filters for distributed-parameter systems by suitably adapting the so-called

Schwarz decomposition methods [19]–[24], which allow to split the overall domain into smaller subdomains and assign each of them to different interconnected processing nodes.

- 2) We exploit the *finite element* (FE) method [25]–[27] in order to approximate the original infinite-dimensional filtering problem into a, possibly large-scale, finite-dimensional one. Combining these two ingredients, we propose a novel distributed *finite element Kalman filter* which generalizes to the more challenging distributed case previous work on FE Kalman filtering [28], [29].
- 3) We show that the parallel FE-based implementation of the Schwarz method on the overall system is equivalent to performing a novel time-discretization scheme on the interconnected subsystems. Furthermore, we verify the well-posedness of the proposed discretization method in terms of numerical stability (i.e., in terms of boundedness and convergence of the time-discretization errors).
- 4) We provide results on the stability of the proposed distributed FE Kalman filter. Last but not least, a practical procedure, which requires the tuning of only one (or few) scalar parameters, is provided to check and guarantee the stability property.

Preliminary ideas on the topic can be found in [30].

The rest of the paper is structured as follows. Section II introduces the basic notation and problem formulation. Then Section III presents the centralized FE Kalman filter for distributed-parameter systems. Section IV shows how to extend such a filter to the distributed setting by means of the parallel Schwarz method and analyzes the numerical stability in terms of boundedness and convergence of the discretization errors. Then, Section V provides results on the exponential stability of the proposed distributed FE Kalman filter while Section VI demonstrates its effectiveness via numerical examples related to the estimation of a bi-dimensional temperature field. Finally, Section VII ends the paper with concluding remarks and perspectives for future work.

## II. PROBLEM FORMULATION

This paper addresses the estimation of a scalar, time-and-space-dependent, field from given discrete, in both time and space, measurements related to such a field provided by multiple sensors placed within the domain of interest. Let  $\Omega$  be a bounded domain of a  $d$ -dimensional Euclidean space  $\mathbb{R}^d$  with boundary  $\partial\Omega$ , where  $d \in \{1, 2, 3\}$ . The spatial coordinate vector is denoted by  $\mathbf{p} \in \Omega$ . The scalar field to be estimated  $x(\mathbf{p}, t)$  is defined over the space-time domain  $\Omega \times \mathbb{R}_+$ , as the solution of a *partial differential equation* (PDE) of the form

$$\frac{\partial x}{\partial t} + \mathcal{A}(x) = f \quad (1)$$

with (possibly unknown) initial condition  $x(\mathbf{p}, 0) = x_0(\mathbf{p})$ ,  $\mathbf{p} \in \Omega$ , and homogeneous boundary conditions

$$\mathcal{B}(x) = 0 \text{ on } \partial\Omega. \quad (2)$$

The dynamic field is observed by a network of sensors  $i \in \mathcal{S} \triangleq \{1, \dots, S\}$  placed at the spatial locations  $\mathbf{s}_i \in \Omega$ , which provide the measurements

$$y_{q,i} = h_i(x(\mathbf{s}_i, t_q)) + v_{q,i} \quad (3)$$

collected at discrete sampling instants  $t_q, q \in \mathbb{Z}_+ = \{1, 2, \dots\}$ , such that  $0 < t_1 < t_2 < \dots$ . In (1)–(3):  $f(\mathbf{p}, t)$  is a forcing term possibly affected by process noise;  $h_i(\cdot)$  is the measurement function of sensor  $i$ ;  $v_{q,1}, \dots, v_{q,S}$  are mutually independent white measurement noise sequences, also independent from the initial state  $x_0(\mathbf{p}) = x(\mathbf{p}, 0)$  for any  $\mathbf{p} \in \Omega$ .

The aim is to design a decentralized Kalman filter for spatially distributed systems, i.e. to solve in a fully distributed fashion the infinite-dimensional filtering problem of estimating the state  $x(\mathbf{p}, t)$  of system (1)–(2) given the locally gathered measurements (3). The proposed solution relies on (i) the FE method [25]–[26] for the approximation of the above problem into a finite-dimensional one, and (ii) a domain decomposition method for the subdivision of the system into interconnected subsystems with possibly overlapping states. The idea is to decompose the original problem on the whole domain of interest into estimation subproblems concerning smaller subdomains, and then assign such subproblems to different *nodes* which can locally process and exchange data in order to estimate their own state. This ensures scalability of the distributed filter for monitoring the target field. To this end, let us consider the set of nodes  $\mathcal{N} = \{1, \dots, N\}$ , subdivide the domain  $\Omega$  into possibly overlapping subdomains  $\Omega_m, m \in \mathcal{N}$ , such that  $\Omega = \bigcup_{m \in \mathcal{N}} \Omega_m$ . Further, let  $\mathbf{y}_q^m \triangleq \text{col}\{y_{q,i} : \mathbf{s}_i \in \Omega_m\}$  denote the vector of local measurements available to node  $m$  at time  $t_q$ . Then, the task of each node  $m$  is to estimate the field  $x$  over the corresponding subdomain  $\Omega_m$  exploiting only the local measurements  $\mathbf{y}_q^m$  and the information coming from the nodes associated to the neighboring subdomains.

Throughout the paper, we make the following assumptions.

A1)  $\mathcal{A}(\cdot)$  and  $\mathcal{B}(\cdot)$  are linear operators over a suitable Hilbert space  $V$ , with  $\mathcal{A}(\cdot)$  self-adjoint.

A2) Under the boundary conditions (2), the quadratic form  $\int_{\Omega} \mathcal{A}(x) x d\mathbf{p}$  is bounded and coercive (i.e., positive definite).

A third and last assumption A3 on the properties of the local measurement function and local observability will be introduced in Section V (to which we refer for a formal definition of local observability and for a discussion of its implications).

An example of the above general problem is the estimation of the temperature field  $x$  over the spatial domain of interest given point measurements of temperature sensors. In this case,  $V$  is usually taken as the Sobolev space  $H^1(\Omega)$ , the measurement function is simply  $h(x) = x$ , while the PDE (1) reduces to the well known *heat equation* with  $\mathcal{A}(x) = -\nabla \cdot (\lambda \nabla(x))$  and  $\mathcal{B}(x) = \alpha \partial x / \partial \mathbf{n} + \beta x$  with  $\alpha(\mathbf{p})\beta(\mathbf{p}) \geq 0$ ,  $\alpha(\mathbf{p}) + \beta(\mathbf{p}) > 0$ ,  $\forall \mathbf{p} \in \partial\Omega$ . Here  $\lambda(\mathbf{p})$  is the *thermal diffusivity*,  $\cdot$  stands for scalar product,  $\nabla \triangleq \partial / \partial \mathbf{p}$  denotes the gradient operator,  $\mathbf{n}$  is the outward pointing unit normal vector of the boundary  $\partial\Omega$ , and  $\partial x / \partial \mathbf{n} = \nabla x \cdot \mathbf{n}$ . Clearly, when the thermal diffusivity is

space-independent, one has  $\mathcal{A}(x) = -\lambda \nabla^2(x)$ , where  $\nabla^2 = \nabla \cdot \nabla$  is the Laplacian operator.

Notice that considering homogeneous boundary conditions as in (2) is not restrictive, since the non-homogeneous case  $\mathcal{B}(x) = g$  on  $\partial\Omega$  can be subsumed into the homogeneous one by means of the change of variables  $z = x - w$ , where  $w$  is any function belonging to  $V$  and satisfying the non-homogeneous boundary conditions.

### III. CENTRALIZED FINITE ELEMENT KALMAN FILTER

In this section, it is shown how to approximate the continuous-time infinite-dimensional system (1) into a discrete-time finite-dimensional linear dynamical system within the FE framework, and how, thanks to this space-time discretization, a centralized filter for field estimation can be directly designed.

By subdividing the domain  $\Omega$  into a suitable set of non overlapping regions, or elements, and by defining a suitable set of basis functions  $\phi_j(\mathbf{p}) \in V$  ( $j = 1, \dots, n$ ) on them, it is possible to write an approximation of the unknown function  $x(\mathbf{p}, t)$  as [25], [26]

$$x(\mathbf{p}, t) \approx \sum_{j=1}^n \phi_j(\mathbf{p}) x_j(t) = \boldsymbol{\phi}^T(\mathbf{p}) \mathbf{x}(t) \quad (4)$$

where:  $x_j(t)$  is the unknown expansion coefficient of function  $x(\mathbf{p}, t)$  relative to time  $t$  and basis function  $\phi_j(\mathbf{p})$ ;  $\boldsymbol{\phi}(\mathbf{p}) \triangleq \text{col}\{\phi_j(\mathbf{p})\}_{j=1}^n$  and  $\mathbf{x}(t) \triangleq \text{col}\{x_j(t)\}_{j=1}^n$ .

The choices of the basis functions  $\phi_j$  and of the elements are key points of the FE method. Typically, the elements (triangles or quadrilaterals in 2D, tetrahedral or polyhedral in 3D) define a FE mesh with vertices  $\mathbf{p}_j \in \Omega$ ,  $j = 1, \dots, n$ . Then each basis function  $\phi_j$  is a piece-wise polynomial which vanishes outside the FEs around  $\mathbf{p}_j$  and such that  $\phi_j(\mathbf{p}_i) = \delta_{ij}$ ,  $\delta_{ij}$  denoting the Kronecker delta.

In order to apply the Galerkin weighted residual method, let the PDE (1) be recast in the following (weak) integral form

$$\int_{\Omega} \frac{\partial x}{\partial t} \psi \, d\mathbf{p} + \int_{\Omega} \mathcal{A}(x) \psi \, d\mathbf{p} = \int_{\Omega} f \psi \, d\mathbf{p} \quad (5)$$

where  $\psi(\mathbf{p})$  is a generic space-dependent weight function. Then, by choosing the test function  $\psi(\mathbf{p})$  equal to the selected basis functions  $\phi_j$  and exploiting the approximation (4) in (5), we get

$$\int_{\Omega} \frac{\partial x}{\partial t} \phi_j \, d\mathbf{p} + \int_{\Omega} \mathcal{A}(x) \phi_j \, d\mathbf{p} = \int_{\Omega} f \phi_j \, d\mathbf{p} \quad j = 1, \dots, n.$$

Stacking (one on top of the other) the above scalar equations into a single vector equation, yields

$$\int_{\Omega} \boldsymbol{\phi} \frac{\partial}{\partial t} (\boldsymbol{\phi}^T \mathbf{x}) \, d\mathbf{p} + \int_{\Omega} \boldsymbol{\phi} \mathcal{A}(\boldsymbol{\phi}^T \mathbf{x}) \, d\mathbf{p} = \int_{\Omega} \boldsymbol{\phi} f \, d\mathbf{p}$$

from which, defining  $\mathcal{A}(\boldsymbol{\phi}) \triangleq \text{col}\{\mathcal{A}(\phi_j)\}_{j=1}^n$  and thanks to the linearity of operator  $\mathcal{A}(\cdot)$ , the usual FE weak form is obtained

[25]–[26]

$$\begin{aligned} & \underbrace{\left[ \int_{\Omega} \boldsymbol{\phi}(\mathbf{p}) \boldsymbol{\phi}^T(\mathbf{p}) \, d\mathbf{p} \right]}_{\mathbf{M}} \dot{\mathbf{x}}(t) + \underbrace{\left[ \int_{\Omega} \boldsymbol{\phi}(\mathbf{p}) [\mathcal{A}(\boldsymbol{\phi}(\mathbf{p}))]^T \, d\mathbf{p} \right]}_{\mathbf{S}} \mathbf{x}(t) \\ &= \underbrace{\int_{\Omega} \boldsymbol{\phi}(\mathbf{p}) f(\mathbf{p}, t) \, d\mathbf{p}}_{\mathbf{u}(t)}. \end{aligned} \quad (6)$$

It is evident how the first two integrals in (6) depend only on basis functions and can be computed *a priori*. In particular, the first integral yields the well known *mass* matrix  $\mathbf{M}$ , while the second depends on the operator  $\mathcal{A}(\cdot)$  and, in the thermal case, is the stiffness matrix  $\mathbf{S}$  [25]. The third integral depends on the forcing term  $f$ , which is assumed to be known, and can hence be computed *a priori*, leading to a time dependent vector contribution  $\mathbf{u}(t)$ .

It is worth pointing out that, in the FE weak form (6), the boundary conditions (2) can be accounted for in two different ways [25], [26]. The so-called *essential* boundary conditions are handled by imposing them on the solution, i.e., by choosing basis functions belonging to  $V_0 = \{x \in V : \mathcal{B}(x) = 0 \text{ on } \partial\Omega\}$ . On the other hand, the so-called *natural* boundary conditions are directly incorporated into the weak form (5). For example, in the case of the heat equation, the (isotherm) homogeneous Dirichlet boundary conditions  $x = 0$  on  $\partial\Omega$  are essential, while the (adiabatic) homogeneous Neumann boundary conditions  $\partial x / \partial \mathbf{n} = 0$  are natural. Of course, by letting the functions  $\alpha$  and  $\beta$  vary on  $\partial\Omega$ , we can also have a problem with mixed essential/natural boundary conditions. In all the cases, the resulting linear differential equation takes the form

$$\mathbf{M} \dot{\mathbf{x}} + \mathbf{S} \mathbf{x} = \mathbf{u} + \boldsymbol{\epsilon} \quad (7)$$

where  $\boldsymbol{\epsilon}$  arises from the approximation error<sup>1</sup> in the finite-dimensional representation (4) of  $x$  in terms of basis functions. Notice that  $\mathbf{M}$  turns out to be positive definite by linear independence of the basis functions  $\phi_j(\cdot)$ . Further,  $\mathbf{S}$  is positive definite as well thanks to the coercivity of the quadratic form in the left-hand side of (5) of assumption A2. Hence the system (7) turns out to be asymptotically stable, the state transition matrix  $-\mathbf{M}^{-1}\mathbf{S}$  being well defined and strictly Hurwitz thanks to the positive definiteness of  $\mathbf{M}$  and  $\mathbf{S}$ . System (7) can be discretized in time by different methods (e.g., backward or forward Euler integration, or the zero-order-hold method) to provide the discrete-time state-space model

$$\mathbf{x}_{k+1} = \mathbf{A} \mathbf{x}_k + \mathbf{B} \mathbf{u}_k + \mathbf{w}_k \quad (8)$$

where the process noise  $\mathbf{w}_k$  has been introduced to account for the various uncertainties and/or imprecisions (e.g. FE approximation, time discretization, and imprecise knowledge of boundary conditions). Specifically, the backward Euler method (here adopted for stability issues) leads to a marching in time

<sup>1</sup> If  $x$  is sufficiently smooth, then the FE approximation error is point-wise bounded and converges to zero as the size of the FE mesh tends to zero.

FE implementation [27] which yields (8) with

$$\mathbf{A} = (\mathbf{I} + \Delta \mathbf{M}^{-1} \mathbf{S})^{-1}, \mathbf{B} = \mathbf{A} \mathbf{M}^{-1} \Delta, \\ \mathbf{u}_k \triangleq \mathbf{u}((k+1)\Delta), \mathbf{x}_k \triangleq \mathbf{x}(k\Delta) = \text{col}\{x_j(k\Delta)\}_{j=1}^n$$

where  $\Delta$  denotes the time integration interval. Notice that  $\mathbf{A}$  is well defined for any  $\Delta > 0$  since both  $\mathbf{M}$  and  $\mathbf{S}$  are positive definite.

In the following, for the sake of notational simplicity, it will be assumed that each sampling instant is a multiple of  $\Delta$ , i.e.,  $t_q = T_q \Delta$  with  $T_q \in \mathbb{Z}_+$ , and we let  $\mathcal{T} = \{T_1, T_2, \dots\}$ ; irregular sampling could, however, be easily dealt with. This amounts to assuming that the numerical integration rate of the PDE (1) in the filter can be higher than the measurement collection rate, which can be useful in order to reduce numerical errors. In a centralized setting where all sensor measurements are available to the filter, the measurement equation (3) takes the discrete-time form

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \quad (9)$$

for any  $k = T_q \in \mathcal{T}$ , where

$$\mathbf{y}_k \triangleq \text{col}\{y_{q,i}\}_{i \in \mathcal{S}}, \mathbf{h}(\mathbf{x}) \triangleq \text{col}\{h_i(\phi^T(\mathbf{s}_i)\mathbf{x})\}_{i \in \mathcal{S}}, \\ \mathbf{v}_k \triangleq \text{col}\{v_{q,i}\}_{i \in \mathcal{S}}$$

In particular, in the case wherein all sensors directly measure the target field  $x$ , i.e.  $h_i(x) = x$  for all  $i \in \mathcal{S}$ , the measurement equation (9) turns out to be linear with  $\mathbf{h}(\mathbf{x}) = \mathbf{C}\mathbf{x}$ , where

$$\mathbf{C} = \text{col}\{\phi^T(\mathbf{s}_i)\}_{i \in \mathcal{S}} \quad (10)$$

Summarizing, the original infinite-dimensional continuous-time problem has been reduced to a much simpler finite-dimensional (possibly large-scale) discrete time filtering problem (a linear one provided that all sensor measurement functions are linear) to which the *Kalman filter*, or *extended Kalman filter* when sensor nonlinearities are considered, can be readily applied. The resulting centralized filter recursion becomes:

$$\hat{\mathbf{x}}_{k|k} = \begin{cases} \hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k (\mathbf{y}_k - \mathbf{h}(\hat{\mathbf{x}}_{k|k-1})) & \text{if } k \in \mathcal{T} \\ \hat{\mathbf{x}}_{k|k-1} & \text{otherwise} \end{cases} \\ \mathbf{P}_{k|k} = \begin{cases} \mathbf{P}_{k|k-1} - \mathbf{L}_k \mathbf{C}_k^T \mathbf{P}_{k|k-1} & \text{if } k \in \mathcal{T} \\ \mathbf{P}_{k|k-1} & \text{otherwise} \end{cases} \\ \hat{\mathbf{x}}_{k+1|k} = \mathbf{A} \hat{\mathbf{x}}_{k|k} + \mathbf{B} \mathbf{u}_k \\ \mathbf{P}_{k+1|k} = \mathbf{A} \mathbf{P}_{k|k} \mathbf{A}^T + \mathbf{Q}_k \quad (11)$$

where

$$\mathbf{C}_k = \frac{\partial \mathbf{h}}{\partial \mathbf{x}}(\hat{\mathbf{x}}_{k|k-1}) \\ \mathbf{L}_k = \mathbf{P}_{k|k-1} \mathbf{C}_k (\mathbf{R}_k + \mathbf{C}_k \mathbf{P}_{k|k-1} \mathbf{C}_k^T)^{-1} \quad (12)$$

for  $k \in \mathcal{T}$ . The recursion is initialized from suitable  $\hat{\mathbf{x}}_{1|0}$  and  $\mathbf{P}_{1|0} = \mathbf{P}_{1|0}^T > \mathbf{0}$ . In (11),  $\mathbf{Q}_k$  and  $\mathbf{R}_k$  denote the covariance matrices of the process noise  $\mathbf{w}_k$  and, respectively, measurement noise  $\mathbf{v}_k$ .

The following two remarks concerning optimality of the Kalman filter and, respectively, handling of sensor nonlinearities are in order.

*Remark 1:* Notice that the process noise  $\mathbf{w}_k$  in (8) arises from the superposition of several uncertainties and/or perturbations (including, e.g., the FE approximation of the continuous field) so that its whiteness and uncorrelation with the initial state, usually assumed in a stochastic framework, do not hold true in practice. As a result, the Kalman filter algorithm (11)–(12), even in the linear case  $\mathbf{h}(\mathbf{x}) = \mathbf{C}\mathbf{x}$ , loses its Bayes optimality but still preserves deterministic least-squares optimality as the minimizer of the following cost function

$$J = (\mathbf{x}_1 - \hat{\mathbf{x}}_{1|0})^T \mathbf{P}_{1|0}^{-1} (\mathbf{x}_1 - \hat{\mathbf{x}}_{1|0}) + \\ \sum_{i=1}^{k-1} (\mathbf{x}_{i+1} - \mathbf{A}\mathbf{x}_i)^T \mathbf{Q}_i^{-1} (\mathbf{x}_{i+1} - \mathbf{A}\mathbf{x}_i) + \\ \sum_{i=1}^k (\mathbf{y}_i - \mathbf{C}\mathbf{x}_i)^T \mathbf{R}_i^{-1} (\mathbf{y}_i - \mathbf{C}\mathbf{x}_i)$$

*Remark 2:* Sensor nonlinearities, provided that the measurement functions  $h_i(\cdot)$  in (3) are invertible, can be handled by applying the inverse measurement functions  $h_i^{-1}(\cdot)$  to the sensor outputs, i.e. by defining transformed sensor outputs  $y'_{q,i} = h_i^{-1}(y_{q,i})$  and considering the transformed linear measurement equations

$$y'_{q,i} = x(\mathbf{s}_i, t_q) + v'_{q,i} \quad (13)$$

in place of (3). This approach has the advantage of eliminating any need for a nonlinear filter. However, while (13) is exact in the ideal, noiseless, case i.e. when  $v_{q,i} = v'_{q,i} = 0$ , it becomes only an approximation in presence of measurement noise. In particular, even if  $v_{q,i}$  in (3) can be reasonably assumed to be zero-mean, white and uncorrelated with the state  $x(\mathbf{s}_i, t_q)$ , non-negligible biases and/or correlations can be induced by the nonlinear transformation  $h_i^{-1}(\cdot)$  in the noise term  $v'_{q,i}$  appearing in (13). For this reason, depending on the particular measurement function under consideration, the use of truly nonlinear filters can be useful also when the sensor nonlinearity is invertible. For non-invertible sensor nonlinearities, nonlinear filters such as, for instance, the extended Kalman filter for sufficiently smooth  $h_i(\cdot)$  or the unscented Kalman filter for arbitrary  $h_i(\cdot)$ , must be used.

#### IV. DISTRIBUTED FINITE ELEMENT KALMAN FILTER

In order to develop a scalable distributed filter for monitoring the target field, the idea is to run in each node  $m \in \mathcal{N}$  a field estimator for the region  $\Omega_m$  exploiting local measurements  $\mathbf{y}_q^m$ , information from the nodes assigned to neighboring subdomains, as well as the PDE model (1) properly discretized in time and space. The proposed approach takes inspiration from the parallel Schwarz method, originally conceived [19] for an iterative solution of boundary value problems. Subsequently, the Schwarz method has received renewed interest [20], [21] in connection with the parallelization of PDE solvers.

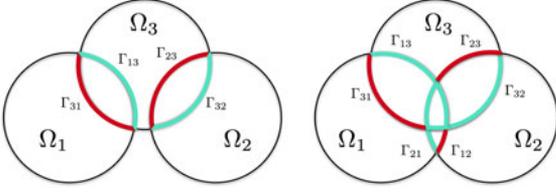


Fig. 1. Definition of interfaces  $\Gamma_{mj}$  in two different configurations with three overlapping subdomains.

Let us define, for any  $m \in \mathcal{N}$ , a partition  $\{\Gamma_{mj}\}_{j \in \mathcal{N}_m}$  of  $\partial\Omega_m$  (the boundary of  $\Omega_m$ ) such that

$$\begin{aligned} \Gamma_{mm} &= \partial\Omega \cap \partial\Omega_m \\ \partial\Omega_m &= \bigcup_{j \in \mathcal{N}_m} \Gamma_{mj} \\ \Gamma_{mj} &\subset \Omega_j, \quad \forall j \neq m \\ \Gamma_{mj} \cap \Gamma_{mh} &= \emptyset, \quad \forall j \neq h \end{aligned} \quad (14)$$

In this way, each piece  $\Gamma_{mj}$  of  $\partial\Omega_m$  for any  $j \in \mathcal{N}_m \setminus \{m\}$  is uniquely assigned to node  $j$ . Notice that in the above definitions, for each node  $m$ ,  $\mathcal{N}_m$  indicates the in-neighborhood of node  $m$ , where  $j$  is called an in-neighbor of node  $m$  whenever  $\Gamma_{mj} \neq \emptyset$  (by definition,  $\mathcal{N}_m$  includes the node  $m$ ). This clearly originates a directed network (graph)  $\mathcal{G} = (\mathcal{N}, \mathcal{L})$  with node set  $\mathcal{N}$  and link set  $\mathcal{L} \triangleq \{(j, m) \in \mathcal{N} \times \mathcal{N} : \Gamma_{mj} \neq \emptyset\}$  (see Fig. 1).

In order to describe the filtering cycle to be implemented in node  $m$  within the sampling interval  $[t_q, t_{q+1})$ , let us assume that at time  $t_q^-$ , before the acquisition of  $\mathbf{y}_q^m$ , such a node is provided with a prior estimate  $\hat{x}_{q|q-1}^m$  as the result of the previous filtering cycles. Let  $\delta$  be the time interval necessary for performing a *distributed prediction* step consisting of an information exchange between neighbors and a local field prediction over a subdomain. Then,  $L_q \triangleq (t_{q+1} - t_q) / \delta$  represents the number of distributed prediction steps (equal to the number of allowed data exchanges) in the  $q$ -th sampling interval. Note that, for the sake of notational simplicity, hereafter it is supposed that  $t_{q+1} - t_q$  is an integer multiple of  $\delta$ , i.e.,  $L_q \in \mathbb{Z}_+$ . Anyway, the method could easily encompass the general case. Then, the above mentioned filtering cycle for the proposed distributed estimation algorithm essentially consists of:

- 1) Correction, i.e. incorporation (assimilation) of the last measurement  $\mathbf{y}_q^m$  into the current estimate;
- 2) Distributed prediction, i.e. alternate exchanges of estimates with the neighborhood  $\mathcal{N}_m$  and predictions over the time sub-intervals  $[t_q + (\ell - 1)\delta, t_q + \ell\delta]$  for  $\ell = 1, \dots, L_q$ , i.e.  $L_q$  times.

The proposed *Parallel Schwarz* filter is detailed in Table I.

Some remarks concerning the above reported algorithm are in order. As it can be seen from step 5, the information received by neighboring nodes is taken into account by explicitly imposing the non-homogeneous Dirichlet interface conditions (16) on  $\Gamma_{mj}$ ,  $j \in \mathcal{N}_m \setminus \{m\}$ . Clearly, a delay is introduced in those terms concerning neighboring nodes which makes the algorithm well-suited for distributed computation. With this respect, it is worth pointing out that the proposed algorithm is based on the

TABLE I  
ALGORITHM 1: PARALLEL SCHWARZ FILTER

1:	Given $\mathbf{y}_q^m$ , update the prior estimate $\hat{x}_{q q-1}^m$ into $\hat{x}_{q q}^m$ .
2:	Initialize the prediction with $\hat{x}_{q,0}^m = \hat{x}_{q q}^m$ and $\hat{x}_{q,-1}^m = \hat{x}_{q q}^m$ .
3:	<b>for</b> $\ell = 1, \dots, L_q$ <b>do</b>
4:	Exchange data with the neighborhood; specifically send to neighbor $j$ the data $\hat{x}_{q,\ell-1}^m$ concerning the sub-boundary $\Gamma_{jm} \subset \partial\Omega_j$ , and get from neighbor $j$ the data $\hat{x}_{q,\ell-1}^j$ concerning the sub-boundary $\Gamma_{mj} \subset \partial\Omega_m$ .
5:	Solve the problem $\frac{\hat{x}_{q,\ell}^m - \hat{x}_{q,\ell-1}^m}{\delta} + \mathcal{A}(\hat{x}_{q,\ell}^m) = f_{q,\ell} \text{ in } \Omega_m \quad (15)$ subject to the Dirichlet boundary conditions $\hat{x}_{q,\ell}^m = \hat{x}_{q,\ell-1}^j \text{ on } \Gamma_{mj} \quad \forall j \in \mathcal{N}_m \setminus \{m\} \quad (16)$ and the linear boundary conditions $\mathcal{B}(\hat{x}_{q,\ell}^m) = 0 \text{ on } \Gamma_{mm} \quad (17)$ where $f_{q,\ell}(\mathbf{p}) \triangleq f(\mathbf{p}, t_q + \ell\delta)$ .
6:	<b>end for</b>
7:	Set $\hat{x}_{q+1 q}^m = \hat{x}_{q,L_q}^m$ for the next cycle.

parallel Schwarz method for evolution problems, which, as well known, enjoys nice convergence properties to the centralized solution as the time discretization step  $\delta$  tends to zero [20]–[21]. Hence, it seems a sensible and promising approach to spread the information through the network.

#### A. Implementation via the Finite-Element Method

In practice, the algorithm, and in particular the solution of the boundary value problem (15)–(17), has to be implemented via a finite dimensional approximation. In particular, we follow the same approach described in Section III for the centralized case by constructing a FE mesh for the global domain  $\Omega$ , and then decomposing such a grid into  $N$  possibly overlapping sub-meshes, according to the domain decomposition. For the sequel, it is important to distinguish vertices lying on the boundary between neighbors (*interface*) from the other vertices of the subdomain. To this end, let  $\text{int}(S)$  denote the interior of a generic set  $S$ . Then, we introduce the sets of indices  $\mathcal{I}_m \triangleq \{i : \mathbf{p}_i \in \text{int}(\Omega_m) \cup \Gamma_{mm}\}$  and  $\mathcal{J}_{mj} \triangleq \{i : \mathbf{p}_i \in \Gamma_{mj}\}$  of the basis functions corresponding to *internal* and, respectively, *interface* vertices of subdomain  $\Omega_m$ . In particular, let  $\mathbf{x}^m \triangleq \text{col}\{x_i : i \in \mathcal{I}_m\}$ ,  $m = 1, \dots, N$ , denote the vector of field values in vertices belonging to  $\text{int}(\Omega_m) \cup \Gamma_{mm}$ , i.e. the *internal* state of subsystem  $m$ . Then, it is possible to extract from (7) the rows relative to states  $\mathbf{x}^m$  so that

$$\begin{aligned} \mathbf{M}^{mm} \dot{\mathbf{x}}^m + \sum_{j \in \mathcal{N}_m \setminus \{m\}} \mathbf{M}^{mj} \dot{\mathbf{x}}^j + \mathbf{S}^{mm} \mathbf{x}^m \\ + \sum_{j \in \mathcal{N}_m \setminus \{m\}} \mathbf{S}^{mj} \mathbf{x}^j = \mathbf{u}^m + \boldsymbol{\epsilon}^m \end{aligned} \quad (18)$$

where the matrices  $\mathbf{M}^{mj}$  and  $\mathbf{S}^{mj}$  take into account the contribution of state variables in vertices  $\mathbf{p}_j \in \Gamma_{mj}$ , and  $\boldsymbol{\epsilon}^m$  accounts for the approximation error in the finite-dimensional representation (4) of  $x$  in terms of basis functions. Notice that both  $\mathbf{M}^{mm}$  and  $\mathbf{S}^{mm}$  are positive definite because so are  $\mathbf{M}$  and  $\mathbf{S}$ . As a result, the ODE (7) can be written as the interconnection of  $N$  subsystems of the form (18).

Each of the subsystems (18) can be discretized in time in the interval  $[t_q, t_{q+1}]$  using a modified backward Euler technique wherein a delay is introduced in those terms concerning neighboring nodes, so that at time  $t_q + \ell\delta$  we obtain the following discrete-time linear descriptor system

$$\begin{aligned} & \mathbf{M}^{mm} \left( \frac{\mathbf{x}_{q,\ell+1}^m - \mathbf{x}_{q,\ell}^m}{\delta} \right) + \mathbf{S}^{mm} \mathbf{x}_{q,\ell+1}^m \\ & + \sum_{j \in \mathcal{N}_m \setminus \{m\}} \left[ \mathbf{M}^{mj} \left( \frac{\mathbf{x}_{q,\ell}^j - \mathbf{x}_{q,\ell-1}^j}{\delta} \right) + \mathbf{S}^{mj} \mathbf{x}_{q,\ell}^j \right] \\ & = \mathbf{u}_{q,\ell+1}^m + \boldsymbol{\epsilon}_{q,\ell+1}^m + \boldsymbol{\tau}_{q,\ell}^m \end{aligned} \quad (19)$$

where  $\mathbf{x}_{q,\ell}^m \triangleq \mathbf{x}^m(t_q + \ell\delta)$ , for  $\ell = 1 \dots, L_q$ , and  $\boldsymbol{\tau}_{q,\ell}^m$  denotes the time discretization error at time  $t_q + \ell\delta$ . The recursion (16) is initialized at time  $t_q$  by setting

$$\begin{aligned} \mathbf{x}_{q,0}^m &= \mathbf{x}^m(t_q), \\ \mathbf{x}_{q,0}^j &= \mathbf{x}^j(t_q), \quad \mathbf{x}_{q,-1}^j = \mathbf{x}^j(t_q), \quad j \in \mathcal{N}_m \setminus \{m\} \end{aligned} \quad (20)$$

The well-posedness of the discretization scheme resulting from (19)–(20) will be analyzed in Section IV-B.

It can be readily seen that such a hybrid Euler time discretization implements the Parallel Schwarz method, described earlier. In fact, it is equivalent to approximate  $x$  in  $\Omega_m$  at time  $t_q + \ell\delta$  as

$$\begin{aligned} x(\mathbf{p}, t_q + \ell\delta) &\approx \sum_{i \in \mathcal{I}_m} \phi_i^m(\mathbf{p}) x_{q,\ell}^{m,i} \\ &+ \sum_{j \in \mathcal{N}_m \setminus \{m\}} \sum_{i \in \mathcal{I}_{m_j}} \phi_i^j(\mathbf{p}) x_{q,\ell-1}^{j,i} \end{aligned} \quad (21)$$

which in turn corresponds to explicitly imposing non-homogeneous Dirichlet interface conditions on  $\Gamma_{mj}$ ,  $j \in \mathcal{N}_m \setminus \{m\}$ , taken from neighboring nodes (like in (16)).

Thanks to the positive definiteness of  $\mathbf{M}^{mm}$  and  $\mathbf{S}^{mm}$ , each discretized-model (19) can be easily transformed into a state-space model of the form

$$\begin{aligned} \mathbf{x}_{q,\ell}^m &= \mathbf{A}^m \mathbf{x}_{q,\ell-1}^m + \sum_{j \in \mathcal{N}_m \setminus \{m\}} \mathbf{A}^{mj} \hat{\mathbf{x}}_{q,\ell-1}^j \\ &+ \sum_{j \in \mathcal{N}_m \setminus \{m\}} \bar{\mathbf{A}}^{mj} \mathbf{x}_{q,\ell-2}^j + \mathbf{B}^m \mathbf{u}_{q,\ell}^m + \mathbf{w}_{q,\ell}^m \end{aligned} \quad (22)$$

where

$$\begin{aligned} \mathbf{A}^m &= (\mathbf{M}^{mm} + \delta \mathbf{S}^{mm})^{-1} \mathbf{M}^{mm} \\ \mathbf{A}^{mj} &= (\mathbf{M}^{mm} + \delta \mathbf{S}^{mm})^{-1} (-\delta \mathbf{S}^{mj} - \mathbf{M}^{mj}) \\ \bar{\mathbf{A}}^{mj} &= (\mathbf{M}^{mm} + \delta \mathbf{S}^{mm})^{-1} \mathbf{M}^{mj} \\ \mathbf{B}^m &= (\mathbf{M}^{mm} + \delta \mathbf{S}^{mm})^{-1} \delta \end{aligned}$$

and  $\mathbf{w}_{q,\ell}^m = (\mathbf{M}^{mm} + \delta \mathbf{S}^{mm})^{-1} \delta (\boldsymbol{\epsilon}_{q,\ell+1}^m + \boldsymbol{\tau}_{q,\ell}^m)$  is the error combining the effects of both spatial and temporal discretizations.

Such interconnected models can be exploited so as to derive a FE approximation of the distributed-state estimation algorithm with Parallel Schwarz method (Algorithm 1 in Table I).

TABLE II  
ALGORITHM 2: DISTRIBUTED FINITE-ELEMENT KALMAN FILTER

- 
- 
- 1: Given  $\mathbf{y}_q^m$ , update the prior estimate  $\hat{\mathbf{x}}_{q|q-1}^m$  and covariance  $\mathbf{P}_{q|q-1}^m$  into  $\hat{\mathbf{x}}_{q|q}^m$  and  $\mathbf{P}_{q|q}^m$  as follows
 
$$\begin{aligned} \hat{\mathbf{x}}_{q|q}^m &= \hat{\mathbf{x}}_{q|q-1}^m + \mathbf{L}_q^m \left( \mathbf{y}_q^m - \mathbf{h}^m \left( \hat{\mathbf{x}}_{q|q-1}^m \right) \right) \\ \mathbf{P}_{q|q}^m &= \mathbf{P}_{q|q-1}^m - \mathbf{L}_q^m (\mathbf{C}_q^m)^T \mathbf{P}_{q|q-1}^m \\ \mathbf{C}_q^m &= \frac{\partial \mathbf{h}^m}{\partial \mathbf{x}} \left( \hat{\mathbf{x}}_{q|q-1}^m \right) \\ \mathbf{L}_q^m &= \mathbf{P}_{q|q-1}^m \mathbf{C}_q^m \left( \mathbf{R}_q^m + \mathbf{C}_q^m \mathbf{P}_{q|q-1}^m (\mathbf{C}_q^m)^T \right)^{-1} \end{aligned}$$
 where  $\mathbf{h}^m \triangleq \text{col} \{h_i : \mathbf{s}_i \in \Omega_m\}$  denote the local measurement function at node  $m$ .
  - 2: Initialize the distributed prediction with  $\hat{\mathbf{x}}_{q,0}^m = \hat{\mathbf{x}}_{q|q}^m$ ,  $\mathbf{P}_{q,0}^m = \mathbf{P}_{q|q}^m$  and  $\hat{\mathbf{x}}_{q,-1}^m = \hat{\mathbf{x}}_{q|q}^m$ ,  $\mathbf{P}_{q,-1}^m = \mathbf{P}_{q|q}^m$ .
  - 3: **for**  $\ell = 1, \dots, L_q$  **do**
  - 4: Exchange data with the neighborhood; specifically send to neighbor  $j$  the estimates  $\hat{\mathbf{x}}_{q,\ell-1}^m$  concerning the sub-boundary  $\Gamma_{mj} \subset \partial\Omega_j$ , and get from neighbor  $j$  the estimates  $\hat{\mathbf{x}}_{q,\ell-1}^j$  concerning the sub-boundary  $\Gamma_{mj} \subset \partial\Omega_m$ .
  - 5: set
 
$$\begin{aligned} \hat{\mathbf{x}}_{q,\ell}^m &= \mathbf{A}^m \hat{\mathbf{x}}_{q,\ell-1}^m + \sum_{j \in \mathcal{N}_m \setminus \{m\}} \mathbf{A}^{mj} \hat{\mathbf{x}}_{q,\ell-1}^j \\ &+ \sum_{j \in \mathcal{N}_m \setminus \{m\}} \bar{\mathbf{A}}^{mj} \hat{\mathbf{x}}_{q,\ell-2}^j + \mathbf{B}^m \mathbf{u}_{q,\ell}^m \end{aligned} \quad (23)$$

$$\mathbf{P}_{q,\ell}^m = \gamma^2 \mathbf{A}^m \mathbf{P}_{q,\ell-1}^m (\mathbf{A}^m)^T + \mathbf{Q}^m \quad (24)$$
 with  $\gamma > 1$ .
  - 6: **end for**
  - 7: Set  $\hat{\mathbf{x}}_{q+1|q}^m = \hat{\mathbf{x}}_{q,L_q}^m$  and  $\mathbf{P}_{q+1|q}^m = \mathbf{P}_{q,L_q}^m$  for the next cycle.
- 
- 

In particular, the numerical solution of (15)–(17) takes the form of the local one-step-ahead predictor for model (22) at time  $t_q + (\ell - 1)\delta$ , whereas the correction step of the local filtering cycle is the usual (extended) Kalman filter update step for the local subsystem. The resulting distributed finite-element (extended) Kalman filter is reported in Table II.

As previously shown, the additional terms  $\sum_{j \in \mathcal{N}_m \setminus \{m\}} \mathbf{A}^{mj} \hat{\mathbf{x}}_{q,\ell-1}^j$  and  $\sum_{j \in \mathcal{N}_m \setminus \{m\}} \bar{\mathbf{A}}^{mj} \hat{\mathbf{x}}_{q,\ell-2}^j$  in equation (22) arise from the non-homogeneous Dirichlet boundary conditions (16). In this respect, it is worth noting that the matrices  $\mathbf{A}^{mj}$  and  $\bar{\mathbf{A}}^{mj}$  are sparse since only the components of the neighbor estimates  $\hat{\mathbf{x}}_{q,\ell-1}^j$  and  $\hat{\mathbf{x}}_{q,\ell-2}^j$  concerning the sub-boundary  $\Gamma_{mj}$  are involved. The positive real  $\gamma > 1$  is a covariance boosting factor whose role, as will be discussed in the stability analysis of the distributed FE-KF, is that of guaranteeing convergence of the estimates. The covariance boosting factor is also necessary in order to compensate for the additional uncertainty associated with the boundary conditions at the interfaces, i.e., for the uncertainty associated with the estimates  $\sum_{j \in \mathcal{N}_m \setminus \{m\}} \mathbf{A}^{mj} \hat{\mathbf{x}}_{q,\ell-1}^j$  and  $\sum_{j \in \mathcal{N}_m \setminus \{m\}} \bar{\mathbf{A}}^{mj} \hat{\mathbf{x}}_{q,\ell-2}^j$ . In fact, such an uncertainty is not explicitly accounted for in (24) due to the fact that the correlation between the estimates of neighboring nodes is not precisely known. The interested reader is referred to [14] for additional insights on this issue in the context of distributed estimation of large-scale interconnected systems. As in the centralized context, the positive definite matrix  $\mathbf{Q}^m$  accounts for the various uncertainties and imprecisions (i.e., discretization errors, imprecise knowledge of the exogenous input  $f$  and of the boundary conditions (17)).

## B. Numerical Stability

As previously shown, in the FE-based implementation the Parallel Schwarz step amounts to performing a hybrid Euler discretization on the interconnection of the  $N$  subsystems (18). Hence, as a preliminary analysis step, it is important to verify the well-posedness of such a modified discretization method in terms of numerical stability (i.e., in terms of boundedness and convergence of the time-discretization errors). To this end, it is convenient to consider the global dynamics of the interconnection.

Let us consider the augmented global state  $\tilde{\mathbf{x}} \triangleq \text{col}\{\mathbf{x}^m, m = 1, \dots, N\}$ , which clearly contains repeated components of the state due to the possibly overlapping nature of the decomposition. Let the vectors  $\tilde{\mathbf{u}}$  and  $\tilde{\mathbf{e}}$  be defined in a similar way. In terms of  $\tilde{\mathbf{x}}$  the interconnection of the  $N$  subsystems of the form (18) gives rise to a global augmented system which obeys the following continuous-time linear dynamics

$$\tilde{\mathbf{M}} \dot{\tilde{\mathbf{x}}} + \tilde{\mathbf{S}} \tilde{\mathbf{x}} = \tilde{\mathbf{u}} + \tilde{\mathbf{e}} \quad (25)$$

Note that the only difference between (7) and (25) is the presence of duplicated states in the latter linear ODE. Nevertheless, the two systems originate an identical state evolution. According to the divide-and-conquer strategy, matrices  $\tilde{\mathbf{M}}$  and  $\tilde{\mathbf{S}}$  can be decomposed as

$$\tilde{\mathbf{M}} = \tilde{\mathbf{M}}_D + \tilde{\mathbf{M}}_F \quad (26)$$

$$\tilde{\mathbf{S}} = \tilde{\mathbf{S}}_D + \tilde{\mathbf{S}}_F \quad (27)$$

with  $\tilde{\mathbf{M}}_D = \text{block-diag}(\mathbf{M}^{11}, \dots, \mathbf{M}^{NN})$ ,  $\tilde{\mathbf{S}}_D = \text{block-diag}(\mathbf{S}^{11}, \dots, \mathbf{S}^{NN})$ , whereas  $\tilde{\mathbf{M}}_F$  and  $\tilde{\mathbf{S}}_F$  take into account the FE interconnection structure among neighboring subsystems. By substituting (26)–(27) into (25), one obtains

$$\tilde{\mathbf{M}}_D \dot{\tilde{\mathbf{x}}} + \tilde{\mathbf{S}}_D \tilde{\mathbf{x}} + \tilde{\mathbf{M}}_F \dot{\tilde{\mathbf{x}}} + \tilde{\mathbf{S}}_F \tilde{\mathbf{x}} = \tilde{\mathbf{u}} + \tilde{\mathbf{e}}. \quad (28)$$

Then, by applying the hybrid Euler time discretization (19), the time-discretized augmented system takes the form

$$\begin{aligned} \tilde{\mathbf{M}}_D \left( \frac{\tilde{\mathbf{x}}_{q,\ell+1} - \tilde{\mathbf{x}}_{q,\ell}}{\delta} \right) + \tilde{\mathbf{S}}_D \tilde{\mathbf{x}}_{q,\ell+1} + \tilde{\mathbf{M}}_F \left( \frac{\tilde{\mathbf{x}}_{q,\ell} - \tilde{\mathbf{x}}_{q,\ell-1}}{\delta} \right) \\ + \tilde{\mathbf{S}}_F \tilde{\mathbf{x}}_{q,\ell} = \tilde{\mathbf{u}}_{q,\ell+1} + \tilde{\mathbf{e}}_{q,\ell+1} + \boldsymbol{\tau}_{q,\ell} \end{aligned} \quad (29)$$

for  $\ell = 0, \dots, L_q - 1$ , where  $\tilde{\mathbf{x}}_{q,\ell} \triangleq \tilde{\mathbf{x}}(t_q + \ell\delta)$ , and, as previously,  $\boldsymbol{\tau}_{q,\ell}$  denotes the time discretization error at time  $t_q + \ell\delta$ . Further, the initialization (20) can be simply rewritten as

$$\tilde{\mathbf{x}}_{q,0} = \tilde{\mathbf{x}}_{q,-1} = \tilde{\mathbf{x}}(t_q) \quad (30)$$

The following result can now be stated which summarizes the numerical stability properties<sup>2</sup> of (29)–(30).

*Theorem 1:* The hybrid Euler time-discretization scheme (29)–(30) is *consistent* with local truncation error of order 1. Further, it is *zero-stable* provided that the following condition

holds

$$\rho(\tilde{\mathbf{M}}_D^{-1} \tilde{\mathbf{M}}_F) < 1 \quad (31)$$

where  $\rho(\cdot)$  denotes the spectral radius.

*Proof:* Let  $\mathcal{D}$  denote the differential operator in the left-hand side of (28), i.e.,

$$\mathcal{D}(\boldsymbol{\xi}, t) = \tilde{\mathbf{M}}_D \dot{\boldsymbol{\xi}}(t) + \tilde{\mathbf{S}}_D \boldsymbol{\xi}(t) + \tilde{\mathbf{M}}_F \dot{\boldsymbol{\xi}}(t) + \tilde{\mathbf{S}}_F \boldsymbol{\xi}(t)$$

for any smooth time-function  $\boldsymbol{\xi}$ . Further, let  $\mathcal{D}_\delta$  denote the discrete-time operator in the left-hand side of (29), i.e.,

$$\begin{aligned} \mathcal{D}_\delta(\boldsymbol{\xi}, t) = \tilde{\mathbf{M}}_D \left( \frac{\boldsymbol{\xi}(t+\delta) - \boldsymbol{\xi}(t)}{\delta} \right) + \tilde{\mathbf{S}}_D \boldsymbol{\xi}(t+\delta) \\ + \tilde{\mathbf{M}}_F \left( \frac{\boldsymbol{\xi}(t) - \boldsymbol{\xi}(t-\delta)}{\delta} \right) + \tilde{\mathbf{S}}_F \boldsymbol{\xi}(t). \end{aligned}$$

As well known, the time-discretization scheme (29) is consistent when, for any smooth time-function  $\boldsymbol{\xi}$  and for any time  $t$ ,  $\mathcal{D}_\delta(\boldsymbol{\xi}, t)$  converges to  $\mathcal{D}(\boldsymbol{\xi}, t)$  as  $\delta$  goes to 0. By taking the Taylor expansion of  $\boldsymbol{\xi}$  in  $t$ , we can write  $\boldsymbol{\xi}(t+\delta) = \boldsymbol{\xi}(t) + \delta \dot{\boldsymbol{\xi}}(t) + \delta^2 \ddot{\boldsymbol{\xi}}(t) + O(\delta^3)$  and  $\boldsymbol{\xi}(t-\delta) = \boldsymbol{\xi}(t) - \delta \dot{\boldsymbol{\xi}}(t) + \delta^2 \ddot{\boldsymbol{\xi}}(t) + O(\delta^3)$ . Hence, after some algebra, we have

$$\begin{aligned} \mathcal{D}_\delta(\boldsymbol{\xi}, t) = \mathcal{D}(\boldsymbol{\xi}, t) + \tilde{\mathbf{M}}_D \delta \ddot{\boldsymbol{\xi}}(t) + \tilde{\mathbf{S}}_D \delta \dot{\boldsymbol{\xi}}(t) - \tilde{\mathbf{M}}_F \delta \ddot{\boldsymbol{\xi}}(t) \\ + O(\delta^2) \end{aligned}$$

which shows that the scheme is consistent and the local truncation error has order 1.

In order to study zero-stability, we start by considering the limit for  $\delta$  going to zero of the time-difference equation (29), which is given by

$$\tilde{\mathbf{M}}_D (\tilde{\mathbf{x}}_{q,\ell+1} - \tilde{\mathbf{x}}_{q,\ell}) + \tilde{\mathbf{M}}_F (\tilde{\mathbf{x}}_{q,\ell} - \tilde{\mathbf{x}}_{q,\ell-1}) = 0. \quad (32)$$

In fact, zero-stability of the time-discretization scheme (29) corresponds to the neutral stability of the discrete-time system (32) (recall that system (32) is neutrally stable when its trajectories remain bounded as  $\ell$  goes to infinity for any initial condition). Then the proof can be concluded by noting that, by defining  $\zeta_{q,\ell+1} = \tilde{\mathbf{x}}_{q,\ell+1} - \tilde{\mathbf{x}}_{q,\ell}$ , system (32) can be rewritten as

$$\begin{bmatrix} \tilde{\mathbf{x}}_{q,\ell+1} \\ \zeta_{q,\ell+1} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & -\tilde{\mathbf{M}}_D^{-1} \tilde{\mathbf{M}}_F \\ \mathbf{0} & -\tilde{\mathbf{M}}_D^{-1} \tilde{\mathbf{M}}_F \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}}_{q,\ell} \\ \zeta_{q,\ell} \end{bmatrix}$$

which is neutrally stable if and only if condition (31) holds. ■

Recall that, in view of the *Dahlquist's Equivalence Theorem*, zero-stability is necessary and sufficient for convergence of a consistent time-discretization scheme [31]. Hence, under condition (31), the hybrid Euler time-discretization scheme (19) turns out to be convergent. For instance, this means that in each interval  $[t_q, t_{q+1}]$  the predicted estimates obtained via the Parallel Schwarz step (23) converge to the solution of a centralized prediction equation of the form

$$\tilde{\mathbf{M}} \dot{\tilde{\mathbf{x}}} + \tilde{\mathbf{S}} \tilde{\mathbf{x}} = \tilde{\mathbf{u}}$$

as the time-discretization step  $\delta$  goes to 0, or equivalently as the number  $L_q$  of distributed prediction steps goes to infinity.

*Remark 3:* It is worth noting that (31) translates into a block diagonal dominance condition for the global system, which requires that the effect of the isolated subsystems on the state

<sup>2</sup> The interested reader is referred to chapter 12 of [31] for an introduction on the concepts of consistency, zero-stability, and convergence of time-discretization methods.

evolution prevails over the effect originated from the interconnections among subsystems. Taking into account the particular structure of the FE mass matrix  $\mathbf{M}$ , which is reflected in the sparse structure of  $\tilde{\mathbf{M}}$ , the numerical stability condition (31) is usually satisfied in practice (see, for instance, the simulation example of Section VI). In addition, in the unlikely case in which condition (31) does not hold, it is possible to modify the hybrid Euler time-discretization scheme (29) (and hence the implementation of the Parallel Schwarz step) so as to retrieve zero-stability. Specifically, by introducing a suitable scalar  $\omega \in (0, 1]$ , one can replace (29) with

$$\begin{aligned} & \tilde{\mathbf{M}}_D \left( \frac{\tilde{\mathbf{x}}_{q,\ell+1} - (2 - \omega)\tilde{\mathbf{x}}_{q,\ell} + (1 - \omega)\tilde{\mathbf{x}}_{q,\ell-1}}{\omega \delta} \right) \\ & + \tilde{\mathbf{S}}_D \tilde{\mathbf{x}}_{q,\ell+1} + \tilde{\mathbf{M}}_F \left( \frac{\tilde{\mathbf{x}}_{q,\ell} - \tilde{\mathbf{x}}_{q,\ell-1}}{\delta} \right) + \tilde{\mathbf{S}}_F \tilde{\mathbf{x}}_{q,\ell} \\ & = \tilde{\mathbf{u}}_{q,\ell+1} + \tilde{\boldsymbol{\epsilon}}_{q,\ell+1} + \boldsymbol{\tau}_{q,\ell} \end{aligned} \quad (33)$$

which is still well-suited for distributed implementation. Notice that such a modified scheme coincides with (29) for  $\omega = 1$ . Further, along the lines of Theorem 1, it is possible to show that (33) is consistent for any value of  $\omega \in (0, 1]$ , and zero-stable provided that

$$\rho(\omega \tilde{\mathbf{M}}_D^{-1} \tilde{\mathbf{M}}_F - (1 - \omega) \mathbf{I}) < 1. \quad (34)$$

In turn, since

$$\rho(\omega \tilde{\mathbf{M}}_D^{-1} \tilde{\mathbf{M}}_F - (1 - \omega) \mathbf{I}) \leq \max\{\omega \rho(\tilde{\mathbf{M}}_D^{-1} \tilde{\mathbf{M}}_F), 1 - \omega\}$$

for any  $\omega \in (0, 1]$ , condition (34) can be always satisfied for suitably small values of  $\omega$  even when condition (31) does not hold. The price to be paid for the improved numerical stability is a slow-down of the information spread.

## V. STABILITY ANALYSIS

In this section, the stability of the estimation error dynamics resulting from application of the distributed finite-element Kalman filter of Algorithm 2 (Table II) is analyzed by supposing the measurement equation in each domain to be linear (as it happens when the sensors directly measure the target field like in (10)). Further, in order to simplify the notation, the interval  $t_{q+1} - t_q$  between consecutive measurements is supposed to be constant, so that in each sampling interval  $[t_q, t_{q+1})$  a fixed number  $L$  of distributed prediction steps is performed. With this respect, we make the following assumption.

A3) For each  $m \in \mathcal{N}$ , the local measurement function is linear, i.e.,  $\mathbf{h}^m(\mathbf{x}^m) = \mathbf{C}^m \mathbf{x}^m$ . Further, local observability holds in the sense that the pair  $((\mathbf{A}^m)^L, \mathbf{C}^m)$  is observable for any  $m \in \mathcal{N}$ .

A set of sensors ensuring local observability in each domain ensures also global observability (i.e., observability of the global state vector given all the measurements). However, the converse need not hold in that local observability requires a sufficient number of sensors to be present in each subdomain. Nevertheless, under global observability, the local observability condition can be satisfied by choosing each subdomain large enough so that a sufficient number of sensors is included inside.

Recalling that the matrices  $\mathbf{A}^m$  arise from space-time discretization of a PDE, some comments on how the local observability assumption A3 maps to the original continuous field are important. In this respect, while the relationship between observability of a continuous field and of its space-time discretization is far from trivial [32], [33], the following considerations can be made: i) from the practical point of view, unless the domain  $\Omega$  has a very specific form, the exact observability of the original PDE solution cannot be directly checked, and one invariably needs to resort to some numerical approximation scheme [32] like the one considered here; ii) on the other hand, it has been proved that, for a convergent discrete approximation scheme, the observability of the discrete numerical model is sufficient (and necessary) for the stability of the related field estimation process (see [32] for a formal statement of this property); iii) finally, it has been recently shown [33] that quantitative observability measures, defined in terms of suitable observability Gramians, carry over in a consistent way from the original PDE to its space-time discretization for any convergent numerical approximation scheme.

Let us first rewrite (29) into the state-space form

$$\begin{aligned} \tilde{\mathbf{x}}_{q,\ell+1} = & \underbrace{(\tilde{\mathbf{M}}_D + \delta \tilde{\mathbf{S}}_D)^{-1} \tilde{\mathbf{M}}_D}_{\tilde{\mathbf{A}}_D} \tilde{\mathbf{x}}_{q,\ell} \\ & + \underbrace{(\tilde{\mathbf{M}}_D + \delta \tilde{\mathbf{S}}_D)^{-1} (-\delta \tilde{\mathbf{S}}_F - \tilde{\mathbf{M}}_F)}_{\tilde{\mathbf{A}}_F} \tilde{\mathbf{x}}_{q,\ell} \\ & + \underbrace{(\tilde{\mathbf{M}}_D + \delta \tilde{\mathbf{S}}_D)^{-1} \tilde{\mathbf{M}}_F}_{\tilde{\mathbf{A}}_F} \tilde{\mathbf{x}}_{q,\ell-1} \\ & + \underbrace{(\tilde{\mathbf{M}}_D + \delta \tilde{\mathbf{S}}_D)^{-1} \delta}_{\tilde{\mathbf{B}}} \tilde{\mathbf{u}}_{q,\ell+1} + \tilde{\mathbf{w}}_{q,\ell} \end{aligned} \quad (35)$$

where, clearly,  $\tilde{\mathbf{A}}_D = \text{block-diag}(\mathbf{A}^1, \dots, \mathbf{A}^N)$  is the block diagonal matrix of state transition matrices, representing the  $N$  isolated subsystems.

Recalling that, in each interval  $[t_q, t_{q+1})$ , the recursion (35) is initialized with the initial conditions (30), it can be easily noticed that at the last distributed prediction step  $\ell = L$  one obtains

$$\tilde{\mathbf{x}}_{q,L} = \tilde{\mathbf{A}}_D^L \tilde{\mathbf{x}}_{q,0} + \tilde{\mathbf{A}}_{F,L} \tilde{\mathbf{x}}_{q,0} + \tilde{\mathbf{B}}_L \tilde{\mathbf{U}}_q + \tilde{\mathbf{D}}_L \tilde{\mathbf{W}}_q \quad (36)$$

where  $\tilde{\mathbf{U}}_q \triangleq \text{col}\{\mathbf{u}_{q,\ell}, \ell = 1, \dots, L\}$ ,  $\tilde{\mathbf{W}}_q \triangleq \text{col}\{\mathbf{w}_{q,\ell}, \ell = 1, \dots, L\}$  and  $\tilde{\mathbf{B}}_L, \tilde{\mathbf{D}}_L, \tilde{\mathbf{A}}_{F,L}$  are suitable matrices with the latter defining the interconnection couplings between subsystems. Noting that, by definition,  $\tilde{\mathbf{x}}_{q,L} = \tilde{\mathbf{x}}_{q+1,0} = \tilde{\mathbf{x}}(T_{q+1}\Delta)$ , the latter equation can be rewritten as

$$\tilde{\mathbf{x}}_{q+1} = \tilde{\mathbf{A}}_D^L \tilde{\mathbf{x}}_q + \tilde{\mathbf{A}}_{F,L} \tilde{\mathbf{x}}_q + \tilde{\mathbf{B}}_L \tilde{\mathbf{U}}_q + \tilde{\mathbf{D}}_L \tilde{\mathbf{W}}_q \quad (37)$$

where  $\tilde{\mathbf{x}}_q \triangleq \tilde{\mathbf{x}}(T_q \Delta)$ .

Similarly, application of step 3 of Algorithm 2 yields, at the last distributed prediction step  $\ell = L$ ,

$$\hat{\mathbf{x}}_{q,L} = \tilde{\mathbf{A}}_D^L \hat{\mathbf{x}}_{q,0} + \tilde{\mathbf{A}}_{F,L} \hat{\mathbf{x}}_{q,0} + \tilde{\mathbf{B}}_L \tilde{\mathbf{U}}_q. \quad (38)$$

where  $\hat{\mathbf{x}}_{q,\ell} \triangleq \text{col}\{\hat{\mathbf{x}}_{q,\ell}^m, m \in \mathcal{N}\}$ . Further, by defining  $\hat{\mathbf{x}}_{q|q} \triangleq \text{col}\{\hat{\mathbf{x}}_{q|q}^m, m \in \mathcal{N}\}$  and  $\hat{\mathbf{x}}_{q|q-1} \triangleq \text{col}\{\hat{\mathbf{x}}_{q|q-1}^m, m \in \mathcal{N}\}$ , the global correction step of Algorithm 2 at time  $t_{q+1}$  can be written as

$$\hat{\mathbf{x}}_{q+1|q+1} = \hat{\mathbf{x}}_{q+1|q} + \tilde{\mathbf{L}}_{q+1}(\tilde{\mathbf{y}}_{q+1} - \tilde{\mathbf{C}} \hat{\mathbf{x}}_{q+1|q}) \quad (39)$$

where  $\tilde{\mathbf{y}}_{q+1} \triangleq \text{col}\{\mathbf{y}_{q+1}^m, m \in \mathcal{N}\}$ ,  $\tilde{\mathbf{L}}_{q+1} = \text{block-diag}(\mathbf{L}_{q+1}^1, \dots, \mathbf{L}_{q+1}^N)$ , and  $\tilde{\mathbf{C}} \triangleq \text{col}\{\mathbf{C}^m, m \in \mathcal{N}\}$ .

Recalling that  $\hat{\mathbf{x}}_{q,L} = \hat{\mathbf{x}}_{q+1|q}$  and  $\hat{\mathbf{x}}_{q,0} = \hat{\mathbf{x}}_{q|q}$ , equations (38) and (39) can be easily combined so as to write  $\hat{\mathbf{x}}_{q+1|q+1}$  as a function of  $\hat{\mathbf{x}}_{q|q}$  and thus obtain a recursive expression for the global estimate. In addition, noting that the global output vector can be written as  $\tilde{\mathbf{y}}_{q+1} = \tilde{\mathbf{C}}\hat{\mathbf{x}}_{q+1} + \tilde{\mathbf{v}}_{q+1}$  with  $\tilde{\mathbf{v}}_{q+1} \triangleq \text{col}\{\mathbf{v}_{q+1}^m, m \in \mathcal{N}\}$ , we can also write a recursive expression for the dynamics of the global estimation error  $\tilde{\mathbf{e}}_q \triangleq \text{col}\{\tilde{\mathbf{x}}_q - \hat{\mathbf{x}}_{q|q}, m \in \mathcal{N}\}$ . Specifically, standard calculations yield

$$\tilde{\mathbf{e}}_{q+1} = (\mathbf{I} - \tilde{\mathbf{L}}_{q+1} \tilde{\mathbf{C}}) (\tilde{\mathbf{A}}_D^L + \tilde{\mathbf{A}}_{F,L}) \tilde{\mathbf{e}}_q + \tilde{\mathbf{v}}_q \quad (40)$$

where the term  $\tilde{\mathbf{v}}_q = (\mathbf{I} - \tilde{\mathbf{L}}_{q+1} \tilde{\mathbf{C}}) \tilde{\mathbf{D}}_L \tilde{\mathbf{W}}_q + \tilde{\mathbf{v}}_{q+1}$  accounts for the time-space discretization errors, for the measurement noise, and for all the other possible uncertainties.

As for the time evolution of the global covariance matrix  $\tilde{\mathbf{P}}_{q|q} \triangleq \text{block-diag}(\mathbf{P}_{q|q}^1, \dots, \mathbf{P}_{q|q}^N)$ , with similar reasoning as above it is an easy matter to see that application of Algorithm 2 leads to the following recursion

$$\begin{aligned} \tilde{\mathbf{P}}_{q+1|q+1} &= (\mathbf{I} - \tilde{\mathbf{L}}_{q+1} \tilde{\mathbf{C}}^T) \tilde{\mathbf{P}}_{q+1|q} \\ &\quad (\mathbf{I} - \tilde{\mathbf{L}}_{q+1} \tilde{\mathbf{C}}^T) \left[ \gamma^{2L} \tilde{\mathbf{A}}_D^L \tilde{\mathbf{P}}_{q|q} (\tilde{\mathbf{A}}_D^L)^T + \tilde{\mathbf{\Phi}} \right] \end{aligned} \quad (41)$$

where  $\tilde{\mathbf{\Phi}} \triangleq \sum_{i=0}^{L-1} \gamma^{2i} \tilde{\mathbf{A}}_D^i \tilde{\mathbf{Q}} (\tilde{\mathbf{A}}_D^i)^T$  and  $\tilde{\mathbf{Q}} \triangleq \text{block-diag}(\mathbf{Q}^1, \dots, \mathbf{Q}^N)$ .

The following stability result can now be stated.

**Theorem 2:** Let assumptions A1-A3 hold and let the matrices  $\tilde{\mathbf{Q}}$  and  $\tilde{\mathbf{R}} \triangleq \text{block-diag}(\mathbf{R}^1, \dots, \mathbf{R}^N)$  be positive definite. Then, the global covariance matrix asymptotically converges to the unique positive solution  $\tilde{\mathbf{P}}(\gamma)$  of the algebraic Riccati equation

$$[\tilde{\mathbf{P}}(\gamma)]^{-1} = \left[ \gamma^{2L} \tilde{\mathbf{A}}_D^L \tilde{\mathbf{P}}(\gamma) (\tilde{\mathbf{A}}_D^L)^T + \tilde{\mathbf{\Phi}} \right]^{-1} + \tilde{\mathbf{C}}^T \tilde{\mathbf{R}}^{-1} \tilde{\mathbf{C}},$$

and the global Kalman gain converges to the steady-state value

$$\begin{aligned} \tilde{\mathbf{L}}(\gamma) &= \left[ \gamma^{2L} \tilde{\mathbf{A}}_D^L \tilde{\mathbf{P}}(\gamma) (\tilde{\mathbf{A}}_D^L)^T + \tilde{\mathbf{\Phi}} \right] \tilde{\mathbf{C}}^T \\ &\quad \times \left\{ \tilde{\mathbf{C}} \left[ \gamma^{2L} \tilde{\mathbf{A}}_D^L \tilde{\mathbf{P}}(\gamma) (\tilde{\mathbf{A}}_D^L)^T + \tilde{\mathbf{\Phi}} \right] \tilde{\mathbf{C}}^T + \tilde{\mathbf{R}} \right\}^{-1}. \end{aligned} \quad (42)$$

Then, the dynamics (40) of the estimation error is exponentially stable if and only if

$$\rho \left\{ \left[ \mathbf{I} - \tilde{\mathbf{L}}(\gamma) \tilde{\mathbf{C}} \right] (\tilde{\mathbf{A}}_D^L + \tilde{\mathbf{A}}_{F,L}) \right\} < 1. \quad (43)$$

*Proof:* Notice first that assumption A2 implies observability of the pair  $(\tilde{\mathbf{A}}_D^L, \tilde{\mathbf{C}})$  which, as it can be easily verified through the PBH test, also implies observability of  $(\gamma^L \tilde{\mathbf{A}}_D^L, \tilde{\mathbf{C}})$  for any real  $\gamma > 0$ . Then, the convergence of  $\tilde{\mathbf{P}}_{q|q}$  to  $\tilde{\mathbf{P}}(\gamma) > 0$  follows from well known results on discrete-time Kalman filtering, since (41) is the standard Kalman filter covariance recursion for a linear system with state matrix  $\gamma^L \tilde{\mathbf{A}}_D^L$  and output matrix  $\tilde{\mathbf{C}}$ .

Further, it is immediate to see that the gain  $\tilde{\mathbf{L}}(\gamma)$  defined in (42) is the steady-state global Kalman gain associated with the steady-state covariance  $\tilde{\mathbf{P}}(\gamma)$ . Notice finally that the matrix  $(\mathbf{I} - \tilde{\mathbf{L}}_{q+1} \tilde{\mathbf{C}}) (\tilde{\mathbf{A}}_D^L + \tilde{\mathbf{A}}_{F,L})$ , which determines the dynamics of the estimation error, exponentially converges to  $[\mathbf{I} - \tilde{\mathbf{L}}(\gamma) \tilde{\mathbf{C}}] (\tilde{\mathbf{A}}_D^L + \tilde{\mathbf{A}}_{F,L})$ , so that the estimation error dynamics is exponentially stable if and only if  $[\mathbf{I} - \tilde{\mathbf{L}}(\gamma) \tilde{\mathbf{C}}] (\tilde{\mathbf{A}}_D^L + \tilde{\mathbf{A}}_{F,L})$  is Schur stable, i.e., if and only if condition (43) is satisfied. ■

In practice, the design of the proposed distributed finite-element Kalman filter requires the tuning of the scalar parameter  $\gamma$ . Specifically, for any given value of  $\gamma$  the stability of the filter can be readily checked by means of condition (43). Then, the tuning of  $\gamma$  can be performed numerically by finding, among the values of  $\gamma$  satisfying the stability condition (43), the one yielding the best estimation accuracy (see Fig. 9 in Section VI for an illustration of these ideas in a specific case study).

In order to understand the role played by the scalar  $\gamma$  in the satisfiability of condition (43) the following result is helpful.

**Proposition 1:** A sufficient condition for (43) to hold is that the scalar  $\gamma$  satisfies the relationship

$$\gamma^L > \left\| \mathbf{I} + (\tilde{\mathbf{A}}_D^L)^{-1} \tilde{\mathbf{A}}_{F,L} \right\|_{\tilde{\mathbf{P}}(\gamma)}, \quad (44)$$

where  $\|\cdot\|_{\mathbf{M}}$  denotes the matrix norm induced by the vector norm  $\|\mathbf{x}\|_{\mathbf{M}} \triangleq \sqrt{\mathbf{x}^T \mathbf{M} \mathbf{x}}$ .

*Proof:* With standard manipulations, it can be seen that  $\tilde{\mathbf{L}}(\gamma)$  and  $\tilde{\mathbf{P}}(\gamma)$  satisfy the relationship

$$\begin{aligned} \tilde{\mathbf{P}}(\gamma) &= [\mathbf{I} - \tilde{\mathbf{L}}(\gamma) \tilde{\mathbf{C}}^T] \left[ \gamma^{2L} \tilde{\mathbf{A}}_D^L \tilde{\mathbf{P}}(\gamma) (\tilde{\mathbf{A}}_D^L)^T + \tilde{\mathbf{\Phi}} \right] \\ &\quad \times [\mathbf{I} - \tilde{\mathbf{L}}(\gamma) \tilde{\mathbf{C}}^T]^T + \tilde{\mathbf{L}}(\gamma) \tilde{\mathbf{R}} [\tilde{\mathbf{L}}(\gamma)]^T \end{aligned}$$

so that

$$[\mathbf{I} - \tilde{\mathbf{L}}(\gamma) \tilde{\mathbf{C}}^T] \left[ \gamma^{2L} \tilde{\mathbf{A}}_D^L \tilde{\mathbf{P}}(\gamma) (\tilde{\mathbf{A}}_D^L)^T \right] [\mathbf{I} - \tilde{\mathbf{L}}(\gamma) \tilde{\mathbf{C}}^T]^T \leq \tilde{\mathbf{P}}(\gamma)$$

and, hence,

$$\left\| [\mathbf{I} - \tilde{\mathbf{L}}(\gamma) \tilde{\mathbf{C}}^T] \tilde{\mathbf{A}}_D^L \right\|_{\tilde{\mathbf{P}}(\gamma)} \leq 1/\gamma^L. \quad (45)$$

Hence, in order to complete the proof, it is sufficient to observe that

$$\begin{aligned} & \left\| \left[ \mathbf{I} - \tilde{\mathbf{L}}(\gamma)\tilde{\mathbf{C}} \right] \left( \tilde{\mathbf{A}}_D^L + \tilde{\mathbf{A}}_{F,L} \right) \right\|_{\tilde{\mathbf{P}}(\gamma)} \\ & \leq \left\| \left[ \mathbf{I} - \tilde{\mathbf{L}}(\gamma)\tilde{\mathbf{C}} \right] \tilde{\mathbf{A}}_D^L \right\|_{\tilde{\mathbf{P}}(\gamma)} \left\| \mathbf{I} + \left( \tilde{\mathbf{A}}_D^L \right)^{-1} \tilde{\mathbf{A}}_{F,L} \right\|_{\tilde{\mathbf{P}}(\gamma)} \\ & \leq \left\| \mathbf{I} + \left( \tilde{\mathbf{A}}_D^L \right)^{-1} \tilde{\mathbf{A}}_{F,L} \right\|_{\tilde{\mathbf{P}}(\gamma)} / \gamma^L \end{aligned}$$

where the latter inequality follows from (45). In fact, this implies  $\left\| \left[ \mathbf{I} - \tilde{\mathbf{L}}(\gamma)\tilde{\mathbf{C}} \right] \left( \tilde{\mathbf{A}}_D^L + \tilde{\mathbf{A}}_{F,L} \right) \right\|_{\tilde{\mathbf{P}}(\gamma)} < 1$  and hence (43) whenever (44) holds. ■

It can be seen from (44) that the smaller is  $\tilde{\mathbf{A}}_{F,L}$  (the part of the dynamics due to interaction between subdomains) as compared to  $\tilde{\mathbf{A}}_D^L$  (the local dynamics in the subdomains), the easier it becomes to achieve stability. In fact, in the limit case of no interaction ( $\tilde{\mathbf{A}}_{F,L} = 0$ ) the condition is satisfied for any  $\gamma > 1$ . In this respect, it is worth pointing out that the quantity  $\left( \tilde{\mathbf{A}}_D^L \right)^{-1} \tilde{\mathbf{A}}_{F,L}$  is usually small because of the structure of the FE matrices and the fact that the interactions are limited to the interfaces. For instance, in the case study of Section VI stability of the filter is guaranteed for a wide range of values of  $\gamma$ . Nevertheless, in general it is not possible to guarantee that a value of  $\gamma$  satisfying (44), or (43), always exists. This state of affairs can be understood by noting that in (44) both the left-hand and the right-hand side increase with  $\gamma$ . In case a suitable  $\gamma$  cannot be found, the stability of the filter can be guaranteed by resorting to a slight modification of the proposed approach which is summarized in the following procedure:

- 1) select the time interval  $\delta$  so that the dynamics of (37) is asymptotically stable;
- 2) pick any  $\gamma > 1$  (for example, by minimizing, the left-hand side of (43));
- 3) find a scalar  $\kappa > 0$  such that

$$\rho \left\{ \left[ \mathbf{I} - \kappa \tilde{\mathbf{L}}(\gamma)\tilde{\mathbf{C}} \right] \left( \tilde{\mathbf{A}}_D^L + \tilde{\mathbf{A}}_{F,L} \right) \right\} < 1; \quad (46)$$

- 4) modify the correction step (39) as follows

$$\hat{\mathbf{x}}_{q+1|q+1} = \hat{\mathbf{x}}_{q+1|q} + \kappa \tilde{\mathbf{L}}_{q+1} (\tilde{\mathbf{y}}_{q+1} - \tilde{\mathbf{C}} \hat{\mathbf{x}}_{q+1|q}). \quad (47)$$

Notice that the stability of (37), obtained from time-discretization of the asymptotically stable system (7), can be preserved by making  $\delta$  suitably small when the time-discretization scheme is zero-stable (a property which either holds when (31) is satisfied or can be enforced by means of the arrangements of Remark 1). Further, under stability of (37), condition (46) can be always satisfied as well for suitably small values of  $\kappa$ . The idea is that the gain of the local Kalman filters should not be too large so that stability is preserved. Hence, in the considered setting, the above-reported procedure is guaranteed to succeed. Finally, it is an easy matter to verify that, under condition (46), the distributed finite-element Kalman filter with the modified correction step (47) leads to an asymptotically stable estimation error dynamics (the proof is analogous to the one of Theorem 2).

As a final remark, we point out that, once the original filtering problem has been recast in the form (38)–(39), the problem of designing the filter gains falls within the wider framework of partition-based distributed Kalman filtering (see [34] and the reference therein for an insight on this problem). The proposed solution has the advantage of requiring the tuning of one (or few) scalar quantities and hence is well-suited to keeping the computational load manageable even when the state vector has a large dimension (as it usually happens in the context of field estimation). Further, the proposed approach requires that only the estimates pertaining to the interfaces are exchanged between neighboring nodes, thus keeping the communication requirements as low as possible.

## VI. SIMULATION EXPERIMENTS

This section provides numerical examples and relative results illustrating the effectiveness of the proposed distributed *finite element Kalman filter* presented in Section IV. Consider the transient *heat conduction* problem, introduced in Section II as a particular example of (1), in a thin polygonal metal plate with constant, homogeneous, and isotropic properties. Assuming that the thickness of the slab is considerably smaller than the planar dimensions, then the temperature can be assumed to be constant along the width direction, and the problem is reduced to two dimensions. Hence, the diffusion process in a thin plate is modelled by the 2D parabolic PDE  $\partial x / \partial t - \lambda (\partial^2 x / \partial \xi^2 + \partial^2 x / \partial \eta^2) = 0$  with boundary condition  $\mathcal{B}(x) = \alpha(\xi, \eta) \partial x / \partial \mathbf{n} + \beta(\xi, \eta) x$  such that  $\alpha(\xi, \eta) \beta(\xi, \eta) \geq 0$ ,  $\alpha(\xi, \eta) + \beta(\xi, \eta) > 0$ ,  $\forall (\xi, \eta) \in \partial\Omega$ . Notice that,  $x(\xi, \eta, t)$  denotes the temperature as a function of time  $t$  and spatial variables  $(\xi, \eta) \in \Omega$ ,  $f = 0$  stands for no inner heat-generation, whereas  $\lambda = 1.11 \times 10^{-4}$  [m<sup>2</sup>/s] is the *thermal diffusivity* of copper at 25 [°C] (Table XII, Appendix 2 in [35]), assumed to be constant in time and space.

A network of  $S = 23$  sensors (Fig. 2) located in the known positions  $\mathbf{s}_i = [\xi_i, \eta_i]^T$  is assumed to collect point temperature measurements at regularly time-spaced instants  $t_q = qT_s$ , with  $T_s = 100$  [s] and standard deviation of measurement noise  $\sigma_v = 0.1$  [K]. The considered sensor network has been chosen to guarantee local observability (assumption A2).

The Matlab PDE Toolbox is used to generate the triangular mesh (252 vertices, 436 elements) shown in Fig. 2 of size  $b = 0.2$  [m] (defined as the length of the longest edge of the element), over the global 2D domain  $\Omega$ . Next, as can be seen from Fig. 2, the domain under consideration is decomposed into  $N = 8$  overlapping subdomains  $\Omega_m$ , i.e.  $\mathcal{N} = \{1, \dots, 8\}$ , each being assigned to a *node* with local processing and communication capabilities. It is worth pointing out that domain decomposition comes with an appropriate partitioning of the original global mesh so that the resulting local grids actually match on the regions of overlap between subdomains.

Domain triangulation allows for a simple construction of basis functions  $\{\phi_j(\xi, \eta)\}_{j=1}^n$ , which are continuous piecewise polynomial functions, such that their value is unity in vertex  $j$

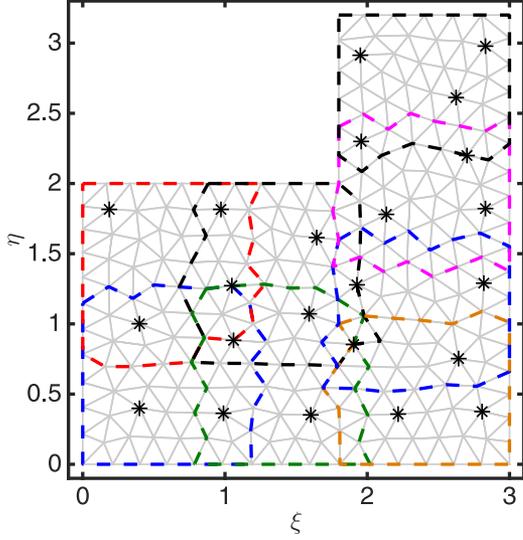


Fig. 2. Global FE mesh (grid of solid lines) generated over  $\Omega$  and domain decomposition into 8 overlapping subdomains (dashed polygons). The position of each sensor is denoted by  $*$ .

and vanishes at the remaining vertices, i.e.

$$\phi_j(\xi_i, \eta_i) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad i, j = 1, 2, \dots, n$$

Here we use continuous piecewise linear functions defined on each element as  $\psi_{\mathcal{E}}(\xi, \eta) = c_0 + c_1\xi + c_2\eta$  with  $(\xi, \eta) \in \mathcal{E}$  and  $c_0, c_1, c_2 \in \mathbb{R}$ , so that each function is uniquely determined by its three nodal values  $x_i = \psi_{\mathcal{E}}(\xi_i, \eta_i)$ ,  $i \in \mathcal{E}$ .

Basis functions are used off-line by the FE centralized filter and in the distributed setup for the element-by-element construction of matrices  $\mathbf{S}$  and  $\mathbf{M}$ , introduced in (6). Then, the state dynamics of the centralized filter can be directly computed, whereas local estimators first need to extract matrices  $\mathbf{M}^{mm}, \mathbf{S}^{mm}$  and  $\mathbf{M}^{mj}, \mathbf{S}^{mj}$  in order to calculate  $\tilde{\mathbf{A}}^m, \tilde{\mathbf{A}}^{mj}$  and  $\tilde{\mathbf{A}}^{mj}$  which finally provide the finite-dimensional model of temperature evolution in  $\Omega_m$  through (22). Notice that these matrices are evaluated for a fixed sampling interval  $\delta = T_s/L$ , where  $L$  denotes the number of distributed prediction iterations  $L_q$  introduced in Section IV, here assumed constant in each sampling interval  $q$ . For a fair comparison between centralized and distributed approaches, a constant time integration interval  $\Delta = 10$  [s] has been chosen for the centralized filter.

Notice that, being  $\{\phi_j(\xi, \eta)\}_{j=1}^n$  functions with a small support defined by the set of triangles sharing node  $j$ , the resulting *mass* and *stiffness* matrices will be sparse, with the same pattern shown in Fig. 3(a). In Fig. 3(b) it can be seen how the structure of the stiffness matrix changes when considering the augmented system (25). The distributed pattern of the networked system is highlighted in Fig. 4, where  $\tilde{\mathbf{A}}_D$  represents each subsystem as isolated, though affected by the evolution of neighbors through  $\tilde{\mathbf{A}}_F$ .

In the following experiments, both FE filters assume the initial temperature field of the plate uniform at  $x_0(\xi, \eta) = 300$  [K], and the a-priori estimate taken as first guess  $\hat{x}_{1|0}(\xi, \eta) = 305$  [K], with diagonal covariance  $\mathbf{P}_{1|0} = 20\mathbf{I}$ . Moreover, a zero-mean

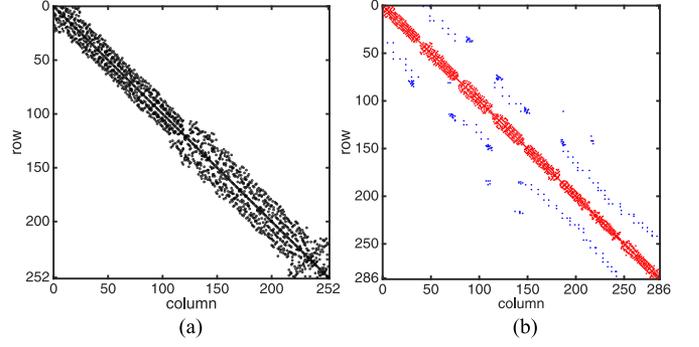


Fig. 3. Sparsity pattern of  $252 \times 252$  matrix  $\mathbf{S}$  (a), and  $286 \times 286$  matrix  $\tilde{\mathbf{S}} = \tilde{\mathbf{S}}_D + \tilde{\mathbf{S}}_F$  (b). (a)  $\mathbf{S}$ : 1626 nonzero elements, (b)  $\tilde{\mathbf{S}}_D$ : 1632 nonzero elements (red);  $\tilde{\mathbf{S}}_F$ : 223 nonzero elements (blue).

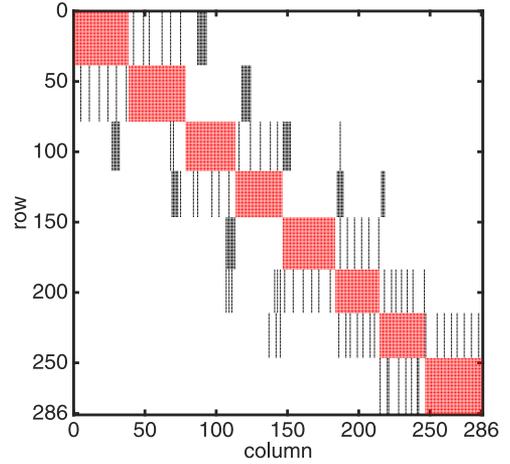


Fig. 4. Sparsity pattern of  $\tilde{\mathbf{A}}_D$  (red) and  $\tilde{\mathbf{A}}_F$  (black).

white noise process has been assumed, with covariance  $\mathbf{Q} = \sigma_w^2 \mathbf{I}$ , where  $\sigma_w = 3$  [K]. Taking into consideration model uncertainty, the *ground truth* of the experiments is represented by a real process simulator implementing a finer mesh (915 vertices, 1695 elements) of size  $b = 0.1$  instead of  $b = 0.2$ , running at a higher sampling rate (1 Hz), and aware of the possibly time-varying boundary conditions of the system. On the other hand, both distributed and centralized filters have no knowledge of the real system boundary conditions, so they simply assume the plate adiabatic on each side.

The performance of the novel distributed *FE Kalman filter* has been evaluated in terms of Root Mean Square Error (RMSE) of the estimated temperature field, averaged over a set of about 300 sampling points uniformly spread within the domain  $\Omega$ , and 500 independent Monte Carlo realizations.

## VI. Scenario 1

In the first example, transient analysis is performed on a thin adiabatic L-shaped plate (seen in Fig. 2) with a fixed temperature along the bottom edge. This is a problem with mixed boundary conditions, namely a non-homogeneous Dirichlet condition on the bottom edge of the plate  $\partial\Omega_1$ , i.e.

$$x = T_1 \quad \text{on } \partial\Omega_1, \quad (48)$$

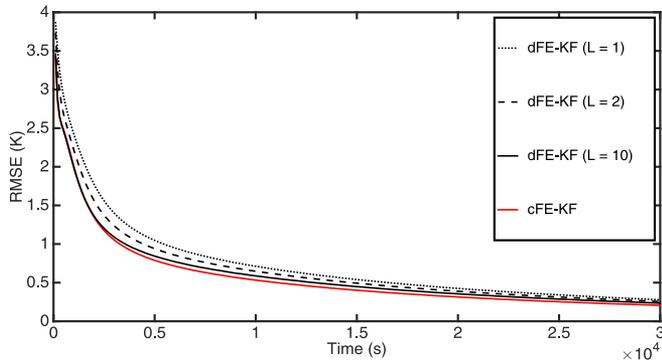


Fig. 5. Scenario 1: Comparison of performance of centralized and distributed FE-KF ( $\gamma = 1.1$ ).

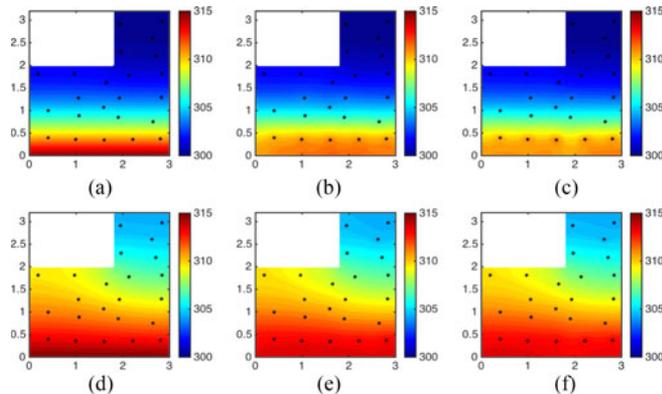


Fig. 6. Scenario 1: *True* (left) and estimated (cFE-KF, center – dFE-KF, right) temperature fields in Kelvin (K) at time steps  $q = 50$  (top) and  $q = 200$  (bottom).

where  $T_1 = 315$  [K], and natural homogeneous Neumann boundary conditions on the remaining insulated sides  $\partial\Omega_2 = \partial\Omega \setminus \partial\Omega_1$ , so that

$$\partial x / \partial \mathbf{n} = 0 \quad \text{on } \partial\Omega_2. \quad (49)$$

The duration of each Monte Carlo run is fixed to  $3 \times 10^4$  [s] (300 samples). Fig. 5 illustrates the performance comparison between centralized (cFE-KF) and distributed (dFE-KF) filters for  $\gamma = 1.1$  and for three different values of the parameter  $L$  adopted in the distributed framework. First of all, it can be seen that both FE algorithms succeed in reconstructing the *true* field of the system based on fixed, point-wise temperature observations. Moreover, the performance of the distributed FE filters is very close, even for  $L = 1$ , to that of the centralized filter, which collects all the data in a central node. Last but not least, in the distributed setting the RMSE behaviour improves by increasing the number  $L$  of distributed prediction steps. This is true for certain values of  $\gamma$ , whereas for others the difference in performance is considerably reduced, as clearly presented in Fig. 9. Note that the covariance boosting factor used in (24) is set to  $\gamma_L = \sqrt[L]{\gamma}$ ,  $\forall L = 1, 2, 10$ , in order to obtain a fairly comparable effect of covariance inflation after  $L$  distributed prediction steps for different distributed filters. Further insight on the performance of the proposed FE estimators is provided in Fig. 6,

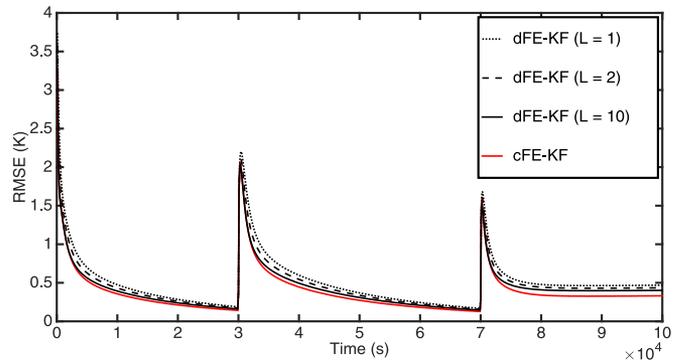


Fig. 7. Scenario 2: Comparison of performance of centralized and distributed FE-KF ( $\gamma = 1.1$ ).

which shows the *true* and estimated temperature fields at two different time steps  $q = 50$  and  $q = 200$ , obtained in a single Monte Carlo experiment by using cFE-KF and dFE-KF with  $L = 10$ .

## VI. Scenario 2

In the second experiment, two time-varying disturbances have been added in order to test the robustness of the proposed FE estimators in a more challenging scenario. To this end, different boundary conditions are considered. Specifically, a time-dependent Dirichlet condition (48) with  $T_1 = 310$  [K] for time steps  $q \in \{0, \dots, 299\}$ , and  $T_2 = 320$  [K] for  $q \in \{300, \dots, 1000\}$ , is set on all nodes of the bottom edge  $\partial\Omega_1$ . The top edge of the plate  $\partial\Omega_3$  is first assumed adiabatic for  $q \in \{0, \dots, 699\}$ , then the inhomogeneous Robin boundary condition

$$\lambda \partial x / \partial \mathbf{n} + \nu x = \nu x_e \quad \text{on } \partial\Omega_3 \quad (50)$$

is applied for  $q \in \{700, \dots, 1000\}$ . This models a sudden exposure of the surface to a fluid, fixed at an external temperature  $x_e = 300$  [K], through a uniform and constant convection heat transfer coefficient  $\nu = 10$  [W/m<sup>2</sup>K]. The remaining edges  $\partial\Omega_2$  where (49) holds, are assumed thermally insulated for the duration of the whole experiment, lasting  $10^5$  [s] (1000 samples).

Performance of the proposed distributed filter has been evaluated for different values of  $L$  over 500 independent Monte Carlo runs and compared to the behavior of the centralized FE Kalman filter. Simulation results, in Fig. 7, show that the proposed FE estimators provide comparable performance to the centralized filter, moreover the gap reduces as  $L$  increases. It is worth pointing out that the peaks appearing in the RMSE plot, displayed in Fig. 7, are due to the abrupt changes of the unknown boundary conditions, which cause considerable jumps of the estimation errors at time steps 300 and 700. Nevertheless, the filters under consideration manage to compensate for the lack of knowledge and effectively reduce the error, even if, due to persistent and cumulative disturbances on the inferred field profile, errors do not converge to zero. The original *ground truth* and the reconstructed fields are depicted in Fig. 8 for  $q = 350$  and  $q = 900$ .

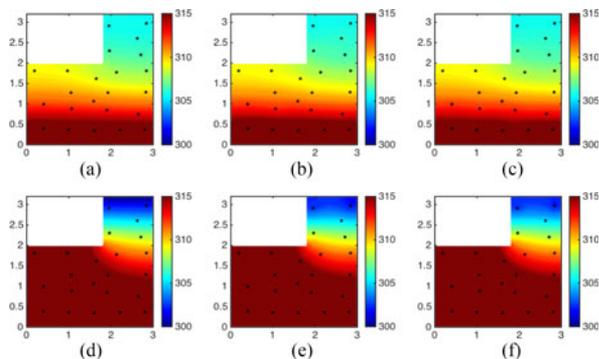


Fig. 8. Scenario 2: *True* (left) and estimated (cFE-KF, center – dFE-KF, right) temperature fields in Kelvin (K) at time steps  $q = 350$  (top) and  $q = 900$  (bottom).

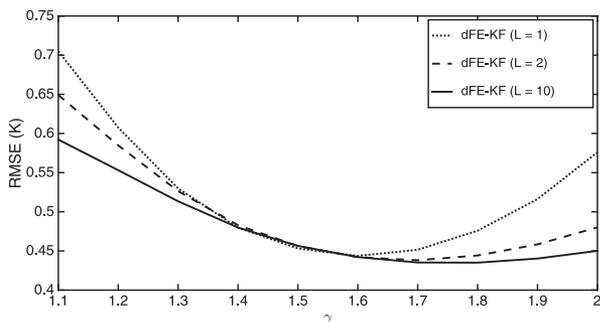


Fig. 9. Scenario 1: Comparison of the mean value of the RMSE for different values of  $\gamma$ .

## VII. CONCLUSION

The paper has dealt with the decentralized estimation of a time-evolving and space-dependent field governed by a linear partial differential equation, given point-in space measurements of multiple sensors deployed over the area of interest. The originally infinite-dimensional filtering problem has been approximated into a finite-dimensional large-scale one via the *finite element* method and, further, a distributed approach inspired by the parallel Schwarz method for domain decomposition has allowed to nicely scale the overall problem complexity with respect to the number of used processing nodes. Combining these two ingredients, a novel computationally efficient distributed finite-element Kalman filter has been proposed to solve in a decentralized and scalable fashion filtering problems involving distributed-parameter systems. Both numerical stability of the considered approximation scheme and exponential stability of the proposed distributed finite-element Kalman filter have been analysed. Simulation experiments have been presented in order to demonstrate the validity of the proposed approach.

The results of this work can be extended to the estimation of fields governed by more general partial differential equations and also be applied to the estimation/localization of diffusive sources.

## REFERENCES

[1] M. Fisher, J. Nocedal, Y. Tremolet and S. J. Wright, “Data assimilation in weather forecasting: A case study in PDE-constrained optimization,” *Optim. Eng.*, vol. 10, no. 3, pp. 409–426, 2009.

[2] J. de Halleux, C. Prieur, J. M. Coron, B. d’Andréa-Novel and G. Bastin, “Boundary feedback control in networks of open channels,” *Automatica*, vol. 39, no. 8, pp. 1365–1376, 2003.

[3] M. Ortner, A. Nehorai, and A. Jeremic, “Biochemical transport modeling and Bayesian source estimation in realistic environments,” *IEEE Trans. Signal Processing*, vol. 55, no. 6, pp. 2520–2532, 2007.

[4] S. Moura, J. Bendtsen, and V. Ruiz, “Observer design for boundary coupled PDEs: application to thermostatically controlled loads in smart grids,” *Proc. IEEE Conf. Decision Control*, pp. 6286–6291, Florence, Italy, 2013.

[5] C. G. Claudel and A. M. Bayen, “Lax-Hopf based incorporation of internal boundary conditions into Hamilton-Jacobi equation. Part I: Theory,” *IEEE Trans. Autom. Control*, vol. 55, no. 6, pp. 1142–1157, 2010.

[6] R. Olfati-Saber, J. A. Fax, and R. Murray, “Consensus and cooperation in networked multi-agent systems,” *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[7] L. Xiao, S. Boyd, and S. Lall, “A scheme for robust distributed sensor fusion based on average consensus,” *Proc. 4th Int. Symp. Inform. Processing Sensor Netw.*, pp. 63–70, Los Angeles, CA, 2005.

[8] G. C. Calafiore and F. Abrate, “Distributed linear estimation over sensor networks,” *Int. J. Control*, vol. 82, no. 5, pp. 868–882, 2009.

[9] G. Battistelli and L. Chisci, “Kullback-Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability,” *Automatica*, vol. 5, no. 3, pp. 707–718, 2014.

[10] G. Battistelli, L. Chisci, and C. Fantacci, “Parallel consensus on likelihoods and priors for networked nonlinear filtering,” *IEEE Signal Processing Lett.*, vol. 21, no. 7, pp. 787–791, 2014.

[11] G. Battistelli, L. Chisci, C. Fantacci, A. Farina, and A. Graziano, “Consensus CPHD filter for distributed multitarget tracking,” *IEEE J. Selected Topics Signal Processing*, vol. 7, no. 3, pp. 508–520, 2013.

[12] U. A. Khan and J. Moura, “Distributing the Kalman filter for large-scale systems,” *IEEE Trans. Signal Processing*, vol. 66, no. 10, pp. 4919–4935, 2008.

[13] S. Stankovic, M. S. Stankovic, and D. M. Stjepanovic, “Consensus based overlapping decentralized estimation with missing observations and communication faults,” *Automatica*, vol. 45, no. 6, pp. 1397–1406, 2009.

[14] M. Farina, G. Ferrari-Trecate, and R. Scattolini, “Moving-horizon partition-based state estimation of large-scale systems,” *Automatica*, vol. 46, no. 5, pp. 910–918, 2010.

[15] H. Zhang, J. Moura, and B. Krogh, “Dynamic field estimation using wireless sensor networks: Tradeoffs between estimation error and communication cost,” *IEEE Trans. Signal Processing*, vol. 57, no. 6, pp. 2383–2395, 2009.

[16] F. Dörfler, F. Pasqualetti, and F. Bullo, “Continuous-time distributed observers with discrete communication,” *IEEE J. Selected Topics Signal Processing*, vol. 7, no. 2, pp. 296–304, 2013.

[17] M. A. Demetriou, “Design of consensus and adaptive consensus filters for distributed parameter systems,” *Automatica*, vol. 46, no. 2, pp. 300–311, 2010.

[18] M. A. Demetriou, “Adaptive consensus filters of spatially distributed systems with limited connectivity,” *Proc., 52nd IEEE Conf. Decision Control*, pp. 442–447, Firenze, Italy, 2013.

[19] H. A. Schwarz, *Gesammelte mathematische abhandlungen. Band I, II*, AMS Chelsea Publishing, Bronx, NY, 1890.

[20] P. L. Lions, On the Schwarz alternating method. I, in *Proc. 1st Int. Symp. Domain Decomposition Methods Partial Differential Equations*, R. Glowinski *et al.*, SIAM, Philadelphia, PA, pp. 1–42, 1988.

[21] M. J. Gander, “Schwarz methods over the course of time,” *Electronic Trans. Numer. Anal.*, vol. 31, pp. 228–255, 2008.

[22] T. F. Chan and T. P. Mathew, “Domain decomposition algorithms,” *Acta Numerica*, vol. 3, pp. 61–143, 1994.

[23] A. Toselli and O. Widlund, *Domain Decomposition Methods—Algorithms and Theory*, Springer-Verlag, Berlin, Germany, 2005.

[24] B. Smith, P. Björstad, and W. Gropp, *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*, Cambridge University Press, New York, NY, 1996.

[25] G. Pelosi, R. Coccioli, and S. Selleri, *Quick Finite Elements for Electromagnetic Waves*, Artech House, Norwood, MA, 2009.

[26] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, NY, 1996.

[27] J.-F. Lee, J.-F., and Z. Sacks, “Whitney elements time domain (WETD) methods,” *IEEE Trans. Magnetics*, vol. 31, no. 3, pp. 1325–1329, 1995.

[28] R. Suga and M. Kawahara, “Estimation of tidal current using Kalman filter finite-element method,” *Comp. Math. Appl.*, vol. 52, no. 8–9, pp. 1289–1298, 2006.

- [29] Y. Ojima and M. Kawahara, "Estimation of river current using reduced Kalman filter finite element method," *Computer Methods Appl. Mech. Eng.*, vol. 198, no. 9–12, pp. 904–911, 2009.
- [30] G. Battistelli, L. Chisci, N. Forti, G. Pelosi and S. Selleri, "Distributed finite element Kalman filter," *Proc. Eur. Control Conf.*, pp. 3700–3705, Linz, Austria, 2015.
- [31] E. Süli and D. F. Mayers, *An Introduction to Numerical Analysis*, Cambridge University Press, Cambridge, UK, 2003.
- [32] S. E. Cohn and D. P. Dee, "Observability of discretized partial differential equation," *SIAM J. Numer. Anal.*, vol. 25 no. 3, pp. 586–617, 1988.
- [33] W. Kang, L. Xu, and F. Giraldo, "Partial observability and its consistency for PDEs," *arXiv:1401.0235v1*, 2014.
- [34] M. Farina and R. Carli, "Partition-based Distributed Kalman Filter with plug and play features," *arXiv:1507.06820*, 2015.
- [35] F. Kreith, R. M. Manglik, and M. S. Bohn, *Principles of Heat Transfer*, Springer-Verlag, New York, NY, 1996.



**Giorgio Battistelli** received the Laurea degree in electronic engineering and the Ph.D. degree in robotics from the University of Genoa, Genoa, Italy, in 2000 and 2004, respectively.

From 2004 to 2006, he was a Research Associate with the Dipartimento di Informatica, Sistemistica e Telematica, University of Genoa. Since 2006 he has been with the University of Florence, Florence, Italy, where he is currently an Associate Professor of automatic control with the Dipartimento di Ingegneria

dell'Informazione. He is currently an editor of the IFAC Journal Engineering Applications of Artificial Intelligence. His current research interests include adaptive and learning systems, real-time control reconfiguration, linear and nonlinear estimation, hybrid systems, sensor networks, and data fusion.

Dr. Battistelli is an Associate Editor of the IEEE Transactions on Neural Networks and Learning Systems, and a member of the conference editorial boards of IEEE Control Systems Society and the European Control Association.



**Luigi Chisci** (SM'13) was born in Florence, Italy, in 1959. He received the M.S. degree in electrical engineering from the University of Florence, Florence, Italy, in 1984 and the Ph.D. in systems engineering from the University of Bologna, Bologna, Italy, in 1989.

Currently he is a Full Professor of control engineering at the University of Florence since December 2004. His educational and research career have been in the area of control and systems engineering. His research interests have

spanned over adaptive control and signal processing, algorithms and architectures for real-time control and signal processing, recursive identification, filtering and estimation, predictive control. He has co-authored over 170 papers of which about 60 on international journals. His current interests concern networked estimation, multi-target multi-sensor tracking, multi-agent systems and sensor data fusion.

Dr. Chisci served on the Conference Editorial Board of the IEEE Control Systems Society as an Associate Editor, from 2000 to 2008.



**Nicola Forti** received the B.Sc. degree in mechanical engineering, the M.Sc. degree in automation engineering and the Ph.D. degree in information engineering from the University of Florence, Florence, Italy, in 2009, 2013 and 2016 respectively. He is currently a Postdoctoral Researcher in the Department of Information Engineering, University of Florence.

In 2015, he has been Visiting Research Scholar in the Department of Electrical and Computer Engineering at Carnegie Mellon University, Pittsburgh, PA for one year. His main research interests include networked estimation, data fusion, security of cyber-physical systems, and multi-sensor multi-target tracking.



**Giuseppe Pelosi** (M'8–SM'91–F'00) was born in Pisa, Italy, in 1952. He received the Laurea (Doctor) degree in physics (summa cum laude) from the University of Florence, Florence, Italy, in 1976.

He is currently with the Department of Information Engineering of the same university, where he is Full Professor of Electromagnetic Fields. He was a Visiting Scientist at McGill University, Montreal, Quebec, Canada, from 1993 to 1995, and a Professor at the University of Nice-

Sophie Antipolis, France, in 2001. He was coauthor of several scientific publications on the aforementioned topics, appeared in international refereed journals, and national and international conferences. He has been guest editor of several special issues of international journals: *IEEE Antennas and Propagation Magazine*, 2013 (with Y. Rahmat-Samii and J. L. Volakis); *IEEE Transactions on Antennas and Propagation*, 2001 (with V. Grikunov and J. L. Volakis); *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, 2000 (with P. Guillon and T. Itoh); *Electromagnetics*, 1998 (with J. L. Volakis); *Annales des Telecommunications* (with J. L. Bernard and P. Y. Ufimtsev), 1995; *COMPEL*, 1994 (with P. P. Silvester); *COMPEL*, 2002; *Alta Frequenza-Rivista di Elettronica*, 1992. He is also coauthor of three books: *Finite Elements for Wave Electromagnetics* (with P. P. Silvester, IEEE Press, 1994), *Finite Element Software for Microwave Engineering* (with T. Itoh and P. P. Silvester, Wiley, 1996), and *Quick Finite Elements for Electromagnetic Fields* (with R. Coccioli and S. Selleri, Artech House, 1998 and 2009). With P. P. Silvester (McGill University), he was promoter of the International Workshop on Finite Elements for Microwave Engineering. This workshop is held every two years, and is the meeting point for finite-element researchers from all over the world. He is also active in the history of the telecommunications engineering field. He has been Associate editor for the Historical Corner of the *IEEE Antennas and Propagation Magazine* since 2006. Among his latest publications in this field is the book, *A Wireless World. One Hundred Years Since the Nobel Prize to Guglielmo Marconi* (series contribution to the History of the Royal Swedish Academy of Sciences, 2012). His research activity is mainly focused on numerical techniques for applied electromagnetics (antennas, circuits, microwave and millimeter-wave devices, scattering problems).

Dr. Pelosi has been a member of the Board of Directors of the Applied Computational Electromagnetics Society (ACES) (1999–2001), of the Board of Directors of the IEEE Central and South Italy Section (1992–1995 and 1995–1998), and Chair of the IEEE Magnetics Chapter of the same Section (1996–1999).



**Stefano Selleri** (S'92–M'96–SM'03) was born in Viareggio, Italy, on December 9, 1968. He received the Laurea degree (with honors) in electronic engineering and the Ph.D. degree in computer science and telecommunications from the University of Florence, Florence, Italy, in 1992 and 1997, respectively.

In 1992 he was a Visiting Scholar at the University of Michigan, Ann Arbor, MI; in 1994 at the McGill University, Montreal, Canada; in 1997 at the Laboratoire d'Electronique of the University of Nice—Sophia Antipolis. From February to July 1998 he was a Research Engineer at the Centre National d'Etudes Telecommunications (CNET) France Telecom. He is currently an Associate Professor at the University of Florence, where he conducts research on numerical modelling of microwave, devices and circuits with particular attention to numerical optimization. Since September 2006 he is Associate editor of the International Journal of Antennas and Propagation.